

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A TRAVÉS DE  
NARRATIVA ESCRITA**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA INGENIERÍA DE SISTEMAS Y  
COMPUTACIÓN  
BOGOTÁ, COLOMBIA  
2021**

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA IDENTIFICACIÓN DE  
SÍNTOMAS DE DEPRESIÓN A TRAVÉS DE NARRATIVA ESCRITA**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**DOCUMENTO PRESENTADO COMO REQUISITO PARA OPTAR AL TÍTULO DE:  
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

**DIRECTOR:**

JUAN CARLOS BARRERO CALIXTO

**LÍNEA DE INVESTIGACIÓN:**

SOFTWARE INTELIGENTE Y CONVERGENCIA TECNOLÓGICA

**GRUPO DE INVESTIGACIÓN:**

GISIC - SEMILLERO SMART

UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA, INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ, COLOMBIA

2021



## Atribución-NoComercial 4.0 Internacional (CC BY-NC 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Advertencia](#).

### Usted es libre de:

**Compartir** — copiar y redistribuir el material en cualquier medio o formato

**Adaptar** — remezclar, transformar y construir a partir del material

La licenciente no puede revocar estas libertades en tanto usted siga los términos de la licencia

---

### Bajo los siguientes términos:



**Atribución** — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciente.



**NoComercial** — Usted no puede hacer uso del material con [propósitos comerciales](#).

**No hay restricciones adicionales** — No puede aplicar términos legales ni [medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia](#).

**NOTA DE ACEPTACIÓN**

---

---

---

---

**FIRMA DIRECTOR DEL PROYECTO**

---

**FIRMA JURADO**

---

**FIRMA JURADO**

---



## **AGRADECIMIENTOS**

Muchas son las personas que han contribuido al presente trabajo directa o indirectamente, es correcto agradecer a Dios principalmente, seguido de agradecer al profesor Juan Carlos Barrero Calixto nuestro director de tesis por su dedicación, aportes, asesoría, gentileza y apoyo, que permitieron que este proyecto tuviera las bases científicas y teóricas necesarias para cumplir todos los objetivos y sortear los obstáculos a lo largo del camino.

Gracias a cada una de las personas que fueron parte de este proyecto, por acompañarnos y aportar con sus valiosos consejos que se ven reflejados a lo largo de este proyecto y en nuestras vidas.

A nuestras familias, por darnos la fuerza imprescindible para poder cumplir nuestras metas.

A la Universidad Católica de Colombia, por brindarnos los espacios educativos para ampliar nuestros conocimientos facilitando nuestro desarrollo como ingenieros, pero más importante, como personas.

Y sobre todo un mensaje especial a todas aquellas personas que sufren o han sufrido de algún problema de Salud Mental ya sea depresión u otro, decirles que no están solos, que este trabajo es una prueba de que el mundo científico quiere darles una mano, apoyarlos y colaborar para que salgamos adelante juntos.

## CONTENIDO

<b>LISTA DE FIGURAS</b>	<b>10</b>
<b>LISTA DE TABLAS</b>	<b>12</b>
<b>LISTA DE ABREVIATURAS</b>	<b>14</b>
<b>GLOSARIO</b>	<b>15</b>
<b>RESUMEN</b>	<b>16</b>
<b>ABSTRACT</b>	<b>17</b>
<b>INTRODUCCIÓN</b>	<b>18</b>
<b>1. GENERALIDADES</b>	<b>20</b>
1.1 LÍNEA DE INVESTIGACIÓN	20
1.2 PLANTEAMIENTO DEL PROBLEMA	20
1.3 ALCANCES Y LIMITACIONES	23
1.4 JUSTIFICACIÓN	24
1.5 OBJETIVOS	26
1.5.1 OBJETIVO GENERAL	26
1.5.2 OBJETIVOS ESPECÍFICOS	26
1.6 MARCOS DE REFERENCIA	27
1.6.1 MARCO CONCEPTUAL	27
1.6.1.1 MACHINE LEARNING	27
1.6.1.2 GOOGLE COLAB	27
1.6.1.3 COMPRESIÓN DE LENGUAJE NATURAL (NLU)	27
1.6.1.4 DEPRESIÓN	28
1.6.1.5 MÉTODO AUTOMÁTICO	28
1.6.1.6 ANÁLISIS DE SENTIMIENTOS	28
1.6.1.7 TRASTORNO	28
1.6.1.8 ONTOLOGÍA	29
1.6.1.9 PROCESAMIENTO DE LENGUAJE NATURAL (PLN)	29
1.6.1.10 LENGUAJE DE PROGRAMACIÓN PYTHON	30
1.6.1.11 ALGORITMO	30
1.6.1.12 RSTUDIO	30
1.6.1.13 SCIPY	30

1.6.1.14 REQUESTS	31
1.6.1.15 BEAUTIFULSOUP	31
1.6.1.16 DEPLACY	31
1.6.1.17 ES_CORE_NEWS_SM	31
1.6.1.18 PKG_RESOURCES	32
1.6.1.19 IMP	32
1.6.1.20 IMP.RELOAD	32
1.6.1.21 COUNTER	32
1.6.1.22 IMPORT FILES	32
1.6.1.23 OPENPYXL	33
1.6.1.24 READ	33
1.6.1.25 SPLIT	33
1.6.1.26 OPENPYXL.LOAD_WORKBOOK	33
1.6.1.27 ITER_ROWS	34
1.6.1.28 CHUNK.TEXT	34
1.6.1.29 MOST_COMMON	34
1.6.1.30 RE.FINDITER	34
1.6.2 MARCO TEÓRICO	35
1.6.2.1 RECUENTO DE INVESTIGACIONES Y DESARROLLOS RELACIONADOS	35
1.6.2.2 TEORÍA DE APRENDIZAJE DE MÁQUINA	36
1.6.2.3 APRENDIZAJE SUPERVISADO	36
1.6.2.4 APRENDIZAJE NO SUPERVISADO	44
1.6.1.5 REDES NEURONALES	45
1.6.1.6 MATRIZ DE CONFUSIÓN	46
1.6.1.7 MÉTRICA: precisión	47
1.6.1.8 MÉTRICA: RECALL	48
1.6.1.9 MÉTRICA: PUNTUACIÓN F1	49
1.6.1.10 MÉTRICA: EXACTITUD	49
1.6.1.11 PROTÉGÉ	50
1.6.1.12 ANÁLISIS DE TEXTO	50
1.6.1.13 ONTOLOGÍA EN SISTEMAS	52
1.7 ESTADO DEL ARTE	53
1.8 METODOLOGÍA	60
1.8.1 FASE 1: CONSTRUCCIÓN DE LA ONTOLOGÍA	61
1.8.2 FASE 2: DISEÑO DE LENGUAJE NATURAL Y DEL MÉTODO AUTOMÁTICO	63
1.8.3 FASE 3: EMPLEO DEL MÉTODO AUTOMÁTICO	64

1.8.4 FASE 4: VALIDAR EL MÉTODO AUTOMÁTICO	66
1.8.5 REDISEÑAR - AFINAR	67
<b>2. ONTOLOGÍA APLICADA A SINTOMATOLOGÍA DE DEPRESIÓN</b>	<b>68</b>
2.1 DETERMINAR EL DOMINIO Y ÁMBITO.	69
2.2 DETERMINAR LA INTENCIÓN DE USO.	73
2.3 REUTILIZAR ONTOLOGÍAS O VOCABULARIOS EXISTENTES.	74
2.4 ENUMERAR LOS TÉRMINOS IMPORTANTES DEL DOMINIO.	76
2.5 DEFINIR JERARQUÍA DE CLASES.	77
2.6 CREAR LAS INSTANCIAS.	79
2.7 VALIDAR CON JUICIO DE EXPERTOS EN LAS ÁREAS DE PSICOLOGÍA E INGENIERÍA	82
2.7.1 VALIDACIÓN AUTÓNOMAS	82
2.7.2 VALIDACIÓN DE EXPERTOS	84
<b>3. DISEÑO DEL MÉTODO AUTOMÁTICO</b>	<b>87</b>
3.1 ANÁLISIS DE INFORMACIÓN	87
3.2 ESPECIFICACIONES PARA EL MÉTODO AUTOMÁTICO ESPERADO	93
3.2.1 DESCRIPCIÓN GENERAL	93
3.2.2 REQUERIMIENTOS Y NECESIDADES DEL MÉTODO AUTOMÁTICO DADOS POR EL ÁREA DE SALUD MENTAL	94
3.2.3 PERSPECTIVA DEL PRODUCTO	98
3.3 PROCESOS FUNDAMENTALES	98
3.4 PSEUDOCÓDIGOS DEL MÉTODO AUTOMÁTICO	104
<b>4. IMPLEMENTACIÓN Y DESARROLLO DEL MÉTODO AUTOMÁTICO</b>	<b>106</b>
4.1 FASE 1. DESARROLLO DEL FORMULARIO DE PREGUNTAS	107
4.2 FASE 2. PROCESO DE ENTENDIMIENTO Y ANÁLISIS DE INFORMACIÓN	109
4.3 FASE 3. CORRELACIÓN DE INFORMACIÓN CON ONTOLOGÍA DESARROLLADA	122
4.4 FASE 4. GENERACIÓN DE INFORME	125
<b>5. VALIDACIÓN DEL MÉTODO AUTOMÁTICO</b>	<b>127</b>
5.1 HERRAMIENTAS UTILIZADAS	128
5.2 PROCESO DE LA HERRAMIENTA.	130
5.3 MEDIDAS DE RENDIMIENTO OBTENIDAS	132
5.3.1 Etiqueta Actividad Física	132
5.3.2 Etiqueta Alimentación	133
5.3.3 Etiqueta de Atención Dispersa	134
5.3.4 Etiqueta de Cargo de Conciencia	135

5.3.5 Etiqueta de Desaliento	136
5.3.6 Etiqueta de Descanso	137
5.3.7 Etiqueta Disminución de interés o placer	138
5.3.8 Etiqueta De Estado De Ánimo Deprimido	139
5.3.9 Etiqueta De Ideas Suicidas	140
5.4 TABLA DE RESULTADOS DE LAS MÉTRICAS	141
<b>6. RESULTADOS Y ANÁLISIS DE RESULTADOS</b>	<b>143</b>
<b>7. CONCLUSIONES</b>	<b>149</b>
<b>8. TRABAJOS FUTUROS</b>	<b>151</b>
<b>BIBLIOGRAFÍA</b>	<b>152</b>

## LISTA DE FIGURAS

- Figura 1. Regresión Lineal Simple.
- Figura 2. Regresión Logística.
- Figura 3. Modelo Knn.
- Figura 4. Árbol de decisión.
- Figura 5. Árbol de decisión Aleatorio.
- Figura 6. Modelo Red Neuronal.
- Figura 7. Matriz de Confusión.
- Figura 8. Funcionamiento tokenización.
- Figura 9. Diagrama general de metodología.
- Figura 10. Diagrama metodológico fase 1: Construcción de la ontología.
- Figura 11. Diagrama metodológico fase 2: Diseño de lenguaje natural y del método automático.
- Figura 12. Diagrama metodológico fase 3: Empleo del método automático.
- Figura 13. Procedimiento de machine learning para análisis de textos.
- Figura 14. Diagrama metodológico fase 4: Validar el método automático.
- Figura 15. Rediseño y afinación de metodología.
- Figura 16. Captura de pantalla reuniones remotas realizadas.
- Figura 17. Gráfico de la ontología exportada de Protégé.
- Figura 18. Gráfico de la ontología con herramienta de visualización especializada.
- Figura 19. Captura del Log generado por Reasoner-Protégé.
- Figura 20. Captura del mensaje de aprobación generado por OntoDebug.
- Figura 21. Diagrama general de método automático esperado.
- Figura 22. Diagrama extracción de la información (ETL).
- Figura 23. Diagrama de limpieza de información
- Figura 24. Diagrama de Análisis del texto.
- Figura 25. Diagrama de Métricas de desempeño Fuente.
- Figura 26. Diagrama de diseño general del método automático.
- Figura 27. Vista del usuario formulario.
- Figura 28. Vista del usuario y macro realizada.
- Figura 29. Script de macro en Google Sheets.
- Figura 30. Vista del usuario.
- Figura 31. Script de la función "Activar librerías".
- Figura 32. Vista del usuario proceso extracción de datos.
- Figura 33. Script proceso extracción de datos.
- Figura 34. Vista del usuario carga de archivo de respuestas.

Figura 35. Vista del usuario validación carga de archivo de respuestas.  
Figura 36. Script carga de archivo de respuestas.  
Figura 37. Vista del usuario selección de código único ID.  
Figura 38. Script selección de código único ID.  
Figura 39. Vista del usuario preparación de información para su análisis.  
Figura 40. Script preparación de información para su análisis.  
Figura 41. Vista del usuario análisis generados.  
Figura 42. Script genérico análisis generados.  
Figura 43. Vista del usuario repetición de un concepto específico.  
Figura 44. Script repetición de un concepto específico.  
Figura 45. Vista del usuario Carga de Ontología.  
Figura 46. Script Carga de Ontología.  
Figura 47. Vista del usuario relación de la ontología con palabras generadas por el consultante.  
Figura 48. Script relación de la ontología con palabras generadas por el consultante.  
Figura 49. Vista del usuario Generación informe de resultados.  
Figura 50. Script Generación informe de resultados.  
Figura 51. Script Implementación de librería MonkeyLearn.  
Figura 52. Implementación de librerías en cuadernillo Google Colab 1.  
Figura 53. Implementación de librerías en cuadernillo Google Colab 2.  
Figura 54. Implementación de librerías en cuadernillo Google Colab 3.  
Figura 55. Proceso de construcción de métricas 1.  
Figura 56. Proceso de construcción de métricas 2.  
Figura 57. Etiqueta Actividad Física.  
Figura 58. Etiqueta Alimentación.  
Figura 59. Etiqueta de Atención Dispersa.  
Figura 60. Etiqueta de Cargo de Conciencia.  
Figura 61. Etiqueta de Desaliento.  
Figura 62. Etiqueta de Descanso.  
Figura 63. Etiqueta Disminución de interés o placer.  
Figura 64. Etiqueta De Estado De Ánimo deprimido.  
Figura 65. Etiqueta De Ideas Suicidas.  
Figura 66. Total de conceptos por síntoma estipulados en la Ontología.  
Figura 67. Manual de uso para el Método automático en apoyo a la identificación de sintomatología depresiva.

## LISTA DE TABLAS

- Tabla 1. Número de personas atendidas por depresión en Colombia de 2009 a 2015.
- Tabla 2. Lista de preguntas generadas.
- Tabla 3. Ontologías y vocabularios existentes que apoyan en la construcción.
- Tabla 4. Enumeración y clasificación de conceptos.
- Tabla 5. Jerarquía de clases e instancias trabajadas en Protégé.
- Tabla 6. Jerarquía de propiedades del objeto trabajadas en Protégé.
- Tabla 7. Métrica de validación ontología.
- Tabla 8. Métrica de validación ontología agrupada.
- Tabla 9. Selección de técnicas y tecnologías - Librerías
- Tabla 10. Selección de técnicas y tecnologías - Funciones
- Tabla 11. Selección de técnicas y tecnologías - Algoritmos y transformaciones
- Tabla 12. Tabla de agrupamiento de requerimientos y necesidades generadas.
- Tabla 13. Tabla de requerimiento, Identificar información del consultante mediante un código único (ID).
- Tabla 14. Tabla de Etiquetas POS universales para el modelo de lenguaje natural
- Tabla 15. Tabla de Resultados de las Métricas.



## **LISTA DE ANEXOS**

- Anexo 1. Consentimiento Informado
- Anexo 2. Estado del arte del proyecto.
- Anexo 3. Preguntas Cuestionario Escrito Ajustadas después de validación.
- Anexo 4. Enumeración y clasificación de conceptos.
- Anexo 5. Ontología sintomatología depresiva.
- Anexo 6. Mapa Ontología.
- Anexo 7. Log generado por Reasoner-Protégé.
- Anexo 8. Métrica de validación para ontología.
- Anexo 9. Requerimientos y necesidades del método automático.
- Anexo 10. Pseudocodigos.
- Anexo 11. Diseño general método automático.
- Anexo 12. Preguntas para identificar tu estado de ánimo.
- Anexo 13. Informe Generado.
- Anexo 14. Manual de uso para el Método automático en apoyo a la identificación de sintomatología depresiva.
- Anexo 15. Método Automático Sintomatología Depresiva.

## LISTA DE ABREVIATURAS

**OMS:** Organización Mundial de la Salud

**NIH:** National Institutes of Health

**SISPRO:** Sistema Integrado de Información de la Protección Social

**PLN:** Procesamiento de Lenguaje Natural

**NLU:** Natural Language understanding (Comprensión de lenguaje natural)

## GLOSARIO

**Método automático:** Es un producto de software el cual constituye una metodología la cual de forma automática lograr un objetivo definido.

**Ontología:** Una ontología es una definición formal de tipos, propiedades, y relaciones entre entidades que existen para un dominio de discurso en particular.

**Protégé:** Protégé es una plataforma que proporciona un conjunto de herramientas para construir modelos de dominio y aplicaciones basadas en ontologías.

**Depresión:** La depresión es un trastorno mental caracterizado por un estado de ánimo bajo, disminución del interés, sentimientos de inutilidad y disminución de la capacidad de concentración.

**DSM-V:** Según el DSM-V, el trastorno depresivo mayor se diagnostica a partir de la presencia de mínimo 5 síntomas y el deterioro en las áreas de ajuste.

**Google Colaboratory:** Es un entorno Jupyter Notebook destinado a la capacitación e investigación en aprendizaje automático. Esta plataforma permite entrenar modelos de machine learning de manera directa en la nube.

## RESUMEN

La OMS registra más de 300 millones de personas con depresión a nivel mundial, esta misma evidencia la ausencia de mecanismos para la identificación de sus síntomas primarios, lo que dificulta la tarea de control y prevención. Para Colombia los casos presentes de depresión han ido en incremento desde el año 2009. Para el año 2015 se reporta un total de casos de depresión del 4,7 % de la población nacional, y en el 2020 en la pandemia se estima que el 35% de los adultos jóvenes han presentado depresión.

Aunque existen diferentes pruebas psicológicas para el diagnóstico de la depresión, hay pocos mecanismos para la identificación temprana de sintomatología depresiva y particularmente mecanismos que apoyen esta labor con soporte en la tecnología, por lo que surge la pregunta de investigación: ¿Qué elementos deben tenerse en cuenta para desarrollar una herramienta tecnológica que apoye en la identificación de síntomas de depresión, en narrativas escritas de una persona?. Para resolver este problema se planteó el diseño, desarrollo e implementación de un método automático que clasifica en los 9 síntomas depresivos las narrativas escritas dadas por un consultante. Este trabajo inició a partir de la creación de un equipo interdisciplinario con el que se construyó una ontología para la sintomatología depresiva, un formulario de preguntas abiertas para obtener las narrativas escritas de los consultantes, y en conjunto se construyó un dataset el cual entrenó al modelo de predicción implementado en el método automático.

Ante el aumento de la depresión a nivel mundial es fundamental el desarrollo de herramientas disruptivas para apoyen en la detección y disminución de este trastorno, que tiene en sus consecuencias el suicidio con un reporte de 800.000 personas fallecidas por año. El resultado final de este proyecto de grado es la consolidación de un método automático que se compone por una serie de procedimientos para el manejo de datos y que con la integración de una ontología con más de 684 palabras posibilita la identificación y detección de posible sintomatología depresiva, aportando a los procesos de detección y tratamiento de la depresión utilizados por profesionales en psicología y salud mental.

**PALABRAS CLAVE:** Salud Mental, Ontología, Análisis Automático De Textos.

## **ABSTRACT**

The WHO registers more than 300 million people with depression worldwide, this evidence shows the absence of mechanisms for the identification of its primary symptoms, which hinders the task of control and prevention. For Colombia, the present cases of depression have been increasing since 2009. For the year 2015 a total of 4.7% of the national population reported cases of depression, and in 2020 in the pandemic it is estimated that 35% of young adults have presented depression.

Although there are different psychological tests for the diagnosis of depression, there are few mechanisms for the early identification of depressive symptomatology and particularly mechanisms that support this work with the support of technology, so the research question arises: What elements should be taken into account to develop a technological tool that supports the identification of symptoms of depression in written narratives of a person? To solve this problem, we proposed the design, development and implementation of an automatic method that classifies the 9 depressive symptoms in the written narratives given by a consultant. This work started with the creation of an interdisciplinary team with which an ontology for depressive symptomatology was built, a form of open questions to obtain the written narratives of the consultants, and together a dataset was built which trained the prediction model implemented in the automatic method.

Given the increase of depression worldwide, it is essential to develop disruptive tools to support the detection and reduction of this disorder, whose consequences include suicide, with a reported 800,000 deaths per year. The final result of this degree project is the consolidation of an automatic method that consists of a series of procedures for data management and that with the integration of an ontology with more than 684 words, enables the identification and detection of possible depressive symptomatology, contributing to the processes of detection and treatment of depression used by professionals in psychology and mental health.

**KEY WORDS:** Mental Health, Ontology, Automatic Text Analysis

## INTRODUCCIÓN

La depresión según el instituto nacional de salud mental (NIH) - EE.UU. por sus siglas en inglés, la entiende y trata como un trastorno del estado de ánimo que causa síntomas de angustia, y llega a afectar el plano cognitivo y emotivo de las personas en cómo se siente, piensa y coordina las actividades diarias, además de afectaciones en él como dormir, comer o trabajar. Para recibir un diagnóstico de depresión, los síntomas deben estar presentes la mayor parte del día, casi todos los días, durante por lo menos dos semanas.<sup>1</sup>

La presencia de la depresión en la persona afecta los procesos de la vida cotidiana y suelen causar afecciones físicas tales como: Problemas orgánicos, cansancio, problemas para dormir, pérdida de apetito, comer compulsivamente, ganar o perder peso, dolor de estómago, retortijones o estreñimiento, dolor de cabeza, ritmo cardíaco acelerado y afectaciones psicológicas como: Suicidio, angustia, apatía, susceptibilidad, baja autoestima, obsesiones, huir de la realidad, amargura, mal humor, pérdida cognitiva.<sup>2</sup> Incluso se puede llegar a violencias contra sí mismos (autolesiones) ó contra otros y en última instancia al suicidio, que para el año 2019 se reporta un total de 800.000 personas fallecidas por año.<sup>3</sup>

Según datos de la organización mundial de la salud (OMS) señalan que la depresión, es un mal que afecta al 4,4% de la población mundial, se calcula que afecta a más de 300 millones de personas. También en sus estudios se encontró que en América Latina, Brasil es el país con mayor prevalencia de depresión, con 5,8%, seguido de Cuba (5,5%) y Paraguay (5,2%). Chile y Uruguay tienen el 5%; Perú, 4,8% mientras que Argentina, Costa Rica y República Dominicana muestran, igual que Colombia, un 4,7%, de identificación de depresión en su población.<sup>4</sup>

Dados estos porcentajes, datos y estadísticas, la depresión evidentemente es un tema que requiere de atención. Ante esto, el presente trabajo de grado pretende mediante el uso de técnicas y algoritmos de aprendizaje de máquina construir un método automático para asistir a los profesionales que tratan el trastorno de la depresión que cada vez es más común entre niños, jóvenes y adultos jóvenes.

---

<sup>1</sup> NIMH “Depresión: Información Básica”. {En línea}. {3 de septiembre de 2020} disponible en: (<https://www.nimh.nih.gov/health/publications/espanol/depresion-sp/index.shtml>)

<sup>2</sup> MARIA DEL CARMEN BRAVO GARCÍA “¿Cuáles son las consecuencias de la depresión?” {En línea}. {3 de septiembre de 2020} disponible en: (<https://blog.cognifit.com/es/consecuencias-depresion/>)

<sup>3</sup> Suicidio. (2019, 2 septiembre). who. {En línea}. {01 de mayo de 2021} disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/suicide>

<sup>4</sup> PERIÓDICO EL TIEMPO - SALUD “Depresión en Colombia es más alta que el promedio en el mundo” {En línea}. {5 de septiembre de 2020} disponible en: (<https://www.eltiempo.com/salud/cifras-sobre-depresion-en-colombia-y-en-el-mundo-segun-la-oms-61454>)

Este proyecto se enmarca dentro de la modalidad de trabajo de grado: trabajo de investigación, la metodología que se implementa tiene un enfoque investigativo y se enfoca en construir un método automático que apoye a los profesionales de la salud en el diagnóstico y clasificación de pacientes que posiblemente padecen del trastorno de depresión, por medio de narrativas escritas de sus pacientes específicamente adultos jóvenes. Vale aclarar que el método automático que se propone en ningún momento se considera como un reemplazo de los procesos clínicos y psicológicos, más bien, se constituye como un aporte metodológico y procedimental a la labor de los profesionales debidamente capacitados para determinar este tipo de trastornos.

# 1. GENERALIDADES

## 1.1 LÍNEA DE INVESTIGACIÓN

**Campo de investigación:** Gestión y Tecnología al Servicio de la Sociedad.

**Grupo de Investigación:** Semillero de Investigación en Desarrollo de Videojuegos, Big Data y Data Analytics - SMART, adscrito al Grupo de investigación en software inteligente y convergencia tecnológica GISIC.

## 1.2 PLANTEAMIENTO DEL PROBLEMA

La organización mundial de la salud (OMS) define la depresión como “un trastorno mental frecuente, que se caracteriza por la presencia de tristeza, pérdida de interés o placer, sentimientos de culpa o falta de autoestima, trastornos del sueño o del apetito, sensación de cansancio y falta de concentración”<sup>5</sup>. Para la OMS este trastorno mental se encuentra en la lista de los trastornos con más intenciones de tratamiento a nivel mundial, para su control, sin embargo, la ausencia de mecanismos para la identificación de sus síntomas primarios, hacen la tarea de control y prevención de la depresión más complicada, por lo menos, en Colombia.

En el año 2017 con motivo de la conmemoración del Día Mundial de la Salud, la OMS publicó una campaña sobre la depresión llamada “hablemos” que pretendía generar una oportunidad única para movilizar la acción en torno a un tema de salud específico que preocupe a las personas de todo el mundo.<sup>6</sup> Durante todo este año fue una de las campañas con más participación y audiencia en el mundo, incluyendo a América Latina que es en donde los trastornos mentales y neurológicos representan el 22% de la carga total de enfermedades. La depresión se encuentra en este grupo y ha sido el diagnóstico más frecuente en los servicios de atención primaria en salud y según datos de la OMS, se encuentra entre los primeros cinco trastornos que generan mayor discapacidad a largo plazo y dependencia emocional a la condición depresiva misma.<sup>7</sup>

---

<sup>5</sup> MINSALUD - COLOMBIA “Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles” {En línea}. {10 de septiembre de 2020} disponible en:

(<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)

<sup>6</sup> OMS “Campañas mundiales de salud pública de la OMS- Aspectos esenciales de la campaña Día Mundial de la Salud 2017” {En línea}. {10 de septiembre de 2020} disponible en:

(<https://www.who.int/campaigns/world-health-day/2017/campaign-essentials/es/>)

<sup>7</sup> MINSALUD - COLOMBIA - DEPRESIÓN “Depresión #hablemos” {En línea}. {15 de septiembre de 2020} disponible en:

(<https://www.minsalud.gov.co/salud/publica/SMental/Paginas/depresion.aspx>)



Para Colombia los casos presentes de depresión han ido en incremento desde el año 2009 según el Boletín de salud mental Depresión Subdirección de enfermedades No Transmisibles publicado por el Sistema Integral de Información de la Protección Social (SISPRO) en el año 2017 y respaldado por el reciente estudio de la Universidad Javeriana 2020 en donde por causas de la pandemia se estima que el 35% de los adultos jóvenes colombianos han presentado depresión.<sup>8</sup>

En el boletín de salud mental publicado por SISPRO es importante resaltar que los casos de depresión identificados en la sociedad Colombiana, presentan un aumento considerable en solamente seis años. (ver Tabla 1)<sup>9</sup>

### Tabla 1

*Número de personas atendidas por depresión en Colombia (CIE10 F321, F322, F323), de 2009 a 2015.*

		Año					
2009	2010	2011	2012	2013	2014	2015	
<b>Depresión Moderada</b>							
17.711	18.534	21.633	26.860	28.457	38.638	36.584	
<b>Depresión grave sin síntoma psicóticos</b>							
4.703	4.801	6.647	6.499	6.879	20.111	8.385	
<b>Depresión grave con síntoma psicóticos</b>							
3.311	3.100	3.963	4.755	2.546	3.137	3.131	

*Tabla 1. Número de personas atendidas por depresión en Colombia de 2009 a 2015*<sup>10</sup>

*Nota:* Esta tabla se realiza a través de la elaboración propia a partir de los Registros Individuales de Prestación de Servicios (RIPS), en la bodega de datos del SISPRO, consultado el 19 febrero 2017.

<sup>8</sup> Ansiedad y depresión, 'verdugos' de los jóvenes en la pandemia. (2021, 12 abril). Revista Pesquisa Javeriana. {En línea}. {01 de mayo de 2021} disponible en: <https://www.javeriana.edu.co/pesquisa/ansiedad-y-depresion-verdugos-de-los-jovenes-durante-la-pandemia>

<sup>9</sup> MINSALUD - COLOMBIA "Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles" {En línea}. {15 de septiembre de 2020} disponible en: (<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)

<sup>10</sup> Hospital San Juan Capestrano "Signos y síntomas de la depresión" {En línea}. {20 de septiembre de 2020} disponible en: (<https://www.sanjuancapestrano.com/depresion/sintomas-efectos/> )

Como se observa en la tabla 1, el aumento de casos de personas que empiezan a presentar síntomas de depresión, son de bastante preocupación, dado que no se cuenta con una diversidad de mecanismos para la identificación de síntomas de depresión primarios. La ausencia de mecanismos para la identificación de síntomas de depresión, son un causante de que los casos de depresión en la sociedad Colombiana no puedan ser tratados a tiempo, para evitar su avance en la persona, lo que conlleva a diferentes consecuencias de la depresión como lo pueden ser el aislamiento social, presentar bajo rendimiento en diferentes actividades, problemas cardiovasculares y demás padecimientos secundarios, hasta llevar al punto de que la persona considere el suicidio o autoagresiones.<sup>11</sup>

Existen diferentes pruebas psicológicas para la identificación de síntomas depresivos como los son, BDI - II (Inventario de depresión de Beck-II), BDI (Inventario de depresión infantil), CDS (Cuestionario de depresión infantil), CET-DE (Cuestionario estructural tetradimensional para la depresión), entre otros. Este tipo de pruebas, mediante la formulación de preguntas y su resolución escrita, permiten medir la gravedad de la depresión en infantes, adolescentes y adultos jóvenes, también, llevar a cabo una evaluación global y específica de la depresión, además, de un diagnóstico y clasificación del humor depresivo, vaciamiento energético, discomunicación y ritmopatía, que se pueden presentar en la persona.<sup>12</sup>

La importancia de identificar síntomas de depresión radica en que este es un trastorno mental grave, frecuente en personas de diferentes contextos, edades, géneros y situaciones. Para el caso de Colombia existen diferentes tipos de tratamientos para analizar y tratar este trastorno. Sin embargo, su accesibilidad es limitada, lo que genera una complejización en la tarea de realizar un seguimiento constante en prevención, detección y tratamiento de la depresión. Es por esto, que al implementar herramientas tecnológicas que permitan hacer un primer análisis de detección de síntomas de depresión, se obtiene una manera más efectiva y rápida de ayudar a los procesos de diagnóstico y cuidado basándonos en el aprendizaje de máquinas e Inteligencia artificial. Esto sería para los psicólogos una herramienta de apoyo en sus futuras consultas, con lo que obtendrán un primer análisis más rápido y confiable antes de iniciar a profundidad en el tratamiento.

---

<sup>11</sup> Hospital San Juan Capestrano “Signos y síntomas de la depresión” {En línea}. {20 de septiembre de 2020} disponible en:

(<https://www.sanjuancapestrano.com/depresion/sintomas-efectos/> )

<sup>12</sup> Pontificia Universidad Javeriana - Cali “Catálogo de pruebas psicológicas” {En línea}. {20 de septiembre de 2020} disponible en:

([https://www2.javerianacali.edu.co/sites/ujc/files/catalogo\\_liip\\_2018.pdf](https://www2.javerianacali.edu.co/sites/ujc/files/catalogo_liip_2018.pdf))

Ante el problema planteado anteriormente, la pregunta de investigación es: ¿Qué elementos deben tenerse en cuenta para desarrollar una herramienta tecnológica que apoye en la identificación de síntomas de depresión, en narrativas escritas de una persona?

### **1.3 ALCANCES Y LIMITACIONES**

#### **ALCANCES**

Se desarrollará un método automático que permita identificar en primer momento los posibles síntomas de depresión en una persona, gracias a una narrativa escrita que la persona desarrolle, donde se tendrá en cuenta los principales conceptos claves a tratar para la identificación y señalización del mismo, con un algoritmo de aprendizaje de máquina, que etiquete y seleccione los términos claves a tratar para así dar un informe preliminar al psicólogo, dicho método automático se desarrollará en el lenguaje de programación Python y el uso de Google colab durante los periodos académico de 2020-II y 2021- I.

Este método automático para la identificación de sintomatología depresiva por medio de narrativas escritas, hace parte de un proyecto Institucional de la Universidad Católica de Colombia, llamado “Fase: Diseño De Un Software De Apoyo Para La Detección De Sintomatología Emocional En Narrativas Textuales”, trabajado por los departamentos y facultades de Ingeniería y Psicológica, este proyecto es guiado por los docentes e investigadores Eliana Ivette Ortiz Garzón, Psicóloga y Juan Carlos Barrero Calixto, Docente del área de ingeniería de Sistemas. Además de esto, el presente trabajo se realiza en paralelo con el proyecto “Método automático para la identificación de sintomatología depresiva por medio de narrativas orales”, realizado por los estudiantes Josué David Vergara Rojas y Daniel Gustavo Forero Chaux, estudiantes del programa de ingeniería de sistemas y computación en la Universidad Católica de Colombia.

#### **LIMITACIONES**

Para el proyecto a realizar solo se tendrá en cuenta la previa identificación de posibles síntomas de depresión que pueden aquejar a las personas que acuden al departamento clínico psicológico de la Universidad Católica de Colombia. Las personas que asistan a participar de estudios con el método automático desarrollado, se les solicitará la aprobación del consentimiento y las normas de la ley de habeas data para el manejo de su información. También se debe tener en cuenta una limitante en el desarrollo del proyecto, dado que no se presenta una construcción de software, con un ciclo tradicional de desarrollo. En este proyecto se presenta el diseño, desarrollo, implementación y validación de un experimento.

## 1.4 JUSTIFICACIÓN

De acuerdo con diferentes datos de la organización mundial de la salud (OMS), la depresión no es solo un trastorno psicológico, sino que es causante de otras enfermedades tales como la diabetes y enfermedades del corazón. El sector poblacional en que suele concentrarse mayormente la depresión es la juventud, puesto que durante esta época se es proclive a cambios, conflictos físicos y psicológicos que muchas veces resultan dramáticos<sup>13</sup>. Así mismo, la detección temprana de la depresión en los adultos jóvenes de edades 15 y 29 años<sup>14</sup>, impulsa la efectividad de diversos tratamientos.

De esta forma, en la actualidad se cuenta con herramientas tecnológicas para detectar la depresión, como son encuestas presenciales y online, que se constituyen como un nuevo medio que permite detectar indicios de depresión en el análisis de un texto escrito largo. Con estas herramientas se pretende analizar el estado de ánimo o sentimientos de las personas consultadas. Así pues, se decidió enfocar este trabajo en el análisis de los textos largos sobre el tema de la depresión, para con ello, generar un aporte tecnológico que ayude en el proceso de detección y diagnóstico temprano de la depresión en adultos jóvenes.

Un proceso tecnológico que permita la posible identificación de síntomas de depresión combina diferentes campos de estudio dentro de la ingeniería de sistemas, tales como, *minería de datos*, *algoritmos de aprendizaje de máquina* y *procesamiento de lenguaje natural*, debido a que el lenguaje humano es complejo de enseñar a las máquinas. Así mismo, en estos procesos informáticos es necesario un aprendizaje supervisado por profesionales del área de psicología, ingeniería y lenguaje.<sup>15</sup> De esto resulta que el trabajo propuesto apoyaría la detección temprana de síntomas depresivos y generaría una acción oportuna por parte de los profesionales del área de salud mental, por lo que este proyecto se presenta como una propuesta tecnológica que apoye a todas las personas que necesiten este tipo de consultas y que no cuenten con los recursos para tal fin.

El proyecto propuesto aporta certeza y confiabilidad en la herramienta de detección y diagnóstico del trastorno depresivo, dado que estos dos aspectos son

---

<sup>13</sup> Krauskopof, Dina. El desarrollo psicológico en la adolescencia. {En línea}. 1999. {17 septiembre de 2020}. Disponible en:

([http://www.scielo.sa.cr/scielo.php?script=sci\\_arttext&pid=S1409-41851999000200004](http://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1409-41851999000200004))

<sup>14</sup> Ministerio de salud {En línea}. {27 de octubre de 2020} disponible en:

(<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)

<sup>15</sup> Héctor G. Barnes” {En línea}. {27 de agosto de 2020} disponible en:

([https://www.elconfidencial.com/alma-corazon-vida/2015-08-28/las-razones-por-las-que-cada-vez-cometemos-mas-errores-ortograficos-y-gramaticales\\_988271/](https://www.elconfidencial.com/alma-corazon-vida/2015-08-28/las-razones-por-las-que-cada-vez-cometemos-mas-errores-ortograficos-y-gramaticales_988271/))

claves para procesos clínicos mayores. Esto se anota puesto que existen diferentes tipos de test hallados en internet, que no están lo suficientemente acreditados, lo cual genera desconfianza e incertidumbre en sus diagnósticos y representan un alto riesgo para las personas que los utilizan, pues existe una alta probabilidad de tener falsos resultados. Por ejemplo, se puede realizar un test que tenga un resultado positivo, mientras que otro test arroje resultado negativo, dando así resultados inciertos y contradictorios.<sup>16</sup>

Por lo tanto, un método automático gratuito que apoye en los procesos de prevención y detección de sintomatología depresiva y que, a su vez, esté avalado por profesionales de salud mental, será primordial para brindar un apoyo y obtener resultados más precisos que mejoren los tratamientos estándar, generando un impacto emocional y económico favorable para los consultantes, además de constituirse como un beneficio laboral, social y tecnológico para todos los profesionales del área de salud mental que lo implementen en sus consultas médicas.

---

<sup>16</sup> Paul Koeck {En línea} {27 de agosto de 2020} “¿Test de depresión? ¿Qué implica un buen test de depresión?” disponible en: (<https://www.mis15minutos.com/depresion/test-de-depresion-que-implica-un-buen-test-de-depresion/>)

## **1.5 OBJETIVOS**

### **1.5.1 OBJETIVO GENERAL**

Desarrollar un método automático, utilizando algoritmos de aprendizaje de máquina, procesamiento de lenguaje natural y una ontología, para apoyar a la identificación de síntomas depresivos en narrativas escritas de adultos jóvenes.

### **1.5.2 OBJETIVOS ESPECÍFICOS**

1. Construir una ontología respecto a síntomas de depresión, a partir de un estado del arte y de las necesidades de los profesionales del área de la psicología clínica, para usarse como insumo en el método automático.
2. Diseñar el método automático, a partir de técnicas de aprendizaje de máquina, para la identificación de sintomatología depresiva en narrativa escrita.
3. Implementar el método automático, para el análisis de textos, aplicando técnicas de análisis de sentimientos y algoritmos de aprendizaje de máquina.
4. Validar el método automático, utilizando las métricas de desempeño: precisión, Recall (“Recordar”), F1-score (“Puntuación F1”) y exactitud, para medir la funcionalidad y aplicabilidad del modelo.

## **1.6 MARCOS DE REFERENCIA**

### **1.6.1 MARCO CONCEPTUAL**

Este marco corresponde a todos los conceptos empleados para el desarrollo del proyecto.

#### **1.6.1.1 MACHINE LEARNING**

El término machine learning consiste en una ciencia de la informática enfocada en el área de inteligencia artificial, la cual es útil en crear sistemas que aprendan por ellos mismos.

Por tanto es una tecnología la cual permite realizar procesos automáticos a una serie de operaciones, lo cual es una gran ventaja al momento de controlar una gran cantidad de información de manera más efectiva.<sup>17</sup>

#### **1.6.1.2 GOOGLE COLAB**

Google colab es un servicio en la nube el cual está basado en jupyter notebook, permite el uso gratuito de GPUs y TPUs de Google, tiene librerías las cuales están enfocadas en machine learning y deep learning, todo esto bajo Python.<sup>18</sup>

La principal ventaja que ofrece esta herramienta es que libera la máquina de tener que llevar el trabajo que por lo general puede llevar demasiado tiempo, y resultar costoso en potencia, permite realizar trabajos incluso si nuestra máquina no tiene los requisitos necesarios para completar una tarea es necesario resaltar que es una herramienta gratuita, también es colaborativo, por tanto permite realizar tareas en la nube y compartir nuestros códigos para trabajar en equipo.

#### **1.6.1.3 COMPRESIÓN DE LENGUAJE NATURAL (NLU)**

La comprensión del lenguaje natural (NLU, por sus siglas en inglés) es una rama de la inteligencia artificial (AI) que utiliza programas informáticos para comprender las entradas hechas en forma de oraciones en formato de texto o habla. NLU utiliza algoritmos para reducir el habla humana a una ontología estructurada. El

---

<sup>17</sup> Jose Martínez Heras - precisión, Recall, F1, Accuracy en clasificación {En línea}. {18 de agosto de 2020} disponible en: (<https://www.iartificial.net/precisión-recall-f1-accuracy-en-clasificacion/>)

<sup>18</sup> Datahack - INTRODUCCIÓN A GOOGLE COLAB PARA DATA SCIENCE {En línea}. {25 de agosto de 2020} disponible en: (<https://www.datahack.es/google-colab-para-data-science/#:~:text=Colab%20es%20un%20servicio%20cloud.disponible%20para%20R%20y%20Scala.>)

objetivo principal detrás de NLU es crear bots habilitados para chat y voz que puedan interactuar de manera efectiva con el público sin supervisión.<sup>19</sup>

#### **1.6.1.4 DEPRESIÓN**

La depresión es el hecho en el cual se sienten síntomas de tristeza, melancolía, se podría describir como estar en un estado abatido o derrumbado, de cierta forma todos los seres humanos son susceptibles a sentirse de esa forma en algún momento de su vida. La depresión puede llegar a un nivel clínico, siendo visto como un trastorno del estado de ánimo, en el cual se experimentan sentimientos de ira, pérdida y frustración durante un periodo de algunas semanas o más.<sup>20</sup>

#### **1.6.1.5 MÉTODO AUTOMÁTICO**

En relación con el presente proyecto un método automático se refiere a la construcción de una metodología la cual de forma automática logre cierto objetivo en este caso la identificación de síntomas de depresión a través de narrativas escritas.<sup>21</sup> El método automático como su nombre lo indica funciona de modo automático en donde se ejecuta una serie de procedimientos por los cuales se realiza un análisis en las narrativas textuales de consultantes que posiblemente presenten síntomas de depresión para identificar estos de manera detallada.

#### **1.6.1.6 ANÁLISIS DE SENTIMIENTOS**

Es también conocido como minería de opinión, consiste en una clasificación masiva de manera automática, se centra en catalogar documentos, textos, dando una connotación, que en este caso la connotación será si tiene depresión o no.<sup>22</sup>

#### **1.6.1.7 TRASTORNO**

Un trastorno es una alteración en el funcionamiento del organismo o una parte de él, puede ser el equilibrio psíquico o mental de una persona, además, en el campo de la medicina es una alteración que afecta física y mentalmente, tiene diversas causas, entre ellas factores genéticos, traumas o enfermedades, de

---

<sup>19</sup> SAS - What is natural language processing nlp {En línea}. {18 de agosto de 2020} disponible en: ([https://www.sas.com/es\\_co/insights/analytics/what-is-natural-language-processing-nlp.html](https://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.html))

<sup>20</sup> Organización Mundial De la Salud - Depresión {En línea}. {18 de agosto de 2020} disponible en: (<https://www.who.int/es/news-room/fact-sheets/detail/depression>)

<sup>21</sup> Pablo Turmero-Métodos automáticos de análisis {En línea}. {19 de agosto de 2020} disponible en: (<https://www.monografias.com/trabajos109/metodos-automaticos-analisis/metodos-automaticos-analisis.shtml>)

<sup>22</sup> Manuel Delgado Tenorio - ¿Qué es el análisis del sentimiento? {En línea}. {19 de agosto de 2020} disponible en: (<https://manueldelgado.com/que-es-el-analisis-del-sentimiento/>)



manera psicológica es un desequilibrio mental, llamados trastornos psicológicos, tienen tratamientos con especialistas, en algunas ocasiones es congénito.<sup>23</sup>

### 1.6.1.8 ONTOLOGÍA

En ciencias de la computación, se refiere a una definición formal de tipos, propiedades y relaciones entre entidades que fundamentalmente existen para un dominio de discurso en particular, siendo una especie de aplicación práctica de la ontología filosófica.<sup>24</sup>

### 1.6.1.9 PROCESAMIENTO DE LENGUAJE NATURAL (PLN)

Es un campo de las ciencias de la computación, inteligencia artificial y lingüística que estudia las interacciones entre las computadoras y el lenguaje humano. El PLN se ocupa de la formulación e investigación de mecanismos eficaces computacionalmente para la comunicación entre personas y máquinas por medio del lenguaje natural, es decir, de las lenguas del mundo. El PLN no trata de la comunicación por medio de lenguas naturales de una forma abstracta, sino de diseñar mecanismos para comunicarse que sean eficaces computacionalmente que se puedan realizar por medio de programas que ejecuten o simulen la comunicación, adicionalmente, permite a las computadoras entender el lenguaje humano, este tipo de lenguaje se divide principalmente de manera escrita y oral, por tanto el PLN permite la manipulación y tratamiento de datos de manera avanzada, para lograr esto, tiene diferentes tipos de uso<sup>25</sup>, los cuales son:

- *Análisis morfológico o léxico*: Se trata de una distinción interna de las palabras que componen una oración .
- *Análisis sintáctico*: Se enfoca en la estructura de la oración basado en un modelo gramatical empleado.
- *Análisis semántico*: Este análisis se centra en la interpretación de las oraciones.
- *Análisis pragmático*: En este análisis incorpora los filtros realizados a las oraciones para dar una interpretación final a un texto especializado.

---

<sup>23</sup> García Higuera - Trastornos psicológicos {En línea}. {25 de agosto de 2020} disponible en: (<http://www.psicoterapeutas.com/Trastornos.html>)

<sup>24</sup> Natalya F. Noy and Deborah L. McGuinness "Ontology Development 101: A Guide to Creating Your First Ontology". {En línea}. {01 de 2020} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).

<sup>25</sup> Alba Sanz - Procesamiento del Lenguaje Natural (PLN o NLP): qué es y para qué se utiliza {En línea}. {25 de agosto de 2020} disponible en: (<https://decidesoluciones.es/procesamiento-del-lenguaje-natural-pln-o-nlp-que-es-y-para-que-se-utiliza/>)

Estos tipo de análisis se utilizan según el objetivo que se va a analizar debido a que cada análisis da un resultado diferente que puede ser utilizado para diferentes fines.

#### **1.6.1.10 LENGUAJE DE PROGRAMACIÓN PYTHON**

Es un lenguaje de programación interpretado cuya filosofía hace hincapié en la legibilidad de su código. Se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, dinámico y multiplataforma.<sup>26</sup>

#### **1.6.1.11 ALGORITMO**

Es una serie ordenada de instrucciones, pasos o procesos que llevan a la solución de un determinado problema.<sup>27</sup>

#### **1.6.1.12 RSTUDIO**

Es un software diseñado para realizar análisis estadísticos, como también gráficos, es un software libre, permite prototipar modelos de manera rápida y adicionalmente con buenos resultados.<sup>28</sup>

#### **1.6.1.13 SCIPY**

Es una librería de Python, compuesta de herramientas y algoritmos matemáticos, tales como optimización, álgebra lineal, integración, interpolación, funciones especiales entre otros, es una herramienta muy utilizada en el procesamiento de lenguaje natural <sup>29</sup>

---

<sup>26</sup> Miguel Angel Álvarez - Qué es Python {En línea}. {25 de agosto de 2020} disponible en: (<https://desarrolloweb.com/articulos/1325.php>)

<sup>27</sup> Mintic - Colombia aprende: ¿QUÉ ES UN ALGORITMO?{En línea}. {25 de agosto de 2020} disponible en: ([http://aprende.colombiaaprende.edu.co/sites/default/files/naspublic/curriculos\\_ex/n1g10\\_fproy/nive11/programacion/unidad1/leccion1.html](http://aprende.colombiaaprende.edu.co/sites/default/files/naspublic/curriculos_ex/n1g10_fproy/nive11/programacion/unidad1/leccion1.html))

<sup>28</sup> Rafael González - R para visualización {En línea}. {25 de agosto de 2020} disponible en: (<https://gonzalezgouveia.com/que-es-r-y-rstudio/>)

<sup>29</sup> UniPython {En línea}. {10 de enero de 2021} disponible en: (<https://uniPython.com/scipy-funciones-principales/>)

#### 1.6.1.14 REQUESTS

Esta librería permite enviar solicitudes HTTP/1.1 con facilidad. Esta se utiliza para enviar cadenas de consultas a las URLs específicas, de una manera automática, no se codifica ningún tipo de formato ni es necesario la intervención manual, esta es altamente utilizada para solicitudes de descargas de paquetes de Python y de demás información perteneciente a un link disponible en la red.<sup>30</sup>

#### 1.6.1.15 BEAUTIFULSOUP

Esta librería de Python es utilizada para el análisis de documentos dado que esta crea un árbol con todos los documentos y la información analizada de archivos con el formato HTML además de ser útil para extraer información de sitios web, permite la extracción de información de archivos XML y permitir la navegación en la información obtenida reduciendo así horas de trabajo en extraer la información de manera manual.<sup>31</sup>

#### 1.6.1.16 DEPLACY

Esta librería permite generar un terminal cuyo trabajo es hacer un análisis y "dibujo", sobre todo el esquema del código, esta librería fue desarrollada específicamente para los cuadernos de Google Colab, dado que permite una integración con el usuario y el cuaderno de manera sencilla, visual y con una mejor conexión de red.<sup>32</sup>

#### 1.6.1.17 ES\_CORE\_NEWS\_SM

Esta librería, aunque es una dependencia de Spacy, permite la multitarea y análisis del lenguaje en el lenguaje español, esta se utiliza para la asignación de vectores de tokens, etiquetas POS, análisis dependencias y entidades, estás con nombres específicos del contexto. La librería es primordial para todo el proceso de análisis textual desde la limpieza del mismo, hasta el análisis lingüístico.<sup>33</sup>

---

<sup>30</sup> Requests. [En línea]. {16 de enero de 2021} PyPI.[En línea] disponible en: <https://pypi.org/project/requests/>

<sup>31</sup> Beautiful Soup Documentation — Beautiful Soup 4.9.0 documentation. (s. f.). Beautiful Soup Documentation. [En línea] {18 de enero de 2021} disponible en: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

<sup>32</sup> Deplacy. [En línea] {20 de enero de 2021}. PyPI. [En línea] disponible en: <https://pypi.org/project/deplacy/>

<sup>33</sup> Spanish · spaCy Models Documentation. (s. f.). Spanish. [En línea] {20 de enero de 2021} disponible en: <https://spacy.io/models/es/>

### 1.6.1.18 PKG\_RESOURCES

Este es un módulo distribuido que proporciona una API para que las diferentes bibliotecas de Python accedan a los archivos y recursos tratados en el mismo cuaderno de desarrollo, este también permite descubrir diferentes complementos para las funciones específicas del lenguaje de Python, proporcionando así soporte en el tiempo de ejecución, cuando se usen extensiones y diferentes funciones en el código, este módulo soporta diferentes paquetes y permite fusionar módulos o librerías distribuidas por separado en la misma API o cuaderno de trabajo.<sup>34</sup>

### 1.6.1.19 IMP

Este módulo proporciona una interfaz para los mecanismos utilizados en la implementación de la función Import, la cual declara y define constantes y funciones en todo el cuaderno de ejecución en Python.<sup>35</sup>

### 1.6.1.20 IMP.RELOAD

Esta librería permite un paso entre un módulo y un paquete, dado que los reinicia para su nueva ejecución y uso en el código, esta función se utiliza para evitar reiniciar todo el intérprete de Python en este caso Google colab y perder información del archivo analizado en el momento, es importante destacar que hace parte del paquete de funciones en este proyecto.<sup>36</sup>

### 1.6.1.21 COUNTER

Esta función se refiere a un contenedor que realiza un seguimiento de cuántas veces existen valores similares en un arreglo o en lista en Python, esta se implementa en algoritmos estructurados que manejan datos o múltiples cantidades de datasets.<sup>37</sup>

### 1.6.1.22 IMPORT FILES

Esta función permite a los diferentes entornos de Jupyter notebook la importación

---

<sup>34</sup> Package Discovery and Resource Access using pkg\_resources — setuptools 51.1.2 documentation. (s. f.). Package Discovery and Resource Access using pkg\_resources. [En línea] {22 de enero de 2021} disponible en: [https://setuptools.readthedocs.io/en/latest/pkg\\_resources.html](https://setuptools.readthedocs.io/en/latest/pkg_resources.html)

<sup>35</sup> imp — Access the import internals — Python 3.9.1 documentation. (s. f.). imp. [En línea] {22 de enero de 2021} disponible en: <https://docs.python.org/3/library/imp.html>

<sup>36</sup> Python Examples of Impreload. (s. f.). Python Examples of imp.reload. [En línea] {22 de enero de 2021} disponible en: <https://www.programcreek.com/Python/example/6950/imp.reload>

<sup>37</sup> Hellmann, D. (s. f.). Counter - Python Module of the Week. Counter - Python Module of the Week. [En línea] {24 de enero de 2021} disponible en: <https://pymotw.com/2/collections/counter.html>

de conjuntos de datos y diferentes modelos para su entrenamiento, como también la habilidad durante la experiencia en la codificación, esta función es importante dado que no solo permite importar sino también descargar datos en Google Colab.

<sup>38</sup>

### 1.6.1.23 OPENPYXL

Esta técnica nos permite leer, escribir archivos xlsx, xslm, xltx, xltm Excel, esta presenta una funcionalidad para leer y escribir de forma nativa desde Python en el formato Office Open XML, ya que presenta facilidades en todos los entornos de notebooks.<sup>39</sup>

### 1.6.1.24 READ

Este método proporciona funciones que interactúan con el sistema operativo incluyendo módulos estándares de Python, donde proporciona funcionalidad para el manejo y descripción de los archivos, este permite funciones como lectura escritura y envío de los mismos archivos o sobre los mismos archivos subidos al notebook.<sup>40</sup>

### 1.6.1.25 SPLIT

Este método divide una cadena en una lista, donde cada palabra se convierte en un elemento de la lista de Python y a su vez puede especificar el separador predeterminado para cualquier espacio específico o carácter en blanco que se encuentre en la cadena original.<sup>41</sup>

### 1.6.1.26 OPENPYXL.LOAD\_WORKBOOK

Esta función controla, administra y maneja la carga y lectura de las diferentes hojas de cálculo de Excel, dado que controla celdas específicas con fórmulas predeterminadas manteniendo siempre el valor almacenado de la última vez que se leyó esta hoja de Excel, a su vez controla diferentes elementos de visual Basic

---

<sup>38</sup> Masum, M.(2020) How to import and Export Data sets Colab - Towards Data Science. Medium [En línea] {24 de enero de 2021} disponible en: <https://towardsdatascience.com/google-colab-import-and-export-datasets-eccf801e2971>

<sup>39</sup> Openpyxl - A Python library to read/write Excel 2010 xlsx/xlsm files — openpyxl 3.0.5 documentation. (s. f.). openpyxl A Python library to read/write Excel 2010 xlsx/xlsm files. [En línea] {26 de enero de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/>

<sup>40</sup> GeeksforGeeks. (2019, 18 junio). Python | os.read() method. [En línea] {26 de enero de 2021} disponible en: [https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read\(\)%20method%20in%20Python.bytes%20left%20to%20be%20read.](https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read()%20method%20in%20Python.bytes%20left%20to%20be%20read.)

<sup>41</sup> Python String split() Method. (s. f.). Python String split() Method. [En línea] {26 de enero de 2021} disponible en: [https://www.w3schools.com/Python/ref\\_string\\_split.asp](https://www.w3schools.com/Python/ref_string_split.asp)

preservándolos de manera que puedan ser editables y leídos por el entorno donde se está trabajando en Python.<sup>42</sup>

#### 1.6.1.27 ITER\_ROWS

Esta función permite un trabajo libre en las diferentes hojas de cálculo de Excel, esta se encarga tanto para proporcionar la extracción de celdas específicas como para que agregue objetos de validación en los diferentes datos de las hojas de cálculo, es utilizada principalmente para el manejo dinámico de las hojas de cálculo y sus datos de Excel.<sup>43</sup>

#### 1.6.1.28 CHUNK.TEXT

Esta función es utilizada como fragmentación de frases nominales, esta es implementada cuando se busca en fragmentos correspondientes a frases nominales individuales que son a menudo las piezas más pequeñas de una oración y las que tienen el significado esencial de la misma.<sup>44</sup>

#### 1.6.1.29 MOST\_COMMON

Esta función devuelve de la lista específica asignada, los elementos más comunes y el recuento de los mismos cuenta desde el elemento más común al menos común de todos no omite ningún tipo de elemento y los recuentos son iguales tanto así que se ordena arbitrariamente.<sup>45</sup>

#### 1.6.1.30 RE.FINDITER

Esta función se utiliza para iterar sobre las diferentes coincidencias de una cadena, esto le proporciona a la implementación información adicional como la ubicación y la coincidencia de los índices en una lista o cadena específica.<sup>46</sup>

---

<sup>42</sup> Simple usage — openpyxl 3.0.5 documentation. (s. f.). Simple usage — openpyxl 3.0.5 documentation. [En línea] {01 de febrero de 2021} disponible en:

<https://openpyxl.readthedocs.io/en/stable/usage.html>

<sup>43</sup> Openpyxl 3.0.5 documentation. (s. f.). iter\_rows openpyxl.worksheet.worksheet module. [En línea] {01 de febrero de 2021} disponible en:

<https://openpyxl.readthedocs.io/en/stable/api/openpyxl.worksheet.worksheet.html>

<sup>44</sup> Extracting Information from Text. (s. f.). Extracting Information from Text. [En línea] {01 de febrero de 2021} disponible en: <https://www.nltk.org/book/ch07.html>

<sup>45</sup> Code Faster with Line-of-Code Completions, Cloudless Processing. (s. f.). [En línea] {01 de febrero de 2021} Code Faster with Line-of-Code Completions. disponible en:

[https://www.kite.com/Python/docs/collections.Counter.most\\_common](https://www.kite.com/Python/docs/collections.Counter.most_common)

<sup>46</sup> Python Language - Iterando sobre los partidos usando `re.finditer` | Python Tutorial. (s. f.). Python Language - Iterando sobre los partidos usando `re.finditer`. [En línea] {01 de febrero de 2021} disponible en:

<https://riptutorial.com/es/Python/example/12097/iterando-sobre-los-partidos-usando--re-finditer>

## 1.6.2 MARCO TEÓRICO

Este marco corresponde a todo lo relacionado con fórmulas, metodologías, investigaciones similares y teorías pertinentes al desarrollo del proyecto efectuando.

### 1.6.2.1 RECUENTO DE INVESTIGACIONES Y DESARROLLOS RELACIONADOS

Se encuentran estudios que han dado importancia a la salud mental de los adolescentes y han utilizado herramientas tecnológicas para la detección de síntomas emocionales en esta población, como el estudio de Xue, Li, Jinfeng, Clifton y Clifford (2014)<sup>47</sup>, investigadores del Reino Unido y China, quienes diseñaron plataformas virtuales donde se pudieron reconocer síntomas relacionados con problemáticas como estrés y depresión.

Para el desarrollo de dichas plataformas, se estudió el comportamiento que tienen los adolescentes en redes sociales como blogs en línea y Twitter; de igual manera, se creó un léxico que incluía las expresiones frecuentemente utilizadas por los adolescentes en dichas redes. Finalmente se implementaron algoritmos de clasificación en plataformas como micro blogs, los cuales permitieron detectar cambios en el estado de ánimo de los adolescentes y alertar a padres y tutores para generar una pronta intervención.

Por otra parte, investigadores alemanes han diseñado aplicaciones para dispositivos móviles que permiten a los usuarios hacer seguimiento de su estado de ánimo y simultáneamente realizar actividades con ejercicios dirigidos al entrenamiento de habilidades que favorezcan hábitos de vida saludables. Un ejemplo de las mismas, es la aplicación móvil "PiaP" (Psychologist in a pocket por sus siglas en inglés) la cual fue diseñada mediante la construcción de un léxico, que incluyó palabras y expresiones que reflejaban la experiencia de la depresión.

Recientemente Farruque, Zaiane y Goebel (2019)<sup>48</sup> presentaron en la Joint European Conference on Machine Learning and Knowledge Discovery in Databases, otro estudio en el que a partir de la identificación de textos en

---

<sup>47</sup> Xue Y., Li Q., Jin L., Feng L., Clifton D.A., Clifford G.D. (2014) Detecting Adolescent Psychological Pressures from Micro-Blog. En: Zhang Y., Yao G., He J., Wang L., Smalheiser N.R., Yin X. (eds) Health Information Science. HIS 2014. Lecture Notes in Computer Science, vol 8423. Springer, Cham. Recuperado de: [https://link.springer.com/chapter/10.1007/978-3-319-06269-3\\_10](https://link.springer.com/chapter/10.1007/978-3-319-06269-3_10)

<sup>48</sup> Farruque N., Zaiane O., Goebel R. (2019) Augmenting Semantic Representation of Depressive Language: From Forums to Microblogs. En: Brefeld U., Fromont E., Hotho A., Knobbe A., Maathuis M., Robardet C. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2019. Recuperado de <https://sci-hub.tw/10.1007/978-3-030-46133-1>

plataformas de internet como blogs en línea, donde las personas se identificaban a sí mismas como depresivas, y a través del diseño previo de un algoritmo automático de aprendizaje, se plantearon modelos predictivos para la identificación del lenguaje depresivo en redes sociales. Esta investigación señala que la forma de expresión de las personas con depresión posee una representación lingüística específica, lo cual facilita la identificación de publicaciones con "lenguaje depresivo" que permite hacer una detección oportuna de los casos, para posteriormente remitir a una valoración psicológica y brindar una intervención oportuna.

### **1.6.2.2 TEORÍA DE APRENDIZAJE DE MÁQUINA**

Esta teoría fundamenta que la máquina puede aprender habilidades sin ser programados explícitamente para esta tarea, por el *“Aprendizaje de máquina es el campo de estudio que da a las computadoras la habilidad de aprender sin ser programadas explícitamente.”*<sup>49</sup>

La alta complejidad de los datos en la realidad generan complicaciones al momento de abordar este tipo de datos por tanto se necesita implementar un algoritmo explícito para este tipo de problemas lo cual en algunas ocasiones resulta imposible por tanto en este tipo de datos es cuando el aprendizaje de máquina es de vital importancia debido a que abordará problemas para los cuales no ha sido programados explícitamente<sup>50</sup>, existen dos tipos de aprendizaje los cuales son:

### **1.6.2.3 APRENDIZAJE SUPERVISADO**

Este aprendizaje se entrena por lo general otorgándoles ciertas características las cuales se denominan también preguntas, adicionalmente se le dan las respuestas que se denominan etiquetas, todo esto con la finalidad de que el algoritmo las combine y pueda hacer predicciones, existen 2 tipos de aprendizaje supervisado.

---

<sup>49</sup> Artur samuel - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {18 de agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)

<sup>50</sup> Nick McCrea - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)



- **REGRESIÓN**

Este tipo de aprendizaje arroja un resultado específico, generalmente las etiquetas suelen ser un valor numérico, por tanto se pueden obtener dígitos como dato resultante.<sup>51</sup>

- ❖ **REGRESIÓN LINEAL SIMPLE**

En un modelo de regresión lineal simple se busca explicar la relación entre una variable dependiente y el conjunto de las variables independientes.<sup>52</sup> En un modelo de regresión lineal simple se trata de explicar la relación de una variable de respuesta Y, y una variable explicativa X.

### **Ecuación 1**

*Regresión simple*

$$y = mx + b$$

*Ecuación 1. Regresión simple.*<sup>53</sup>

*Nota:* La ecuación de regresión lineal simple tiene la anterior expresión

En donde m y b números reales, m es la pendiente de la recta y b es la ordenada en el origen. El modelo de regresión lineal simple apreciado de manera gráfica puede detallarse en la Figura 1, sin embargo esta es una representación generalizada, puede tener modificaciones.

---

<sup>51</sup> Juan Zambrano - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de agosto de 2020} disponible en:

(<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)

<sup>52</sup> Víctor Román - Machine Learning Supervisado: Fundamentos de la Regresión Lineal {En línea}. {28 de agosto de 2020} disponible en:

(<https://medium.com/datos-y-ciencia/machine-learning-supervisado-fundamentos-de-la-regresi%C3%B3n-lineal-bbcb07fe7fd>)

<sup>53</sup> Víctor Román - Machine Learning Supervisado: Fundamentos de la Regresión Lineal {En línea}. {28 de agosto de 2020} disponible en:

(<https://medium.com/datos-y-ciencia/machine-learning-supervisado-fundamentos-de-la-regresi%C3%B3n-lineal-bbcb07fe7fd>)

## Figura 1

### Regresión lineal Simple

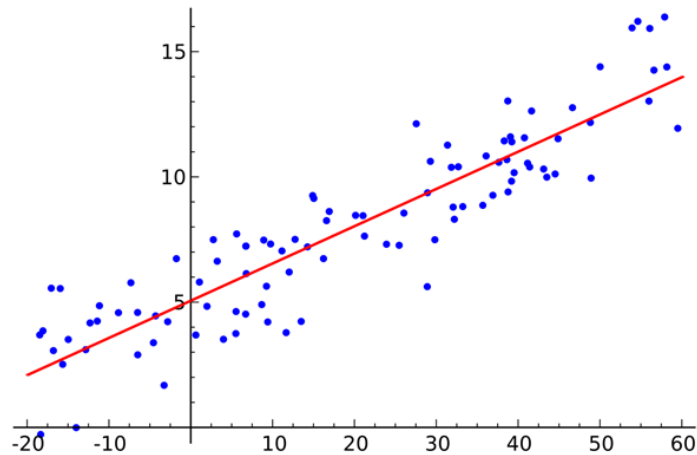


Figura 1. Regresión Lineal Simple <sup>54</sup>

Nota: Figura que muestra la regresión lineal simple de manera gráfica.

## ❖ REGRESIÓN LOGÍSTICA

Es un método de regresión que permite estimar la probabilidad de una variable cuantitativa, de una clasificación binaria, en que las observaciones se clasifican en un grupo u otro dependiendo del valor que tome la variable como predictor.<sup>55</sup>

### Ecuación 2

#### Regresión Logística

$$\frac{1}{1+e^{-t}}$$

Ecuación 2. Regresión Logística<sup>56</sup>

---

<sup>54</sup> Víctor Román - Machine Learning Supervisado: Fundamentos de la Regresión Lineal {En línea}. {28 de agosto de 2020} disponible en:

(<https://medium.com/datos-y-ciencia/machine-learning-supervisado-fundamentos-de-la-regresi%C3%B3n-lineal-bbcb07fe7fd>)

<sup>55</sup> Joaquin Amat Rodrigo - Regresión logística simple y múltiple {En línea}. {28 de agosto de 2020} disponible en:

([https://www.cienciadedatos.net/documentos/27\\_regresion\\_logistica\\_simple\\_y\\_multiple](https://www.cienciadedatos.net/documentos/27_regresion_logistica_simple_y_multiple))

<sup>56</sup> Joaquin Amat Rodrigo - Regresión logística simple y múltiple {En línea}. {28 de agosto de 2020} disponible en:

([https://www.cienciadedatos.net/documentos/27\\_regresion\\_logistica\\_simple\\_y\\_multiple](https://www.cienciadedatos.net/documentos/27_regresion_logistica_simple_y_multiple))

Nota: La ecuación de regresión logística tiene la anterior expresión.

Una representación generalizada de la gráfica obtenida al realizar una ecuación de regresión logística, es el que se puede apreciar en la figura 2, esta tiene particularmente esta forma creciente y curva que la caracteriza.

## Figura 2

### Regresión Logística

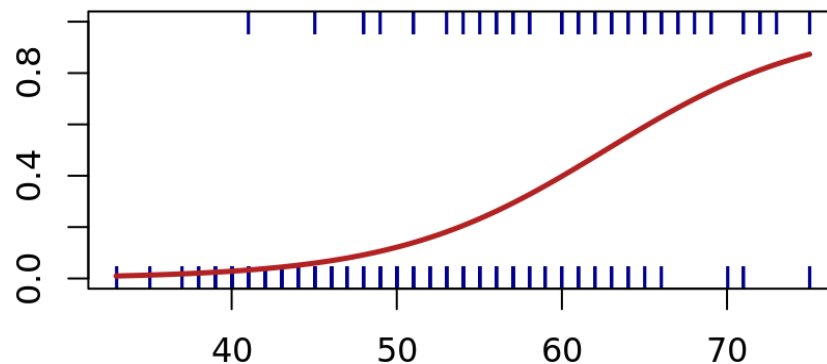


Figura 2. Regresión Logística<sup>57</sup>

Nota: Figura que muestra la regresión logística de manera gráfica.

## • CLASIFICACIÓN

Es una técnica para determinar la clase a la que pertenece la variable dependiente según una o más variables independientes<sup>58</sup> este tipo de aprendizaje, busca clasificar en grupos los diferentes elementos del dataset.<sup>59</sup>

## ❖ K-NN

---

<sup>57</sup> Joaquín Amat Rodrigo - Regresión logística simple y múltiple {En línea}. {28 de agosto de 2020}: ([https://www.cienciadedatos.net/documentos/27\\_regresion\\_logistica\\_simple\\_y\\_multiple](https://www.cienciadedatos.net/documentos/27_regresion_logistica_simple_y_multiple))

<sup>58</sup> Sitio Big Data - Clasificación de Aprendizaje automático supervisado (En línea). {24 de agosto de 2020} disponible en: (<https://sitiobigdata.com/2019/12/24/clasificacion-de-aprendizaje-automatico-supervisado/>)

<sup>59</sup> Juan Zambrano - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)

Es uno de los algoritmos de clasificación más simples, se utiliza para que identifique los puntos de datos que están separados en varias clases para poder predecir la clasificación de un nuevo punto de muestra, clasifica los casos nuevos en función de una medida de similitud, es muy bueno cuando el conjunto de entrada es pequeño, sin embargo con grandes conjuntos tiene problemas.<sup>60</sup>

Knn calcula la distancia entre puntos, para esto utiliza la fórmula:

### Ecuación 3

#### Distancia Ecludiana

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Ecuación 3. Distancia euclidiana.<sup>61</sup>

*Nota:* La ecuación de distancia euclidiana tiene la anterior expresión.

La fórmula de arriba toma n como el número de dimensiones, o también se podría decir las características de nuestro modelo.

La *figura 3. Modelo Knn*, es una representación gráfica del algoritmo de clasificación K-NN en donde la estrella roja representa los puntos de prueba, y esta misma se encuentra rodeada por puntos amarillos y azules los cuales representan dos clases distintas. El procedimiento a seguir es hallar la distancia desde el punto de prueba a cada punto de detalle en la gráfica, por cada punto hay una distancia específica por lo tanto en este gráfico si hay diez puntos, hay 10 distancias, y con esto se determina la distancia más corta existente para predecir que esta pertenece a la misma clase o no.<sup>62</sup>

<sup>60</sup> Clasificación de Aprendizaje automático supervisado {En línea}. {29 de agosto de 2020} disponible en:

(<https://sitiobigdata.com/2019/12/24/clasificacion-de-aprendizaje-automatico-supervisado/>)

<sup>61</sup> Clasificación de Aprendizaje automático supervisado {En línea}. {29 de agosto de 2020} disponible en:

(<https://sitiobigdata.com/2019/12/24/clasificacion-de-aprendizaje-automatico-supervisado/>)

<sup>62</sup> Shubham Panchal - k Nearest Neighbor Classifier (kNN) - Machine Learning Algorithms {En línea}. {29 de agosto de 2020} disponible en:

(<https://medium.com/@equipintelligence/k-nearest-neighbor-classifier-knn-machine-learning-algorithm-ed62feb86582>)

### Figura 3

#### Modelo Knn

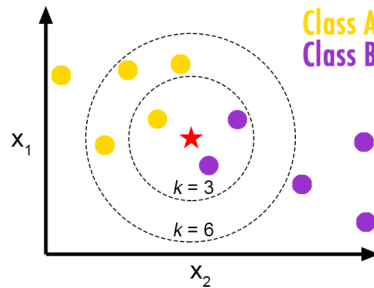


Figura 3. Modelo Knn<sup>63</sup>

Nota: Figura que muestra el funcionamiento del modelo Knn de manera gráfica.

#### ❖ SVM

Las svm son un tipo de máquinas de Aprendizaje, que se basa en el concepto de planos de decisión, que separa uno entre un conjunto de objetos. Los vectores de soporte son los puntos que definen el margen máximo de separación del hiperplano (un hiperplano es una recta, divide el plano en dos mitades.<sup>64</sup>).

Son vectores porque estos puntos tienen el mismo número de elementos como de dimensiones, por tanto son representados por un vector de dimensiones.<sup>65</sup>

#### Ecuacion 4

#### Formula Svm

$$g(x) = w^t x + b$$

<sup>63</sup> Shubham Panchal - k Nearest Neighbor Classifier (kNN) - Machine Learning Algorithms {En línea}. {29 de agosto de 2020} disponible en:

(<https://medium.com/@equipintelligence/k-nearest-neighbor-classifier-knn-machine-learning-algorithm-ed62feb86582>)

<sup>64</sup> Esacademic - Hiperplano {En línea}. {29 de agosto de 2020} disponible en: (<https://esacademic.com/dic.nsf/eswiki/574819>)

<sup>65</sup> Jose Martínez Heras - Máquinas de Vectores de Soporte (SVM) {En línea}. {29 de agosto de 2020} disponible en: (<https://www.iaartificial.net/maquinas-de-vectores-de-soporte-svm/>)

Ecuación 4. Fórmula svm<sup>66</sup>

*Nota:* La fórmula del modelo SVM tiene la anterior expresión.

Donde W son los vectores ortogonales, x es la variable dependiente y b la variable independiente teniendo en cuenta que x también es un vector.

## ❖ ÁRBOL DE DECISIÓN

Es un algoritmo basado en árboles se considera uno de los mejores y más utilizados métodos de aprendizaje supervisado.<sup>67</sup> Este método potencia modelos predictivos con alta precisión, adicionalmente ofrecen facilidad de interpretación, mapean muy bien, se suelen utilizar en diversos problemas de ciencias de datos debido a las características mencionadas anteriormente.

### Ecuación 5

*Gini Score*

$$\text{Classification: } G = \sum(pk * (1 - pk))$$

Ecuación 5. Gini score<sup>68</sup>

*Nota:* La fórmula de Gini Score tiene la anterior expresión.

El parámetro gini es muy importante en el algoritmo de árbol de decisión, debido a que da una idea de que tan bien se dividió el árbol de acuerdo a como se establece la distribución de las clases de respuesta, PK es la proporción de entradas de la misma clase presentes en un grupo particular.

---

<sup>66</sup> Jose Martínez Heras - Máquinas de Vectores de Soporte (SVM) {En línea}. {29 de agosto de 2020} disponible en: (<https://www.iaartificial.net/maquinas-de-vectores-de-soporte-svm/>)

<sup>67</sup> Sitio Big Data - Árbol de decisión en Machine Learning (Parte 1) {En línea}. {29 de agosto de 2020} disponible en: (<https://sitiobigdata.com/2019/12/14/arbol-de-decision-en-machine-learning-parte-1/>)

<sup>68</sup> Sitio Big Data - Árbol de decisión en Machine Learning (Parte 1) {En línea}. {29 de agosto de 2020} disponible en: (<https://sitiobigdata.com/2019/12/14/arbol-de-decision-en-machine-learning-parte-1/>)

## Figura 4

### Árbol de decisión

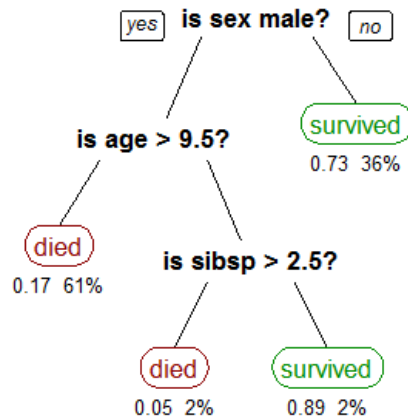


Figura 4. Árbol de decisión<sup>69</sup>

Nota: Figura que muestra el funcionamiento del árbol de decisión de manera gráfica.

Como se aprecia en la *Figura 4. Árbol de decisión*, un árbol se dibuja de arriba hacia abajo, su raíz está en la parte de arriba, el texto en negrilla representa la condición, basándose en que árbol se va a dividir en ramas, la rama que no se divide más es la hoja, en este caso las hojas serán representadas tanto verdes como rojas independientemente si sobrevivió o murió.<sup>70</sup>

### ❖ BOSQUE ALEATORIO

El bosque aleatorio es un modelo de predicción el cual se basa en las reglas binarias, por lo tanto tiene 2 estados de 1 o 0, con esta regla se reparte las observaciones, para lograr predecir el valor de salida. Este método engloba técnicas supervisadas, sin parámetros, las cuales consiguen segmentar los elementos, en regiones simples, en las cuales se pueden manejar las interacciones.

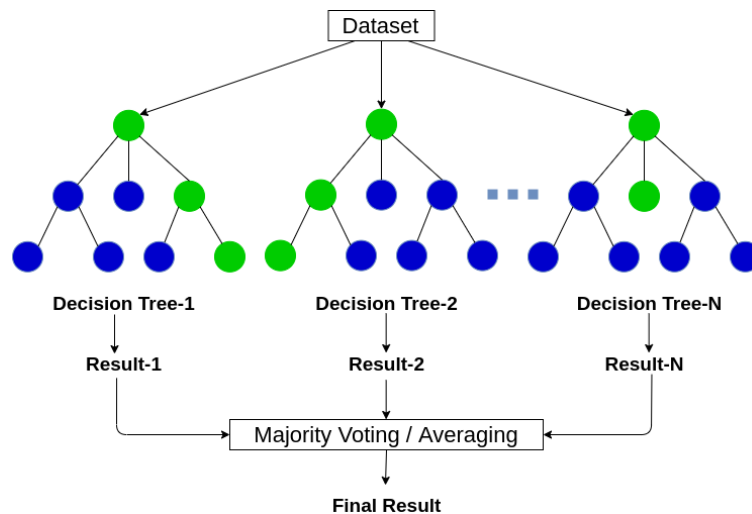
<sup>69</sup> Prashant Gupta - Decision Trees in Machine Learning {En línea}. {29 de agosto de 2020}; (<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)

<sup>70</sup> Prashant Gupta - Decision Trees in Machine Learning {En línea}. {29 de agosto de 2020} disponible en: (<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)

En la *Figura 5. Árbol de decisión Aleatorio*, se representan los modelos de bosque aleatorio los cuales están conformados por árboles de decisión, cada uno es entrenado con una muestra un poco diferente, la predicción de un nuevo elemento se obtiene de las agregaciones de cada árbol de manera individual, tal como se aprecia el proceso gráfico a continuación:

**Figura 5**

*Árbol de Decisión Aleatorio*



*Figura 5. Árbol de decisión Aleatorio* <sup>71</sup>

*Nota:* Figura que muestra el funcionamiento del árbol de decisión aleatorio de manera gráfica.

#### 1.6.2.4 APRENDIZAJE NO SUPERVISADO

Este tipo de aprendizaje a diferencia del aprendizaje supervisado solo tiene las características, sin proporcionar ninguna etiqueta, por tanto su función es agrupar. El algoritmo deberá clasificarse por similitud y de esta manera podrá crear grupos, sin la capacidad de definir cómo es la individualidad de cada integrante del grupo.

<sup>72</sup>

- **K-MEANS**

<sup>71</sup> ABHISHEK SHARMA, - Decision Tree vs. Random Forest – Which Algorithm Should you Use (mayo de 2020) {En línea}. {10 de enero de 2021} disponible en:

(<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)

<sup>72</sup> Juan Zambrano - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de agosto de 2020} disponible en:

(<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)



Es un algoritmo de aprendizaje no supervisado, el cual agrupa objetos en  $k$  grupos basándose en las características de estos. El agrupamiento se realiza a través de minimizar la suma de las distancias entre cada objeto y centroide de su grupo o dicho de otra forma cluster, suele ser común usar la distancia cuadrática <sup>73</sup>este algoritmo consta de 3 pasos:

1. **Inicialización:** una vez escogido el número de grupos,  $k$ , se establecen  $k$  centroides en el espacio de los datos.
2. **Asignación objetos a los centroides:** cada objeto de los datos es asignado a su centroide más cercano.
3. **Actualización centroides:** se actualiza la posición del centroide de cada grupo tomando como nuevo centroide la posición del promedio de los objetos pertenecientes a dicho grupo.

Los pasos de asignación de objetos a los centroides y su actualización se realiza hasta que los centroides no se muevan, este algoritmo resuelve una fórmula de optimización (minimizar):

## Ecuación 6

*Fórmula de optimización*

$$\min_{\mathbf{S}} E(\boldsymbol{\mu}_i) = \min_{\mathbf{S}} \sum_{i=1}^k \sum_{\mathbf{x}_j \in \mathcal{S}_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2$$

*Ecuación 6. Fórmula de optimización*<sup>74</sup>

*Nota:* En esta ecuación La fórmula de optimización  $\mathbf{S}$  es el conjunto de datos, sus elementos son  $\mathbf{x}_j$  representados por vectores,  $k$  son los clusters con su correspondiente centroide.

### 1.6.1.5 REDES NEURONALES

Las redes neuronales son un modelo basado en el cerebro humano, como su nombre lo indica, se conforma por un conjunto de nodos, cada conjunto está conformado por neuronas, cada neurona tiene un valor y un peso, todas las

<sup>73</sup> Universidad de Oviedo - El algoritmo k-means aplicado {En línea}. {29 de agosto de 2020} disponible en: ([https://www.unioviado.es/compnum/laboratorios\\_py/kmeans/kmeans.html](https://www.unioviado.es/compnum/laboratorios_py/kmeans/kmeans.html))

<sup>74</sup> Universidad de Oviedo - El algoritmo k-means aplicado {En línea}. {29 de agosto de 2020} disponible en: ([https://www.unioviado.es/compnum/laboratorios\\_py/kmeans/kmeans.html](https://www.unioviado.es/compnum/laboratorios_py/kmeans/kmeans.html))

neuronas están conectadas entre sí de un nodo a otro, cada conexión también posee su propio valor.

## Ecuación 7

*Fórmula de Red Neuronal*

$$y = f(\sum w_i x_i)$$

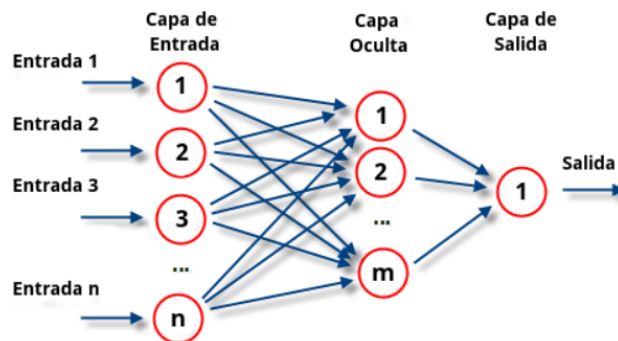
*Ecuación 7. Fórmula de Red Neuronal<sup>75</sup>*

*Nota:* Esta figura muestra la ecuación de una red neuronal.

La *Figura 6. Modelo Red Neuronal*, describe el proceso por el cual las redes reciben un conjunto de valores por un nodo, que son las aquí llamadas neuronas encargadas de procesar la información. El conjunto de nodos forma la red neuronal, la cual transforma los valores de entrada y continúan siendo modificados por cada conjunto de nodos hasta llegar a la salida.<sup>76</sup>

## Figura 6

*Modelo Red Neuronal*



*Figura 6. Modelo Red Neuronal<sup>77</sup>*

*Nota:* Figura que muestra de manera gráfica el modelo de una red neuronal.

<sup>75</sup>Goica21, Monografias.com. (s. f.). Redes neuronales (página 2) - Monografias.com. monografías. <https://www.monografias.com/trabajos12/redneuro/redneuro2.shtml>

<sup>76</sup>Atria Innovation - (22 de octubre 2019) " {En línea}. {10 de enero de 2021} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

<sup>77</sup>Atria Innovation - (22 de octubre 2019) " {En línea}. {10 de enero de 2021} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

### 1.6.1.6 MATRIZ DE CONFUSIÓN

Es una tabla que describe el rendimiento de un modelo supervisado de Machine Learning en los datos de prueba, donde se desconocen los verdaderos valores. Se llama “matriz de confusión” porque hace que sea fácil detectar dónde el sistema está confundiendo dos clases.<sup>78</sup>

#### Figura 7

*Matriz de Confusión*

		ACTUAL VALUES	
		POSITIVE	NEGATIVE
PREDICTED VALUES	POSITIVE	TP	FP
	NEGATIVE	FN	TN

Figura 7. Matriz de Confusión <sup>79</sup>

*Nota:* En esta figura se muestra la matriz de confusión.

La Figura 7. Matriz de Confusión, es una representación generalizada de una matriz de confusión, esta matriz de confusión está compuesta por 4 tipos de clasificadores que son los siguientes:

1. True Positives (TP): cuando la clase real del punto de datos era 1 (Verdadero) y la predicha es también 1 (Verdadero).
2. Verdaderos Negativos (TN): cuando la clase real del punto de datos fue 0 (Falso) y el pronosticado también es 0 (Falso).
3. False Positives (FP): cuando la clase real del punto de datos era 0 (False) y el pronosticado es 1 (True).
4. False Negatives (FN): Cuando la clase real del punto de datos era 1 (Verdadero) y el valor predicho es 0 (Falso).”

<sup>78</sup> Juan Ignacio barrios - Everything you Should Know about Confusion Matrix for Machine Learning La matriz de confusión y sus métricas (2019)” {En línea}. {18 de agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

<sup>79</sup> Aniruddha Bhandari - Everything you Should Know about Confusion Matrix for Machine Learning {En línea}. {29 de agosto de 2020} disponible en: (<https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/>)

### 1.6.1.7 MÉTRICA: precisión

Se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud. Cuanto menor es la dispersión mayor la precisión. Se representa por la proporción entre el número de predicciones correctas (tanto positivas como negativas) y el total de predicciones.

#### Ecuación 8

*Ecuación de precisión*

$$precision = \frac{TP}{TP+FP}$$

*Ecuación 8. Precisión*<sup>80</sup>

*Nota: En la figura anterior se muestra la ecuación de la métrica de precisión.*

Donde el mejor caso da como resultado 1, por lo tanto, todas las predicciones positivas, fueron reales. Mientras que el peor caso es 0, donde todas las predicciones positivas, fueron falsas.<sup>81</sup>

### 1.6.1.8 MÉTRICA: RECALL

Es el número de elementos identificados correctamente como positivos del total de positivos verdaderos.<sup>82</sup>

#### Ecuación 9

*Ecuación de Recall*

$$recall = \frac{TP}{TP+FN}$$

*Ecuación 9. Recall*<sup>83</sup>

*Nota: En la figura anterior se aprecia la ecuación de la métrica recall.*

---

<sup>80</sup>Juan Ignacio barrios - La matriz de confusión y sus métricas (2019)" {En línea}. {18 de agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

<sup>81</sup>Juan Ignacio barrios - La matriz de confusión y sus métricas (2019)" {En línea}. {18 de agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

<sup>82</sup>Juan Ignacio barrios - La matriz de confusión y sus métricas (2019)" {En línea}. {18 de agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

<sup>83</sup>Juan Ignacio barrios - La matriz de confusión y sus métricas (2019)" {En línea}. {18 de agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

Esta métrica busca informar sobre la cantidad que el modelo de machine learning es capaz de identificar, si el resultado es 1 esto significa que ha encontrado todo el modelo como información relevante, se entenderá como una recuperación de información perfecta, por tanto entre más cercano de a 1 mejor será el resultado.

#### 1.6.1.9 MÉTRICA: PUNTUACIÓN F1

F1 es una medida general de la precisión de un modelo que combina precisión y recuperación, de esa forma extraña que la suma y la multiplicación simplemente mezclan dos ingredientes para hacer un plato por completo. Es decir, un buen puntaje F1 significa que tiene bajos falsos positivos y bajos falsos negativos, por lo que está identificando correctamente las amenazas reales y no le molestan las falsas alarmas. Un puntaje F1 se considera perfecto cuando es 1, mientras que el modelo es un fracaso total cuando es 0.<sup>84</sup>

#### Ecuación 10

*Métrica Puntuación F1*

$$F1 = 2 * \left( \frac{precision * recall}{precision + recall} \right)$$

*Ecuación 10. Métrica puntuación F1<sup>85</sup>*

*Nota: En la ecuación anterior se muestra como es el cálculo de la métrica puntuación F1*

El valor F1 asume que nos importa de igual forma la precisión y la exhaustividad. Esto no tiene que ser así en todos los problemas. Por ejemplo, cuando necesitamos predecir si hay riesgo de que un trozo de basura espacial se choque con un satélite, podemos valorar más la exhaustividad a riesgo de tener una peor precisión.

#### 1.6.1.10 MÉTRICA: EXACTITUD

La exactitud mide el porcentaje de casos que el modelo ha acertado, es la fracción de predicciones que el modelo realizó correctamente. En la clasificación binaria, la exactitud también se puede calcular en términos de positivos y negativos. Para

---

<sup>84</sup>Jose Martínez Heras - precisión, Recall, F1, Accuracy en clasificación {En línea}. {18 de agosto de 2020} disponible en: (<https://www.iartificial.net/precisión-recall-f1-accuracy-en-clasificacion/>)

<sup>85</sup>Jose Martínez Heras - precisión, Recall, F1, Accuracy en clasificación {En línea}. {18 de agosto de 2020} disponible en: (<https://www.iartificial.net/precisión-recall-f1-accuracy-en-clasificacion/>)

este proyecto la implementación exclusiva de la métrica de exactitud sola no muestran el panorama completo del conjunto de datos, es por eso que esta se integra con otras métricas de validación.<sup>86</sup>

### Ecuación 11

$$Exactitud = \frac{TP+TN}{TP+TN+FP+FN}$$

*Ecuación 11. Exactitud*<sup>87</sup>

Nota: En el gráfico anterior se muestra la ecuación de la métrica de desempeño de exactitud

La métrica de exactitud tiene limitaciones, debido a que no funciona bien con las clases desequilibradas porque pueden tener muchos elementos de la misma clase, en otras palabras se refiere a que tan cerca están los valores predichos de los valores reales.

#### 1.6.1.11 PROTÉGÉ

Es un editor de ontologías y un sistema de adquisición de conocimiento. Está disponible como código abierto.<sup>88</sup> Adicionalmente fue construido por desarrolladores de Stanford del centro informático biomédico en la escuela universitaria Stanford de medicina.

Protégé es un medio por el cual se busca adquirir, representar y procesar información, sobre la salud humana, el sistema Protégé busca el desarrollo y mantenimiento de ontologías.<sup>89</sup>

#### 1.6.1.12 ANÁLISIS DE TEXTO

El análisis de textos, consiste en analizar cada elemento de este, desglosando cada elemento que integra el texto, buscando su estructura, sentido y finalidad. Es el proceso de deducir el significado de un texto, con el fin de medir la opinión de las personas.

---

<sup>86</sup> Developers Google-Clasificación:Exactitud{En línea}{29 de agosto de 2020} disponible en:

(<https://developers.google.com/machine-learning/crash-course/classification/accuracy?hl=es-419>)

<sup>87</sup> Jose Martínez Heras - precisión, Recall, F1, Accuracy en clasificación {En línea}. {18 de agosto de 2020} disponible en: (<https://www.iartificial.net/precisión-recall-f1-accuracy-en-clasificacion/>)

<sup>88</sup> Maria Manuela Valenzuela Sosa - Herramienta Protégé para la Web Semántica {En línea}. {19 de agosto de 2020} disponible en:

(<https://es.slideshare.net/Mvalenzuelasosa/herramienta-protg-para-la-web-semntica>)

<sup>89</sup> Universidad de Stanford - Protégé {En línea}. {25 de agosto de 2020} disponible en:

(<https://Protégé.stanford.edu/about.php>)

- **ANÁLISIS DE SENTIMIENTOS**

Este campo por lo general estudia los elementos subjetivos, definidos como expresiones lingüísticas de los estados particulares en contexto, el análisis de sentimientos es muy importante en el monitoreo de redes sociales, permite segmentar las opiniones públicas por tanto es fundamental tener en cuenta los beneficios del análisis de sentimientos debido a que son numerosos e importantes permitiendo entender de manera rápida y eficaz las actitudes de la población seleccionada.<sup>90</sup>

Mediante el análisis de sentimientos, se busca entender la relación de texto a analizar, después de entender la relación y opinión que tiene el texto analizado se busca conocer el sentido que tiene esa opinión, como también la intensidad de esa opinión.<sup>91</sup>

- **TOKENIZACIÓN**

La tokenización es un término utilizado para dividir un texto en un conjunto de caracteres, se podría decir que es una gran herramienta para el análisis de textos, la mejor forma de verlo es ver cada token como una palabra, dependiendo el tipo de problemas el token puede ser utilizado de diversas formas.

La forma de elegir los tokens en los textos depende mayormente del problema que se abarque, en algunas ocasiones simplemente dividiendo el texto por espacio será suficiente, sin embargo es importante entender, que tokenizar los textos permitiría utilizar otras herramientas para poder procesar los textos de una mejor manera.<sup>92</sup> Para una representación gráfica del proceso de tokenización puede dirigirse al ejemplo respetando en la *Figura 8. Funcionamiento tokenización*. El cual describe el proceso de tokenización de la frase “El gato es el animal” generando los tokens “El”, “gato”, “es”, “animal”.

### **Figura 8**

#### *Funcionamiento Tokenización*

---

<sup>90</sup> Brandwatch - Entendiendo el análisis de sentimiento: qué es y para qué se usa {En línea}. {24 de agosto de 2020} disponible en: (<https://www.brandwatch.com/es/blog/analisis-de-sentimiento/>)

<sup>91</sup> Manuel Delgado - ¿Qué es el análisis del sentimiento? {En línea}. {24 de agosto de 2020} disponible en: (<https://manueldelgado.com/que-es-el-analisis-del-sentimiento/>)

<sup>92</sup> Antonio Feregrino Bolaños - Introducción al análisis de textos {En línea}. {10 de enero de 2021} disponible en: (<https://tacosdedatos.com/analisis-texto>)

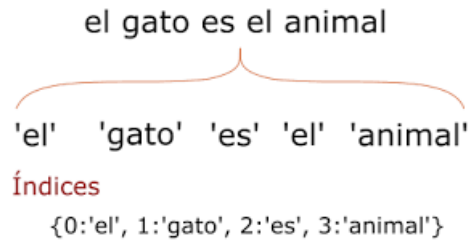


Figura 8. Funcionamiento tokenización. <sup>93</sup>

*Nota: En la figura anterior se muestra el funcionamiento de la tokenización.*

- **LEMATIZACIÓN**

La lematización es un proceso lingüístico, que consiste en hallar como una palabra está relacionada con el lema correspondiente, el lema es la forma estándar en la cual está definida una palabra, en síntesis es el representante de esa palabra en sus otras expresiones verbales.

- **STEMMING**

Este proceso es básicamente recortar las palabras a una base común, es muy útil para poder detectar la raíz de estas, adicionalmente la gran ventaja de esta herramienta es que funciona de una manera óptima en varios idiomas, al contrario que la lematización no necesita de una base de datos para guardar sus palabras simplemente recorta las una base la cual se puede utilizar para mejorar el resultado del procesamiento de texto <sup>94</sup>

### 1.6.1.13 ONTOLOGÍA EN SISTEMAS

La palabra ontología en ciencias de la computación, se refiere a una definición formal de tipos, propiedades, y relaciones entre entidades que realmente o fundamentalmente existen para un dominio de discurso en particular. En este sentido, sería una especie de aplicación práctica de la ontología filosófica, con una taxonomía dada a un campo de estudio específico de las ciencias de la computación, que es el área que nos compete. <sup>95</sup>

<sup>93</sup> Brandwatch - Entendiendo el análisis de sentimiento: qué es y para qué se usa {En línea}. {24 de agosto de 2020} disponible en: (<https://www.brandwatch.com/es/blog/analisis-de-sentimiento/>)

<sup>94</sup> Elena Álvarez Mellado - La lematización de textos{En línea}. {10 de enero de 2021} disponible en: (<https://es.slideshare.net/lirondos/la-lematizacion-de-textos>)

<sup>95</sup> Natalya F. Noy and Deborah L. McGuinness "Ontology Development 101: A Guide to Creating Your First Ontology". {En línea}. {01 octubre de 2020} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).



## 1.7 ESTADO DEL ARTE

En la exploración y búsqueda de estudios, investigaciones, proyectos y artículos similares, a través de las bases de datos de biblioteca disponibles para la Universidad Católica de Colombia como lo son *Scopus*, *Sciencedirect*, entre otros. Se implementan ecuaciones y conceptos específicos de búsqueda como: DEPRESIÓN, TRATAMIENTOS DEPRESIVOS EN COLOMBIA, INFORMÁTICA EN LA DEPRESIÓN, MÉTODO AUTOMÁTICO PARA TEXTOS, LENGUAJE NATURAL y demás, para así identificar variedad de publicaciones respecto a los avances más importantes en el área de la salud mental, específicamente depresión y todo lo convenido a la relación entre esta área y el área de la informática, el análisis de textos y la aplicabilidad de la ingeniería de sistemas en este proyecto. Las publicaciones destacadas se comentarán a continuación.<sup>96</sup>

En primer lugar se encuentra el documento *BOLETÍN DE SALUD MENTAL DEPRESIÓN SUBDIRECCIÓN DE ENFERMEDADES NO TRANSMISIBLES*<sup>97</sup> realizado por el ministerio de salud de Colombia en marzo de 2017 se presenta toda la información de salud mental, en especial, la depresión, presente en la población Colombiana. Qué será de interés para el área de salud y de comunicaciones en general. Este boletín también señala índices de la depresión en América Latina y el Caribe, además de explicar conceptos del grupo de enfermedades y consecuencias que acoge la depresión, entre esos la morbilidad, mortalidad y discapacidades. También menciona a la Organización Mundial de la Salud que da a conocer su reporte oficial, en el cual aparece que la depresión es una enfermedad mental que se encuentra en uno de los primeros trastornos con más frecuencia en todo el mundo. También en este mismo boletín el sistema integral de información de la protección social reporta las atenciones e

---

<sup>96</sup> Podrás encontrar las varias investigaciones que fueron llevadas a cabo para la consolidación del estado del arte en el Anexo número 2, de este proyecto de grado.

<sup>97</sup> Ministerio de salud, del gobierno de Colombia .”Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles” [En línea] {03 de febrero de 2021} disponible en: (<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)

identificaciones de casos de depresión en Colombia, que han ido en incremento desde el año 2009 hasta el 2015.

La relación que presenta este boletín de Salud Mental, con el proyecto que aquí se adelanta, consiste en ser una primera guía para conocer el área de trabajo en el cual estamos construyendo este método automático. Toda la información del área de depresión se presenta de una manera general y centrada en Colombia lo que nos permite saber a qué población irá dirigido este proyecto en nuestro país y si este tema es de importancia para todas las entidades de salud mental públicas tanto en Colombia como en América Latina y el Caribe.

En el sentido de indagar información en otros países de latino América, se encuentra la investigación publicada por las universidades *Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil* y la *Universidad Federal de Minas Gerais, Belo Horizonte, Brasil*, titulado SELF-COMPASSION AND SYMPTOMS OF STRESS, ANXIETY, AND DEPRESSION (“AUTOCOMPASIÓN Y SÍNTOMAS DE ESTRÉS, ANSIEDAD Y DEPRESIÓN”)<sup>98</sup> el cual trata temas sobre autocompasión y síntomas de estrés, ansiedad y depresión a través de una investigación acompañada de un documento. Se analizó y se buscó una relación entre el significado de comentarios realizados por diferentes tipos de personas y su relación respecto a síntomas de sintomatología depresiva.

Estos estudios fueron realizados en personas masculinas y femeninas en diferentes rangos de edad, en donde la relación más directa fue la que presentan la gente que asiste a psicoterapia con relación a personas que no asisten a la psicoterapia. La documentación de la misma muestra una serie de conocimientos prácticos y teóricos obtenidos de esta investigación, los cuales se modelaron en diferentes sistemas de gestión de datos que fueron próximamente utilizados para llevar a cabo una serie de estudios y construir nuevas hipótesis que expliquen, distintas relaciones ya sea desde la conducta social, la madurez, y la afectación de las terapias psicológica con la salud mental de las personas. Estos estudios resultan útiles al presente proyecto en la medida en que muestran cómo diferentes investigaciones en el área de identificación y relación de sintomatología depresiva con el lenguaje escrito, son realizadas en este momento y brinda una serie de pautas y modelos de investigación.

---

<sup>98</sup> Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. "Self-compassion and Symptoms of Stress, Anxiety, and Depression" [En línea] {03 de febrero de 2021} disponible en: (<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85082776505&origin=resultslist&sort=plf-f&src=s&st1=sintomas+depresi%c3%b3n&st2=&sid=60b9ad8431f907b88a1e1093beb09ff8&sot=b&sdt=b&sl=33&s=TITLE-ABS-KEY%28sintomas+depresi%c3%b3n%29&relpos=2&citeCnt=1&searchTerm=>)

Según el modelo visto en la investigación se encontró un artículo titulado *TEXT MINING ANALYSIS OF COMMENTS IN THAI LANGUAGE FOR DEPRESSION FROM ONLINE SOCIAL NETWORKS (ANÁLISIS DE MINERÍA DE TEXTOS DE LOS COMENTARIOS EN TAILANDÉS PARA LA DEPRESIÓN DE LAS REDES SOCIALES EN LÍNEA)*<sup>99</sup> el cual analiza la relación de las frases y las palabras comúnmente encontradas en los comentarios de depresión en la red social Twitter. Para esto se utilizaron diferentes reglas de asociación de datos, al igual que diferentes tipos de comentarios en la red social, es de resaltar que estos comentarios son variables respecto al lugar geográfico del estudio y las normas que aplican en este. Se utilizaron diferentes software para reunir y difundir la información obtenida en esta investigación. Entre los resultados obtenidos y expuestos en la documentación de la investigación anteriormente realizada señalan una relación entre comentarios depresivos y la posible sintomatología que presente la persona que lo escribió.

Esta investigación resulta significativa porque propone un primer acercamiento que muestra cómo las herramientas tecnológicas tanto de recolección, análisis de datos y publicación de resultados, son esenciales para las investigaciones de salud mental en las comunidades actuales. Así mismo permitió un acercamiento a una lista de diferentes palabras o conceptos identificados en los comentarios de Twitter, en donde se relacionan directamente con la presencia de sintomatología depresiva en la persona.

Teniendo en cuenta la información del artículo anterior, es importante entender cómo machine learning se puede usar para la detección de depresión, en la investigación *MACHINE LEARNING FOR DEPRESSION SCREENING IN ONLINE COMMUNITIES (APRENDIZAJE AUTOMÁTICO PARA DETECCIÓN DE DEPRESIÓN EN COMUNIDADES EN LÍNEA)*<sup>100</sup> en la cual se presenta un enfoque computacional para la clasificación y detección de posibles usuarios que presenten síntomas de depresión o algún otro síntoma primario de una

---

<sup>99</sup> Chaiwuttisak, P. Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand. "Text Mining Analysis of Comments in Thai Language for Depression from Online Social Networks " [En línea] {05 de febrero de 2021} disponible en:

(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85088469231&origin=resultslist&sort=plf-f&src=s&st1=Depression+and+Software&st2=&sid=10a2d25e36c518b4f31ef2fda413dbc6&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Depression+and+Software%29&relpos=1&citeCnt=0&searchTerm=>)

<sup>100</sup> Trifan, A. , Antunes, R. Oliveira, J.L. DETI/IEETA, University of Aveiro, Aveiro, Portugal {En línea}. "Machine Learning for Depression Screening in Online Communities " [En línea] {05 de febrero de 2021} disponible en:

(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85089220717&origin=resultslist&sort=plf-f&src=s&st1=Language+for+Depression&st2=&sid=4507d7f667902348aa8d12e0318915fd&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Language+for+Depressio n%29&relpos=1&citeCnt=0&searchTerm=>)

enfermedad mental. Este artículo se basa en la documentación de un estudio de evaluación cruzada en donde con conjuntos de datos públicos relacionados en una misma red social, permite la detección y tratamiento interpretativo de los datos. El entorno de trabajo y de desarrollo de este estudio fue realizado en una red social debido a que son una fuente potencial de información, permite saber las opiniones de personas de una manera digital. La finalidad de este estudio es que con los resultados presentados sea posible tratar métodos de Big Data con métodos de detección de enfermedades de salud mental.

Es importante resaltar que muestra cómo el aprendizaje automático en grandes cantidades de datos, nos permite detectar sintomatología depresiva o alguna otra sintomatología respectiva al área de enfermedades de salud mental, además de que este puede ser utilizado como referente técnico de estudio para trabajar conceptos de interoperabilidad entre sistemas y el área de la salud mental.

Teniendo en cuenta que para detectar síntomas depresivos es necesario encontrar un comportamiento en el trabajo *A UNIFIED NEUROCOGNITIVE MODEL OF SEMANTICS LANGUAGE SOCIAL BEHAVIOUR AND FACE RECOGNITION IN SEMANTIC DEMENTIA (UN MODELO NEUROCOGNITIVO UNIFICADO DE SEMÁNTICA, LENGUAJE, COMPORTAMIENTO SOCIAL Y RECONOCIMIENTO FACIAL EN LA DEMENCIA SEMÁNTICA.)*<sup>101</sup> se muestra la construcción de un modelo donde se relacione el área neurocognitiva, la semántica, el lenguaje, el comportamiento social y el reconocimiento facial, todos con base en la demencia semántica.

Es uno de los trabajos con más interés para la neurociencia cognitiva en el mundo, debido a que en este trabajo de investigación se genera un modelo unificado, en donde se capturan las variables cognitivas y de la conducta que presentan los pacientes. Es aquí donde se muestra el grave deterioro semántico que generan diferentes afecciones mentales, como también que afecta la capacidad de producción verbal la cual está relacionada con diferentes comportamientos sociales y la presencia de demencia en las personas.

Este modelo de datos muestra la semántica del lenguaje, el comportamiento social y un reconocimiento facial todos estos relacionados con enfermedades cognitivas

---

<sup>101</sup> Ding, J.a,b, Chen, K.c, Liu, H.d, Huang, L.e, Chen, Y.a,f, Lv, Y.g, Yang, Q.h, Guo, Q.e , Han, Z.a , Lambon Ralph, M.A.i University, Beijing, China, Department of Neurosurgery, Baylor College of Medicine, Houston, TX, United States, Fudan University, Shanghai, China. "A unified neurocognitive model of semantics language social behaviour and face recognition in semantic dementia" [En línea] {05 de febrero de 2021} disponible en: [https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85085265414&origin=resultslist&sort=plf-f&src=s&st1=Language+for+Depression&st2=&sid=4507d7f667902348aa8d12e0318915fd&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Language+for+Depressio n%29&relpos=4&citeCnt=0&searchTerm=\)](https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85085265414&origin=resultslist&sort=plf-f&src=s&st1=Language+for+Depression&st2=&sid=4507d7f667902348aa8d12e0318915fd&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Language+for+Depressio n%29&relpos=4&citeCnt=0&searchTerm=)

y mentales, específicamente la demencia. Realizan un estudio de diferentes ciencias y universidades del mundo, en donde muestran su interés por entender cómo funciona el comportamiento social y el lenguaje cuando se está en presencia de personas que sufren de demencia o de alguna enfermedad mental.

Por tanto es fundamental entender a fondo el comportamiento que genera la depresión en el artículo titulado *DEPRESSION AND DISCLOSURE BEHAVIOR VIA SOCIAL MEDIA: A STUDY OF UNIVERSITY STUDENTS IN CHINA (COMPORTAMIENTO DE DEPRESIÓN Y DIVULGACIÓN A TRAVÉS DE LAS REDES SOCIALES: UN ESTUDIO DE ESTUDIANTES UNIVERSITARIOS EN CHINA)*<sup>102</sup> se muestra cómo el estrés severo y la depresión constituyen serios desafíos para la salud tanto personal como la salud pública en la sociedad, este artículo expone a las redes sociales como una tendencia que ha prevalecido entre los estudiantes universitarios y demás adultos jóvenes en distintos lugares del mundo. Las redes sociales se convierten en una plataforma para divulgar información personal que puede ser entendida como referencias y datos que muestran una nueva cara de la depresión.

Este artículo aporta una nueva vía para encontrar textos dirigidos al proceso de identificación y prevención de la depresión, dado que nos muestra tanto referentes teóricos sobre los diferentes trastornos mentales, su forma de estudio y también pautas de cómo seleccionar las posibles palabras que identifiquen conceptos de depresión, para prevenir el crecimiento o el desarrollo de nuevas etapas depresivas en la persona.

Así mismo es necesario analizar la toma de decisiones, en la investigación *DECISION MAKING WITH NATURAL LANGUAGE BASED PREFERENCES AND PSYCHOLOGY-INSPIRED HEURISTICS (TOMA DE DECISIONES CON PREFERENCIAS BASADAS EN EL LENGUAJE NATURAL Y HEURÍSTICAS INSPIRADAS EN LA PSICOLOGÍA)*<sup>103</sup> se expone un estudio referente a la toma de decisiones con preferencias basadas en el lenguaje natural y en la heurística de la psicología, todo esto para la compra de cualquier bien o servicio.

Este estudio es asistido por un sistema de software de procesamiento para la toma de decisiones que se caracterizan por la señalización de la intención de compra

---

<sup>102</sup> Maria Li Zou, Mandy Xiaoyang Li, Vincent Cho. "Depression and disclosure behavior via social media: A study of university students in China" [En línea] {05 de febrero de 2021} disponible en: (<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S2405844020302139>)

<sup>103</sup> Ingrid Nunes, Simon Miles, Michael Luck, Simone Barbosa, Carlos Lucena. "Decision making with natural language based preferences and psychology-inspired heuristics" [En línea] {05 de febrero de 2021} disponible en: (<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/abs/pii/S0952197615000603>)

en el lenguaje natural previo a la misma, en este artículo se propone una técnica de toma de decisiones automatizadas, que elige como opción un conjunto de datos disponibles para identificar las preferencias y prioridades que son expresadas en el lenguaje natural de la persona. Así mismo explora los términos y señala las expresiones de diferentes tomas de decisiones.

Este trabajo aporta una visión que ayuda a comprender cómo el lenguaje natural es un tipo de indicador de un sentimiento interno, ya sea para llevar a cabo una compra de un bien o servicio. Por tanto aporta a la identificación de la presencia y la confirmación de sintomatología depresiva presente en la persona respecto a un texto que esté tomando y que el procesamiento de lenguaje natural pueda resaltar la inclinación a algún posible síntoma de esta enfermedad mental.

Entendiendo la importancia del lenguaje natural, es fundamental la comprensión de textos, en TEXT COMPREHENSION: MODELS IN PSYCHOLOGY (COMPRESIÓN DE TEXTOS: MODELOS EN PSICOLOGÍA)<sup>104</sup> se trabajan los modelos de comprensión de textos que se encargan de describir y explicar los procesos involucrados en la comprensión y el recuerdo de la información verbal, toda esta comprensión textual se evalúa tanto en documentos fijos, como en documentos en proceso de escritura, todo esto en línea.

Se estudian los autores de estos textos mientras se lleva a cabo el proceso de escritura, se examinan de los mismos textos todos aquellos índices que representan el estado mental de la persona y también la semántica del texto, para así mismo tener una comprensión general tanto de las personas como del texto escrito. Este artículo científico presenta una relación con modelos computacionales tanto en la comprensión de textos, como en la producción de identificación de palabras claves de los textos.

En este trabajo se resaltan los modelos computacionales que analizan textos y presentan el paso a paso para identificar las palabras esenciales presentes en los textos permitiendo su comprensión y análisis, tanto de la persona como del texto que está escribiendo.

Teniendo en cuenta los modelos computacionales, comprender el lenguaje es primordial en el documento SOCIAL PSYCHOLOGY AND LANGUAGE (PSICOLOGÍA SOCIAL Y LENGUAJE)<sup>105</sup> se realiza una argumentación, en donde

---

<sup>104</sup> S.R.Goldman, M.B.W.Wolfe - International Encyclopedia of the Social & Behavioral Sciences {En línea}. "Text Comprehension: Models in Psychology" {10 de febrero de 2021} disponible en: (<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/B0080430767015424>)

<sup>105</sup> JJ Bradac, S.A. Reid - Encyclopedia of Language & Linguistics (Second Edition) {En línea}. "Social Psychology and Language" {10 de febrero de 2021} disponible en:



exponen a los seres humanos como criaturas que utilizan símbolos esenciales y exclusivamente para la comunicación entre los mismos, Más sin embargo los humanos también son animales psicológicos y es de ahí el surgimiento de la psicolingüística, la cual es una rama de la psicología en donde se relacionan los textos y los símbolos que utilizan con el significado de los mismos, siempre en relación con la sociedad, generando de ahí la sociolingüística.

En el documento se demuestra porque el lenguaje escrito es esencial para el análisis de las personas y el análisis social de las comunidades,este documento expone el lenguaje escrito como uno de los medios comunicativos más antiguos que posee la historia humana y es por esto mismo donde se encuentra la importancia que ha tenido este con el pasar de los años y los cambios de época y formas de vivir que se han presentado en las personas.

Al comprender la importancia del lenguaje, se resalta que en tiempos modernos al hablar del lenguaje es importante resaltar las redes sociales, en el artículo titulado *FEATURE ENGINEERING FOR DEPRESSION DETECTION IN SOCIAL MEDIA (INGENIERÍA DE FUNCIONES PARA LA DETECCIÓN DE DEPRESIÓN EN LAS REDES SOCIALES)*<sup>106</sup> Se realizó una investigación para la detección y el riesgo de presentar depresión, esto gracias a un conjunto de datos recopilados de diferentes mensajes escritos por usuarios de la red social Reddit.

El fin principal de este documento de investigación es mostrar la clasificación de los usuarios en dos grupos, un grupo que es en donde se presenta el caso de riesgo de presentar depresión y un caso de no riesgo a presentar depresión.

Este documento considera diferentes puntos de características para la detección de la depresión entre los usuarios que escribieron estos mensajes, toda mediante el procesamiento de mensajes de textos, que permiten examinar modelos de palabras, su gramática y los diferentes conjuntos de donde se valora la aplicabilidad de las características morfológicas de las palabras.

Además de esto se realizó una comparación entre los resultados obtenidos, adicionalmente aportan un modelo para la aplicabilidad de aprendizaje máquina y procesamiento de lenguaje en textos, realizando tareas de clasificación,

(<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/B0080448542013079>)

<sup>106</sup> Stankevich, M.,Isakov,V.,Devyatkin, D.,Smirnov,-Institute for Systems Analysis,Federal Research Center "Computer Science and Control" of RAS, Moscow, Russian Federation" Feature engineering for depression detection in social media"{En línea}{10 de febrero de 2021} disponible en:

(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85052014689&origin=resultslist&sort=plf-f&src=s&st1=depression+AND+machine+learning&st2=&sid=9fab56fd23cdd630985120d318a1db8&sot=b&sdt=b&sl=45&s=TITLE-ABS-KEY%28depression+AND+machine+learning%29&relpos=0&citeCnt=8&searchTerm=>)

identificación y señalamiento, teniendo como fundamento los grupos de riesgo de personas que presenten o no presenten sintomatología depresiva.

## 1.8 METODOLOGÍA

Para la realización del método automático que apoye en la identificación de síntomas de depresión por medio de análisis de textos escritos por adultos jóvenes, se trabajó con un apoyo parcial del marco de metodología general de las metodologías ágiles, dado que estas nos permiten adaptar la forma de trabajo de los proyectos, a una serie de condiciones establecidas consiguiendo así flexibilidad e inmediatez en la respuesta, el desarrollo y métodos de amoldar el proyecto para cumplir con el entorno solicitado.

Sin embargo por la naturaleza de este proyecto dado que no es construcción de software, se estipula una metodología a la medida del proyecto con un enfoque investigativo, en donde se presenta flexibilidad en el desarrollo del mismo ya que sus resultados son tanto para uso de ingeniería de sistemas, como para profesionales de salud y psicólogos interesados en el mismo.<sup>107</sup> Se especifican similitudes con metodologías ágiles en las fases de desarrollo, diseño, construcción y pruebas. Sin embargo para este proyecto se resaltó la fase de correcciones y afinamiento, para que la gestión de este trabajo sea continua y se perfeccione en cada ciclo de su funcionamiento.

Específicamente para la realización de la metodología se desarrollaron una serie de fases donde cada una depende de la anterior. Cada una de las fases presentadas en el diagrama de metodología lleva consigo una serie de tareas específicas las cuales van desde la construcción de la ontología, el diseño, selección y desarrollo de algoritmos para el procesamiento de lenguaje natural y de aprendizaje de máquina, luego de esto se valida el método usando métricas de desempeño. Existe una última fase metodológica en la cual se realizan las afinaciones y correcciones si son necesarias al proyecto, las cuales se identificaron en la validación del método.

---

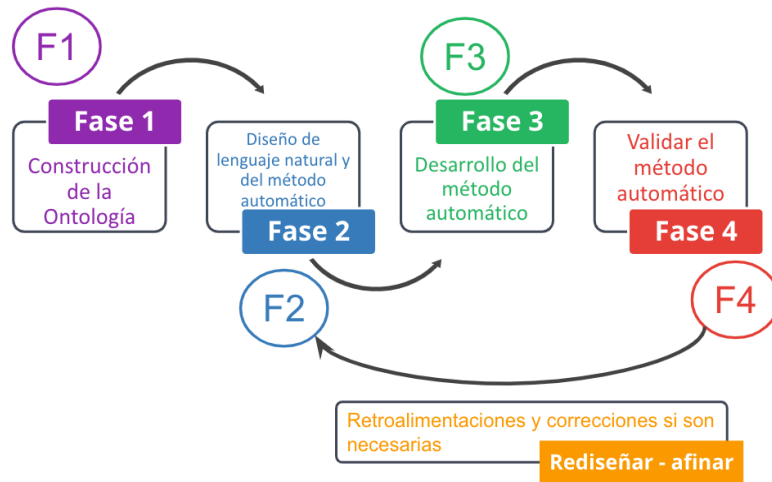
<sup>107</sup> Vanessa Rosselló Villán “Las metodologías ágiles más utilizadas y sus ventajas dentro de la empresa”. {En línea}. {10 de febrero de 2021} disponible en: (<https://www.iebschool.com/blog/que-son-metodologias-agiles-agile-scrum/>).



Se observa en la Figura 9. Diagrama general de metodología, las fases metodológicas que se llevarán a cabo para el desarrollo de este proyecto de grado.

**Figura 9**

*Diagrama General De Metodología*



*Figura 9. Diagrama general de metodología. Fuente: [El autor].*

*Nota:* En esta figura se muestran las fases metodológicas del proyecto.

### 1.8.1 FASE 1: CONSTRUCCIÓN DE LA ONTOLOGÍA

**Figura 10**

*Diagrama Metodológico Fase 1*



Figura 10. Diagrama metodológico fase 1: Construcción de la ontología. Fuente:[El autor]

Nota: En esta figura se aprecia la primera fase metodológica del proyecto.

La Figura 10. Diagrama metodológico fase 1: Construcción de la ontología representa el proceso implementado en donde se utilizó la guía "ONTOLOGY DEVELOPMENT 101"<sup>108</sup> que fue establecido por la Universidad de Stanford para unificar los criterios de definición y construcción de una ontología en ingeniería de software, esta ontología será usada como insumo para el método automático. Basados en esto se seguirán las siguientes etapas:

- a. Determinar el dominio y ámbito.
- b. Determinar la intención de uso.
- c. Reutilizar ontologías o vocabularios existentes.
- d. Enumerar los términos importantes del dominio.
- e. Definir jerarquía de clases.
- f. Crear las instancias.
- g. Validar con juicio de expertos en las áreas de psicología e ingeniería, pertenecientes a la Universidad Católica de Colombia, sobre la construcción, contenido final y aplicabilidad de la ontología propuesta.

Para la primera etapa metodológica de este trabajo, se estipulan como requerimientos necesarios, una computadora en donde el software Protégé mencionado anteriormente se ejecute y trabaje sin alteraciones, al igual que una investigación previa en literatura de salud mental centrada específicamente en literatura de conceptos relacionados con la presencia del trastorno mental de la depresión en las personas, para esto mismo se requiere el acceso a librerías de investigación de la literatura pertinente, ya sea en las diferentes bases de datos de

<sup>108</sup> Natalya F. Noy and Deborah L. McGuinness "Ontology Development 101: A Guide to Creating Your First Ontology". {En línea}. {15 de febrero de 2021} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).

la biblioteca disponibles por la Universidad Católica de Colombia.

## 1.8.2 FASE 2: DISEÑO DE LENGUAJE NATURAL Y DEL MÉTODO AUTOMÁTICO

Figura 11

*Diseño Metodológico Fase 2*

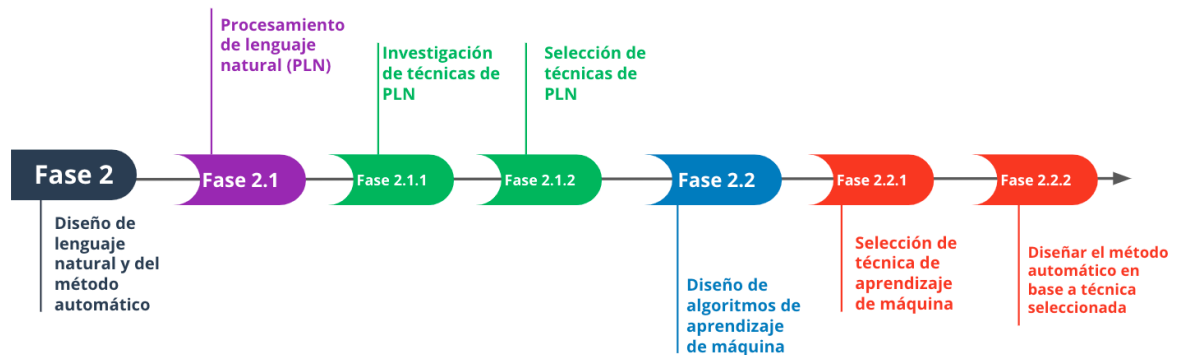


Figura 11. Diagrama metodológico fase 2: Diseño de lenguaje natural y del método automático. Fuente: [El autor]

*Nota:* En esta figura se muestra la segunda fase del proyecto, en la cual se ven aspectos del lenguaje natural y del método automático.

La Figura 11. Diagrama metodológico fase 2: Diseño de lenguaje natural y del método automático, describe el desarrollo de esta fase metodológica, en esta se relacionan dos sub-metodologías que son:

### **FASE 2.1: PROCESAMIENTO DE LENGUAJE NATURAL**

**2.1.1** Investigación de técnicas y librerías de software, para llevar a cabo el proceso de lenguaje natural.

**2.1.2** Selección de técnicas de procesamiento de lenguaje natural que se ajusten a la naturaleza de los datos.

## **FASE 2.2: DISEÑO DEL MÉTODO AUTOMÁTICO**

**2.2.1** Selección de la técnica o las técnicas de aprendizaje de máquina que deban ser utilizadas para la identificación de la sintomatología depresiva.

**2.2.2** Diseñar el método automático utilizando las distintas técnicas seleccionadas, que deben conectarse con la ontología y que consideren la validación del método automático.

Para la segunda etapa metodológica de este trabajo, se estipulan como requerimientos necesarios, un acceso por medio de la red o físico a literatura centrada en el procesamiento de la información por medio de la técnica de lenguaje natural, además de esto, un entorno de trabajo de acceso libre, como lo es Google Colab, para la investigación y pruebas primarias de las librerías de software que permita tanto procesos de lenguaje natural como de aprendizaje de máquina, para llevar a cabo una selección de las técnicas que más se ajusten a este trabajo.

Así mismo, el entorno de Google Colab que se estipula en esta etapa, se utilizará para el diseño preliminar y correcciones del método automático, donde se reúnan las técnicas de aprendizaje de máquina y de lenguaje natural.

## **1.8.3 FASE 3: EMPLEO DEL MÉTODO AUTOMÁTICO**

### **Figura 12**

*Diagrama Metodológico Fase 3*



*Figura 12. Diagrama metodológico fase 3: Empleo del método automático. Fuente: [El Autor]*

*Nota:* En esta figura se muestra la fase 3 del proyecto en el cual se ven aspectos como la

recolección de la información, el dataset con lenguaje natural y la aplicación en el ambiente de programación.

La Figura 12. Diagrama metodológico fase 3: Empleo del método automático representa el desarrollo de la tercera fase metodológica, en esta se trabaja con base a las siguientes tres fases:

**3.1.1** Recolección de la información de los usuarios.

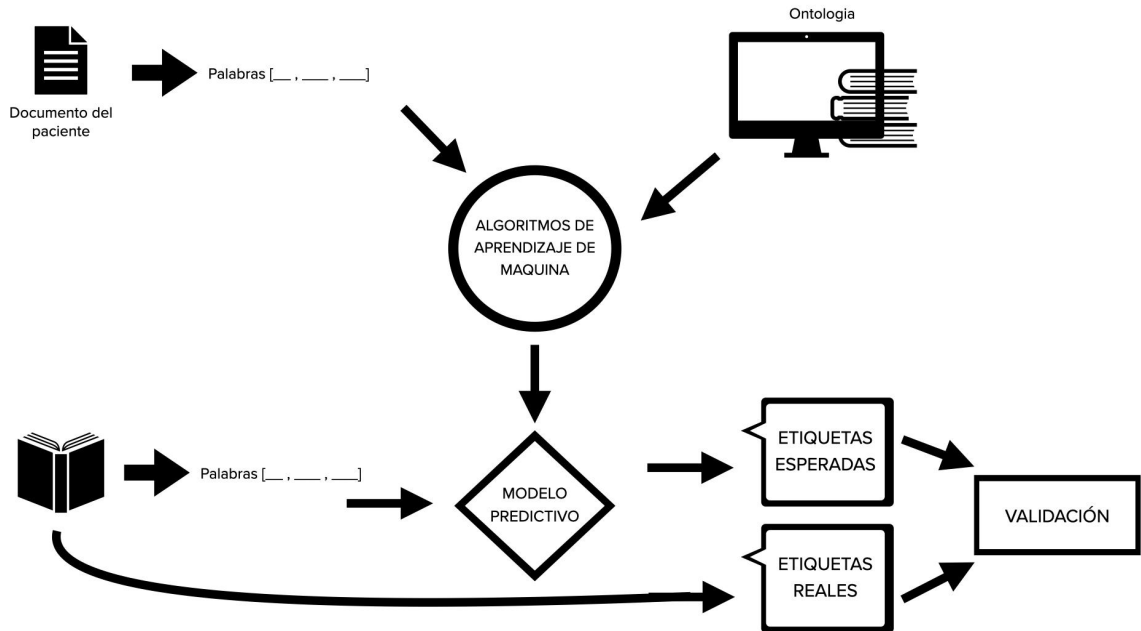
**3.1.2** Procesar los textos a su lenguaje natural para crear el dataset que será comprendido por el método automático.

**3.1.3** Aplicación en un ambiente de programación, que permita codificar el método automático diseñado en relación con la ontología.

Para la fase metodológica número 3, se desarrolló la Figura 13. Procedimiento de machine learning para análisis de textos, que permite comprender de una mejor manera lo esperado que realice el método automático, de igual manera presenta una explicación por etapas más detallada:

**Figura 13**

*Procedimiento de Machine Learning para Análisis de Textos*



*Figura 13. Procedimiento de machine learning para análisis de textos. Fuente:[El Autor]*

*Nota:* En este figura se muestra el procedimiento utilizado para el análisis de textos utilizando machine learning.

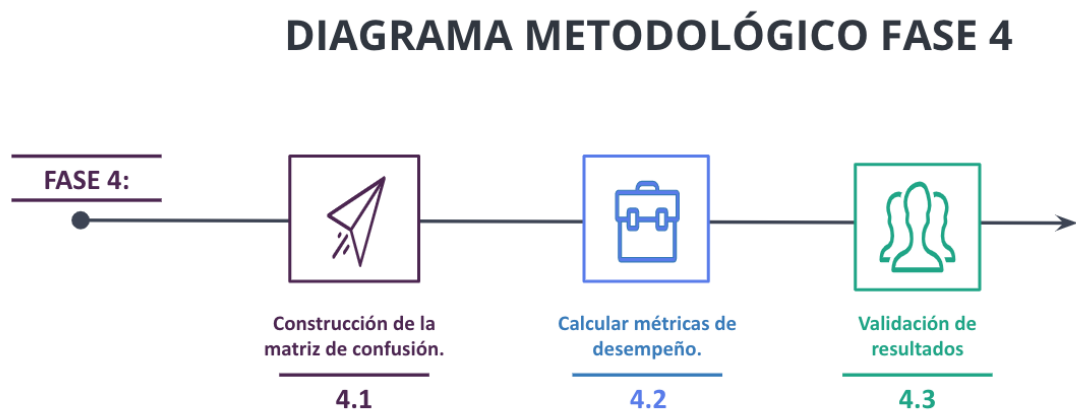
Para la tercera etapa metodológica de este trabajo de grado, se estipulan como requerimientos necesarios, un dataset creado previamente en donde se almacene la información de los usuarios y sus respectivos textos, a trabajar en el método automático desarrollado, al igual que los diferentes documentos aprobados por el usuario como lo son la Ley Habeas Data y el consentimiento de información, estos se pueden encontrar en la sección de Anexos de este documento, como Anexo 1.

Además, como complemento para los requerimientos para esta fase metodológica es necesario contar con una o dos máquinas de cómputo, donde el método automático desarrollado en la segunda fase metodológica de este trabajo se encuentre implementado y listo para uso, además de que esto, la máquina de cómputo ya debe contar con toda la información disponible en sus archivos y memoria respecto a la ontología de términos de sintomatología depresiva, las librerías de lenguaje natural y de aprendizaje de máquina.

#### 1.8.4 FASE 4: VALIDAR EL MÉTODO AUTOMÁTICO

**Figura 14**

*Diagrama Metodológico Fase 4*



*Figura 14. Diagrama metodológico fase 4: Evaluar el método automático. Fuente: [El autor]*

*Nota:* En esta figura se puede apreciar el diagrama metodológico de la fase 4 el cual consiste en evaluar el método automático.

La última fase metodológica del proyecto de grado, se representa en la Figura 14. Diagrama metodológico fase 4: Validar el método automático, es en esta fase en donde se válida la funcionalidad y viabilidad de la aplicación de este método, en los trabajos futuros del departamento de psicología de la Universidad Católica de Colombia.

Para el desarrollo de esta fase metodológica, se trabaja con base a las siguientes tres fases:

- 4.1 Construcción de la matriz de confusión.
- 4.2 Calcular las métricas de desempeño: precisión, Recall (“Recordar”), F1-score (“Puntuación F1”) y exactitud, al método automático realizado.
- 4.3 Validación técnica de los resultados obtenidos y validación por juicio de expertos de salud mental de la Universidad Católica de Colombia.

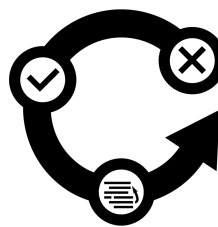
Para la cuarta etapa metodológica de este trabajo, se estipulan como requerimientos necesarios, herramientas de software especializadas en construcciones de métricas de desempeño específicas como las mencionadas anteriormente.

Además de estas herramientas de software, es necesario como requerimiento de esta fase metodológica, contar con una serie de reuniones y charlas de retroalimentación por parte de docentes, expertos en salud mental y demás interesados en el proyecto, de la Universidad Católica de Colombia o de otras instituciones externas, pero afines a la temática del proyecto, para que con esto se pueda dar una validación del mismo y empezar con tareas de rediseño y afinación del trabajo.

### 1.8.5 REDISEÑAR - AFINAR

#### Figura 15

*Rediseño y Afinación de Metodología*



*Figura 15. Rediseño y afinación de metodología. Fuente:[El Autor]*

*Nota:Esta figura muestra el rediseño y afinación de la metodología.*

La Figura 15. Rediseño y afinación de metodología, es una representación simbólica de la etapa de rediseño y afinación del proyecto, es una fase metodológica propia de los proyectos de software, dado que acogen a una serie de retroalimentaciones y correcciones en el proyecto, ya sea que se identificaron en la fase 4 de la validación del método o también en el proceso de pruebas donde posibles errores puedan ser identificados y trabajados.

Esta etapa de retroalimentación debe seguir el método científico, y un análisis que permita realizar diferentes ajustes al proyecto, ya sea desde las fases metodológicas de diseño y desarrollo, para volver a salir el método automático a ejecución.

Para la última etapa metodológica de este trabajo, los requerimientos que se estipulan, son los mismos requerimientos mencionados anteriormente en la fase 4, la fase de validación del método automático. Esta va de la mano con esta última fase dada que en esta es en donde de las diferentes retroalimentaciones o afinaciones señaladas en la fase 4, se empiezan a trabajar y se vuelve así al ciclo de trabajo de los proyectos que presenta esta metodología, basándose en el esquema de metodologías ágiles.

## 2. ONTOLOGÍA APLICADA A SINTOMATOLOGÍA DE DEPRESIÓN

Para poder realizar la construcción de la ontología respectiva a la sintomatología depresiva que será utilizada como insumo en el método automático desarrollado en este proyecto, se parte desde la documentación y guía "*ONTOLOGY DEVELOPMENT 101*"<sup>109</sup>, la cual fue establecida por la Universidad de Stanford, para llevar a cabo la unificación de criterios para la definición y la construcción de una ontología en Ingeniería de Software.

El desarrollo de este capítulo parte de la premisa de que las ontologías se han convertido en componentes centrales de muchas aplicaciones, incluyendo este método automático para la identificación de sintomatología depresiva. Es importante responder a la pregunta de "¿Por qué se quiere desarrollar una ontología?". Algunas razones del porqué se fundamentan en el principio de que las ontologías nos permiten compartir el entendimiento de la información en forma estructurada permitiendo que sea clara y entendible entre personas o software.

Las ontologías permiten la reutilización del conocimiento del dominio, para que con el mismo, se logren realizar supuestos a relaciones entre diferentes conocimientos que estén incluidos en el mismo dominio, así también se permite establecer supuestas relaciones o demarcar las separaciones del conocimiento que estén agrupados en un dominio. El documento "*ONTOLOGY DEVELOPMENT 101*" aporta una guía de cómo abordar el problema de construir la ontología, este a su vez presenta una metodología interna para crear las ontologías en el *software Protégé*, para que así las ontologías desarrolladas se basen en sistemas declarativos que representan el conocimiento.

---

<sup>109</sup> Natalya F. Noy and Deborah L. McGuinness "Ontology Development 101: A Guide to Creating Your First Ontology". {En línea}. {05 de marzo de 2021} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).



Es importante resaltar que los autores de este documento han especificado todo el desarrollo y mantenimiento de ontologías para una serie de entornos diferentes, como lo pueden ser *Protégé*, *Ontolingua* y *Chimaera*. En este proyecto trabajaremos con el software *Protégé*, para la construcción de la ontología, dado que es el identificado en la documentación para la construcción de ontologías en cualquier sistema centrado de objetos.

Siguiendo el proceso metodológico que especifica la documentación revisada para la formulación de ontologías y, también, el diagrama procedimental estipulado para este trabajo de grado en su fase 1 (*ver apartado metodológico*), se procede a establecer una serie de etapas que permite observar el correcto desarrollo de la ontología de palabras que identifica la sintomatología depresiva, que es el objetivo central de este trabajo.<sup>110</sup>

## **2.1 DETERMINAR EL DOMINIO Y ÁMBITO.**

Según la documentación "*ONTOLOGY DEVELOPMENT 101*" en la primera etapa metodológica para la realización de ontologías se debe determinar el dominio y el alcance de la ontología, la cual se implementó para dar respuesta a las preguntas fundamentales del desarrollo, que son las siguientes:

- ¿Cuál es el dominio que cubrirá la ontología?
- ¿Para qué vamos a utilizar la ontología?
- ¿Para qué tipo de preguntas debería dar respuesta la información de la ontología?
- ¿Quién usará y mantendrá la ontología?

Las respuestas a estas preguntas se van moldeando y generando a medida que este proyecto fue avanzando, es importante considerar que este es un proyecto interdisciplinario en el cual tanto profesionales de la psicología como profesionales de la ingeniería de sistemas entrarán en valor, por lo tanto el modelo, alcance y dominio de esta ontología debe ser de viabilidad para cualquiera de los profesionales mencionados anteriormente, es por eso que en las respuestas a las preguntas que se presentan a continuación se señala un interés tanto de la salud mental, como de la parte técnica y de la viabilidad del software.

### **❖ ¿Cuál es el dominio que cubrirá la ontología?**

El dominio que cubrirá la ontología desarrollada se ubica en el dominio específico

---

<sup>110</sup> Ontology101 - Protégé Wiki. (s. f.). Protégé Wiki.Stanford.Edu. {En línea}. {05 de marzo de 2021} disponible en: <https://Protégéwiki.stanford.edu/wiki/Ontology101>

del conocimiento a las palabras que presenten una relación directa con la posible identificación de sintomatología depresiva, los conceptos (cosas o palabras) deben señalar directamente a cualquiera de los síntomas principales de la depresión que puede sufrir una persona joven adulta, es decir de 18 a 26 años<sup>111</sup>. El dominio de esta ontología va relacionada directamente con el tipo de usuario (paciente) al cual va destinado el uso y aplicación de este método automático en el señalamiento de la posible presencia de sintomatología depresiva que actualmente está presentando.

#### ❖ **¿Para qué vamos a utilizar la ontología?**

La ontología va a utilizarse de dos maneras. La primera manera en la que se utilizara la ontología acá desarrollada, será como un insumo en forma de fuente de datos (palabras), con el cual el método automático realizará la respectiva clasificación e identificación de las mismas palabras en un texto generado por un usuario (paciente). Y la segunda manera de utilizar la ontología, será como el referente teórico y de soporte a las diferentes clasificaciones ó sesgos que realice el método automático.

Dado que la ontología debe proporcionar y mantener conceptos claves para la clasificación en síntomas de depresión respecto a palabras escritas por el usuario, la realización de la ontología se presenta con base a información veraz y clave, esta información fue extraída de diferentes fuentes de literatura psicológica, respecto a síntomas de depresión, clasificación de la depresión, relación entre palabras escritas y sentimientos encontrados. También se señala que diferentes palabras que se encuentran en la ontología fueron las mismas que se señalan en los diferentes documentos del apartado del estado del arte en el presente proyecto.

#### ❖ **¿Para qué tipo de preguntas debería dar respuesta la información de la ontología?**

El tipo de preguntas al que debe dar respuesta la información de la ontología, va directamente relacionado con un formulario de preguntas establecido por expertos de la psicología clínica, basándose en el referente teórico de documentos y pruebas psicológicas como los son: El Inventario de Depresión de Beck-II (BDI-II)<sup>112</sup>, la escala Zung de depresión (específicamente enfocada en el apartado de

---

<sup>111</sup> Colombia, S. P. S. D. M. Y. (s. f.). Páginas - Ciclo de Vida. [minsalud.gov.co/](https://www.minsalud.gov.co/). {En línea}. {05 de Marzo de 2021} disponible en:

[https://www.minsalud.gov.co/proteccion-social/Paginas/cicloVida.aspx#:~:text=6%20%2D%2011%20a%C3%B1os\)-Adolescencia%20\(12%20%2D%2018%20a%C3%B1os\).o%20mas\)%20envejecimiento%20y%20vejez](https://www.minsalud.gov.co/proteccion-social/Paginas/cicloVida.aspx#:~:text=6%20%2D%2011%20a%C3%B1os)-Adolescencia%20(12%20%2D%2018%20a%C3%B1os).o%20mas)%20envejecimiento%20y%20vejez)

<sup>112</sup> Psi.uba.ar. 2021. {En línea}. {15 de Marzo de 2021} disponible en:

[\[https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios\\_catedras/obligatorias/070\\_psicoterapias1/material/inventario\\_beck.pdf\]](https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios_catedras/obligatorias/070_psicoterapias1/material/inventario_beck.pdf)

validaciones)<sup>113</sup>, el manual diagnóstico y estadístico DSM 5<sup>114</sup> y CIE 10<sup>115</sup>.

El desarrollo del formulario de preguntas que fueron establecidas por los expertos de la psicología clínica, se desarrolló por medio de varias reuniones con expertos de la Salud Mental tanto practicantes del departamento de psicología, como profesionales que ejercen en el mismo.

## Figura 16

### *Reuniones Remotas Realizadas*

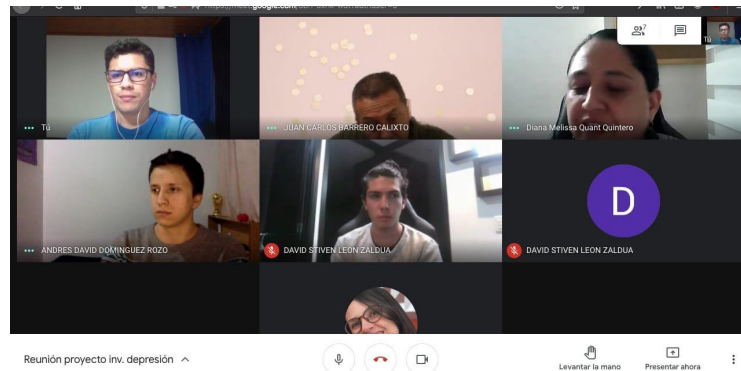


Figura 16. Captura de pantalla reuniones remotas realizadas. Fuente: [El Autor]

*Nota:* En la figura se observa cómo se realizaron reuniones remotas para la consolidación de conceptos para la ontología,

Como se puede detallar en la Figura 16. *Captura de pantalla reuniones remotas realizadas*, se llevaron a cabo reuniones masivas las cuales tenían tareas específicas, en un primer momento se empezó con la determinación del uso de las preguntas, seguido de la clasificación de las preguntas, luego qué tipo de palabras se deben implementar y cuál debe ser la redacción de las preguntas para generar más información detallada en las respuestas.

<sup>113</sup> Escala de Autoevaluación para la Depresión de Zung (SDS). [En línea] Disponible en: Mentalhealthministries.net website: {En línea}. {15 de Marzo de 2021} disponible en: [http://www.mentalhealthministries.net/resources/flyers/zung\\_scale/zung\\_scale\\_sp.pdf](http://www.mentalhealthministries.net/resources/flyers/zung_scale/zung_scale_sp.pdf)

<sup>114</sup> di AGnósticos, C. (s/f). Guía de consulta. Disponible en: Edu.co website: {En línea}. {15 de Marzo de 2021} disponible en: <https://www.eafit.edu.co/ninos/reddelaspreguntas/Documents/dsm-v-guia-consulta-manual-diagnostico-estadistico-trastornos-mentales.pdf>

<sup>115</sup> Diagnósticos, C.-10-E. (s/f). MANUAL DE CODIFICACIÓN. [En línea] Disponible en Gob.es website: {En línea}. {15 de Marzo de 2021} disponible en: [https://www.mscbs.gob.es/estadEstudios/estadisticas/normalizacion/CIE10/UT\\_MANUAL\\_DIAG\\_2016\\_prov1.pdf](https://www.mscbs.gob.es/estadEstudios/estadisticas/normalizacion/CIE10/UT_MANUAL_DIAG_2016_prov1.pdf)

Todos estos procesos se validan en reunión final, la cual especificó cuáles son las preguntas finales que serán establecidas para la ejecución y desarrollo de este método automático en las consultas. Mencionando a los profesionales del área de salud mental y profesionales de ingeniería de sistemas y computación, que hicieron posible esta tarea encontramos:

El equipo que propuso las preguntas estuvo compuesto por:

- Laura Orozco (*Estudiante Psicología*)
- Johana Orozco (*Estudiante Psicología*)
- Andrés Domínguez (*Estudiante Ing. Sistemas y Comp.*)
- Cristian Robayo (*Estudiante Ing. Sistemas y Comp.*)
- Josue Vergara (*Estudiante Ing. Sistemas y Comp.*)
- Juan Barrero (*Profesional Ing. Sistemas*)
- Eliana Ivette Ortiz (*Profesional Psicología*)

El equipo que validó estas preguntas estuvo compuesto por:

- Diana Melissa Quant (*Profesional Psicología*)
- Ricardo Santa (*Profesional Ing. Sistemas*)
- Holman Bolivar (*Profesional Ing. Sistemas*)

Algunas de las preguntas desarrolladas que buscaban con su respuesta de manera abierta identificar la presencia o la ausencia de síntomas de depresión en la persona, se muestran en la siguiente tabla:

**Tabla 2**

*Lista De Preguntas Generadas*

CRITERIOS	PREGUNTA	FORMATO
<b>ESTADO DE ÁNIMO ANHEDONIA</b>	1. Piensa en algunas situaciones, en las dos últimas semanas, en las que te hayas sentido triste o desanimado(a) y descríbenos qué pensabas y qué sentías	ESCRITO

<b>AGITACIÓN/ RETRASO PSICOMOTOR FATIGA DETERIORO EN LOS ÁMBITOS DISMINUCIÓN DE LA CONCENTRACIÓN</b>	2. ¿En las dos últimas semanas te has sentido con menos energía (de lo normal) o con dificultades para concentrarte en las actividades diarias? Describe nos más detalladamente... ¿En las últimas semanas te has sentido más acelerado o lento (de lo normal) al enfrentar las actividades diarias? Describe nos más detalladamente...	ESCRITO
...	...	...

Tabla 2. Lista de preguntas generadas. Fuente:[El Autor]

*Nota: En estas tabla se muestran las preguntas que fueron generadas.*

Generalizando, el tipo de preguntas al que debe dar respuesta la información de la ontología, son aquellas preguntas y respuestas que van directamente relacionadas con el estado de ánimo y diferentes valores, como lo son la condición física, la condición mental, el apetito, la motivación en las acciones diarias y las relaciones interpersonales que tenga la persona (consultante) a tratar.

Para apreciar minuciosamente, cada una de las preguntas a las que se hace referencia en este apartado, y que debe dar respuesta con la información de la ontología, diríjase al anexo 3.

### ❖ ¿Quién usará y mantendrá la ontología?

La ontología desarrollada para este proyecto va a ser usada por la facultad de Psicología y la facultad de Ingeniería, ambas pertenecientes a la Universidad Católica de Colombia, a su vez por profesionales y estudiantes interesados en el proyecto, en la usabilidad ó métodos de aplicación del mismo.

Este al ser un proyecto interdisciplinario las labores de desarrollo, uso y mantenimiento fueron trabajadas por diferentes estudiantes de ambas facultades que centraron sus esfuerzos en permitir que la ontología logre ser utilizada por el entorno de desarrollo Google Colab y por las diferentes librerías físicas para el manejo y extracción de información.

Para la parte técnica, se destaca que el entorno de trabajo de Google Drive mantiene, asegura y provee mecanismos para el desarrollo de proyectos futuros y demás avances que se realicen en este proyecto o relacionados con el mismo.

## 2.2 DETERMINAR LA INTENCIÓN DE USO.

La ontología que se desarrolló en este proyecto, tiene diferentes intenciones de

uso, no solo se aplicaría para ser el insumo del método automático trabajado, sino que a su vez tenía como intención de uso, compartir el entendimiento común de la estructura de la información sobre conceptos que identifiquen sintomatología depresiva en las personas<sup>116</sup> y diferentes agentes de software cómo Protégé.

Entre las diferentes intenciones de uso de esta ontología, se encontraba el de permitir encaminar este proyecto a la reutilización del conocimiento respecto a los conceptos que identifican la sintomatología depresiva, a su vez aplicar estos conocimientos en un dominio diferente, donde con estos se espera clasificar y trabajar las relaciones para un análisis e identificación de la depresión, todo esto en una sesión médica impartida en el departamento clínico, de la facultad de psicología de la Universidad Católica de Colombia.

### 2.3 REUTILIZAR ONTOLOGÍAS O VOCABULARIOS EXISTENTES.

El proceso de reutilización de ontologías y vocabularios existentes es una variable que debe considerarse, dado que con este proceso se permite verificar y refinar el conocimiento construido apoyándose en el conocimiento construido por otras personas ó investigaciones, para esa tarea en particular utilizamos una serie de ontologías clínicas, y de uso investigativo, alojadas en diferentes bases de datos académicas. Al estar estas ontologías en el marco de investigación hace que las mismas puedan ser reutilizables, con un acceso en la web ó en la literatura.

Entre las diferentes bibliotecas y bases de datos académicas de la universidad<sup>117</sup>, encontramos las siguientes ontologías que permiten determinar y establecer un marco, tanto metodológico como conceptual, para delimitar una ontología enfocada en el análisis del sentimiento por medio de datos sociales.

En este sentido, resultaron sumamente valiosos los siguientes trabajos, descritos en la Tabla 3. Ontologías y vocabularios existentes que apoyan en la construcción.

#### Tabla 3

*Ontologías y vocabularios existentes que apoyan en la construcción*

#### Ontologías y vocabularios existentes

---

<sup>116</sup> Se refiere a personas, como a todos los involucrados en ese proyecto tanto ingenieros de sistemas, como psicólogos y expertos de la salud mental.

<sup>117</sup>Bases de Datos. (s. f.). U Católica Bases de Datos. {En línea}. {20 de marzo de 2021} disponible en:

[https://login.ucatolica.basesdedatosezproxy.com/public/menu\\_ext.htm#cat%C3%A1logo](https://login.ucatolica.basesdedatosezproxy.com/public/menu_ext.htm#cat%C3%A1logo)

Documento	Aporte
Ontology-Based Approach to Social Data Sentiment Analysis: Detection of Adolescent Depression Signals <sup>118</sup>	El objetivo de este estudio fue la definición de una ontología y terminología de la depresión en adolescentes, que fuese implementada como un marco referencial para analizar los datos de las redes sociales y evaluar con lógica las distintas descripciones entre clases, al igual que la aplicabilidad de esta ontología en el área del análisis de sentimientos.
Development and Evaluation of an Adolescents' Depression Ontology for Analyzing Social Data <sup>119</sup>	Este estudio tiene como objetivo desarrollar y evaluar una ontología para la depresión en los adolescentes, que será utilizada para recopilar y analizar datos sociales. Esta ontología se desarrolla de acuerdo con la metodología de 'ONTOLOGY DEVELOPMENT 101'. Los conceptos de esta ontología se extrajeron de guías clínicas y de literatura relacionada. Esta ontología está compuesta por cinco sub-ontologías que representan factores de riesgo, signos y síntomas, medición, resultado diagnóstico y atención de manejo.
From Affective Science to Psychiatric Disorder: Ontology as a Semantic Bridge <sup>120</sup>	El aporte de esta ontología, se basa en la contribución, propuesta de un marco ontológico para capturar explícitamente las complejas interrelaciones entre las entidades afectivas y los trastornos psiquiátricos, de las personas, con el fin de facilitar el mapeo y la integración entre la ciencia afectiva y el diagnóstico psiquiátrico. También se incluyen en el estudio aclaraciones a las categorías de diagnóstico psiquiátrico, sistemas de información clínica, la integración y traducción de resultados de investigación entre disciplinas.
Mobile cloud-based depression diagnosis using an ontology and a Bayesian network <sup>121</sup>	Este estudio pretende obtener un diagnóstico preciso de depresión, para esto desarrolla un modelo de ontología basado en la terminología utilizada para describir la depresión, para esto utiliza una red bayesiana que logra inferir la probabilidad de deprimirse. Los resultados que se muestran en este método puede ser útil para inferir un diagnóstico de depresión.
Ontology-based ubiquitous monitoring and treatment against depression <sup>122</sup>	El objetivo de esta ontología y artículo, fue el de proponer un modelo de tratamiento y monitoreo ubicuo mejorado y representado con una ontología, para ayudar a las personas a superar los desafíos de los trastornos mentales. En este proyecto es implementado para facilitar la

<sup>118</sup> Jung, H., Park, H. A., & Song, T. M. (2017). Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals. *Journal of medical internet research*, 19(7), e259. [Consultado el 20 de marzo de 2021].

<sup>119</sup> Jung, H., Park, H. A., & Song, T. M. (2016). Development and Evaluation of an Adolescents' Depression Ontology for Analyzing Social Data. *Studies in health technology and informatics*.

<sup>120</sup> Larsen, R. R., & Hastings, J. (2018). From affective science to psychiatric disorder: Ontology as a semantic bridge. *Frontiers in psychiatry*, 9, 487. [Consultado el 20 de Marzo de 2021].

<sup>121</sup> Chang, Y. S., Fan, C. T., Lo, W. T., Hung, W. C., & Yuan, S. M. (2015). Mobile cloud-based depression diagnosis using an ontology and a Bayesian network. *Future Generation Computer Systems*, 43, 87-98. [Consultado el 20 de Marzo de 2021].

	TCC en línea, para el tratamiento de la depresión en la etapa actual.
Ontology-Based Inference for Supporting Clinical Decisions in Mental Health <sup>123</sup>	Este estudio realiza una revisión de la literatura que involucra artículos científicos y las guías más actuales de la OMS en salud mental, posteriormente trasladamos este conocimiento a un modelo computacional formal, construyendo la ontología propuesta. Con diferentes motores de clasificaciones se desarrolla un artefacto computacional semántico para el almacenamiento y generar conocimiento para ayudar a los profesionales de la salud mental en las decisiones clínicas.

Tabla 3. Ontologías y vocabularios existentes que apoyan en la construcción. Fuente: [El Autor]

*Nota:* En la tabla anterior se muestran las ontologías y vocabularios que fueron de gran utilidad en la construcción de la ontología.

Los anteriores, aunque son sistemas representativos del conocimiento respecto a la depresión, su sintomatología y diferentes metodologías para su identificación, no muestran, específicamente, un listado de palabras que señalen la presencia de sintomatología depresiva en la persona que escribe.

Es por eso que para el desarrollo de esta ontología, recurrimos también a vocabularios y diferentes documentaciones como los cuestionarios psicológicos tales como, cuestionario BDI, La escala Zung de depresión (En esta ,específicamente, se centra en a sus validaciones) y al manual diagnóstico estadístico DSM5 y el cuestionario CIE 10.<sup>124</sup>

## 2.4 ENUMERAR LOS TÉRMINOS IMPORTANTES DEL DOMINIO.

Se desarrolla una lista de todos los términos sobre los cuales se realizan las respectivas declaraciones y clasificaciones de las palabras específicas que determinan la posible presencia de depresión en una persona, es de resaltar que los siguientes términos van relacionados con un tipo específico de sintomatología.

Aunque es complejo llevar a cabo la clasificación de un término específico, para cada síntoma que presenta la presencia de depresión en la persona, se determina que existen relaciones simultáneas entre categorías de depresión y términos. El paso siguiente fue desarrollar la jerarquía de clases y definir las propiedades de

<sup>122</sup> Hu, B., Hu, B., Wan, J., Dennis, M., Chen, H. H., Li, L., & Zhou, Q. (2010). Ontology-based ubiquitous monitoring and treatment against depression. *Wireless communications and mobile computing*, 10(10), 1303-1319. [Consultado el 20 de Marzo de 2021].

<sup>123</sup> Yamada, D. B., Bernardi, F. A., Miyoshi, N. S. B. (2020, June). Ontology-Based Inference for Supporting Clinical Decisions in Mental Health. In *International Conference on Computational Science* (pp. 363-375). Springer, Cham. [Consultado el 20 de Marzo de 2021].

<sup>124</sup> Las palabras extraídas fueron validadas en el Tesoro de la UNESCO y el diccionario de la Real Academia de la Lengua Española.



los conceptos es decir qué relación lleva estrechamente cada concepto con su categoría principal de sintomatología depresiva y cuáles de estos mismos se encuentran entrelazados con otras categorías.

La lista de términos, conceptos y adjetivos que se muestran en la Tabla 4. Enumeración y clasificación de conceptos. fue presentada en un primer momento como referente para el desarrollo de la ontología. Está tras varias correcciones y afinaciones, fue finalmente estipulada como la lista de conceptos más indicada y correspondiente a las diferentes clasificaciones de depresión, para este proyecto de grado.

**Tabla 4**

*Enumeración y clasificación de conceptos.*

Síntomas depresivos	General	Agrupación	Conceptos	Derivaciones
1	Ánimo	Estado de ánimo deprimido	Desaliento, Pena, Tristeza ...	Tengo un vacío, me ahogo, Oscuridad, Estoy en el fondo...
2	Apatía	Disminución de interés o placer	No siento deseo, Indiferente, Sensible ...	Nada motivado, Desesperado, Desesperanzado, Reducido ...
3	Alimentación	Pérdida de peso	Indiferente, Dejatez, Descuidado ...	Desinteresado, Desánimo, Desganado...
		Aumento de peso	Apetito, Hambre, Inapetente...	Fatigado, Cansado, Intranquilo ...
...				
9	Descanso	Hipersomnia	Debilidad, Desaliento, Cotidiano...	Mamado, debilitado, flojo ...
		Insomnio	Sensible, Hábito, Susceptible ...	Igual, Hábitos, no duermo ...

*Tabla 4. Enumeración y clasificación de conceptos. Fuente:[El Autor]*

*Nota:* En esta tabla se muestra la enumeración clasificación de conceptos clave en el proyecto.

La tabla 4 representa un fragmento de la tabla original, en la que se clasifican los conceptos referentes a cada uno de los nueve síntomas principales para la determinación e identificación de la depresión. Se hace esta corta tabla en representación de la tabla general debido a su magnitud y extensión, la tabla original es añadida a este trabajo como el anexo número 4.

## 2.5 DEFINIR JERARQUÍA DE CLASES.

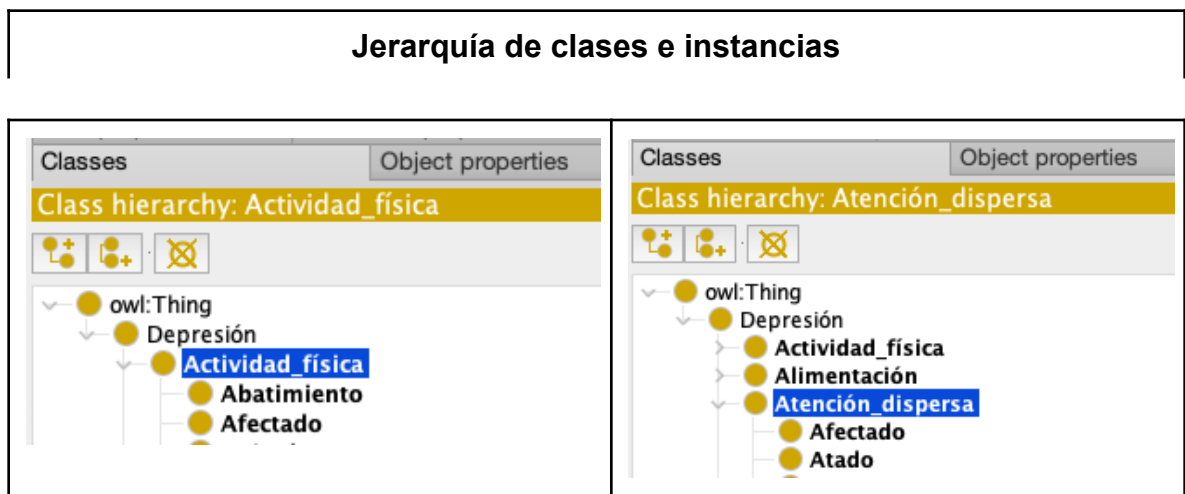
Para el proceso de definir las clases y la jerarquía de los conceptos en esta ontología, fue utilizado uno de los enfoques específicos de los diferentes mencionados en el documento “Ontology Development 101”.

El enfoque seleccionado es de arriba hacia abajo (*A top-down*) este es un proceso de desarrollo que comienza con la definición de los conceptos más generales del dominio, que en este caso es la depresión y posteriormente se especializa en los conceptos o su clasificación que son los diferentes síntomas para la detección de la depresión, y por último estipulamos los conceptos finales aquellos con los cuales trabajara directamente el método automático.

A continuación, se presenta el desglose realizado entre los diferentes niveles de generalidad en los síntomas de depresión y los conceptos que posiblemente los identifiquen en una persona, estas figuras fueron realizadas en el entorno Protégé.

**Tabla 5**

*Jerarquía De Clases e Instancias Trabajadas En Protégé*



<p>Classes   Object properties</p> <p><b>Class hierarchy: Alimentación</b></p>	<p>Classes   Object properties</p> <p><b>Class hierarchy: Cargo_de_conciencia</b></p>
...	...

Tabla 5. Jerarquía de clases e instancias trabajadas en Protégé. Fuente:[El Autor]<sup>125</sup>

Nota: En la tabla anterior se muestra la jerarquía de clases e instancias trabajadas en Protégé.

**Tabla 6**

*Jerarquía de Propiedades Del Objeto Trabajadas en Protégé*

<b>Jerarquía de propiedades del objeto</b>	
<p>Classes   Object properties</p> <p><b>Object property hierarchy: Actividad física</b></p>	
...	.....

<sup>125</sup> Para consultar las nueve jerarquía de clases e instancias que componen a esta Ontología, diríjase al anexo 5, en la sección “Class hierarchy”

Tabla 6. Jerarquía de propiedades del objeto trabajadas en Protégé.<sup>126</sup>

*Nota:* En la anterior tabla se muestran las propiedades del objeto, como las relaciones y clases de la ontología.

Las tablas 5 y 6 representan las relaciones, conceptos y clases, pertenecientes a la ontología desarrollada en *Protégé*. Se hace esta corta tabla en representación de las clases e instancias debido a su magnitud y extensión, la Ontología final es añadida a este trabajo como el anexo número 5.

## 2.6 CREAR LAS INSTANCIAS.

Para llevar a cabo el desarrollo de la creación de las respectivas clases en esta ontología, se seguía el mismo orden jerárquico de arriba hacia abajo, es decir si una clase A es una superclase de B entonces cada instancia de B es también una instancia de A.

Es por esto que aunque el árbol final era jerárquico existían instancias que tenían una relación directa con la otra clase, es decir existen conceptos específicos que podrían ser los señalados para uno o más tipos de sintomatología depresiva, es por esto que el grafo generado en la figura 23 es similar a una telaraña, dado que las ranuras de este cuerpo están entrelazadas por algunos conceptos entre sí.

### Figura 17

*Gráfico De La Ontología Exportada De Protégé*

---

<sup>126</sup> Para consultar la jerarquía de propiedades de los objetos que componen a esta Ontología, diríjase al anexo 5, en la sección "*Object Properties*"

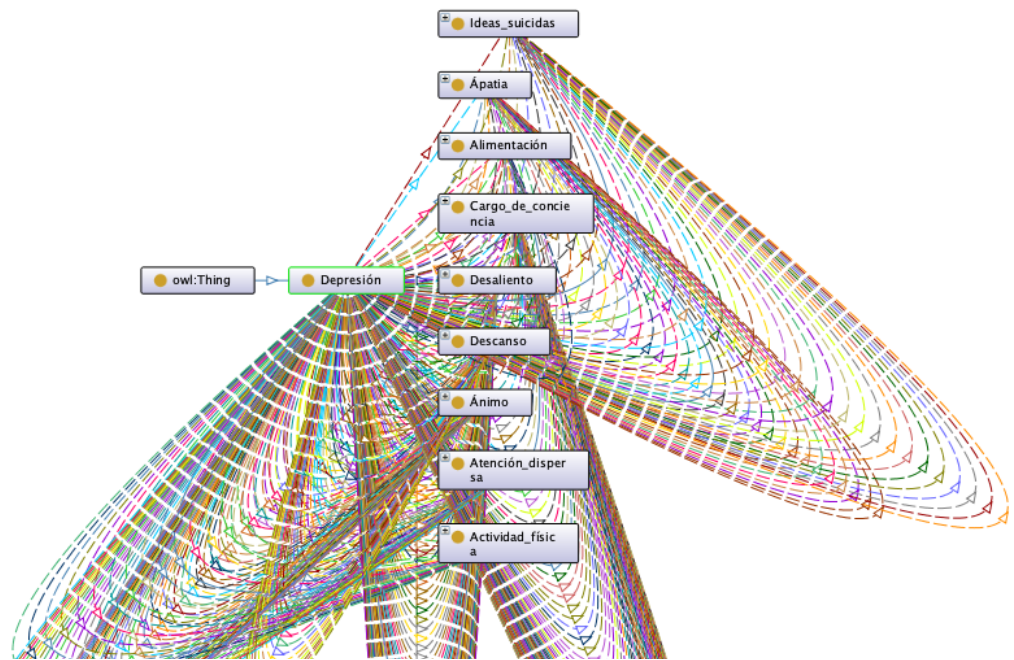


Figura 17. Gráfico de la ontología exportada de Protégé. Fuente: [El Autor]<sup>127</sup>

Nota: En este gráfico se aprecia la ontología que fue elaborada en protégé.

La *Figura 17. Gráfico de la ontología exportada de Protégé*, es similar a una telaraña dado que los conceptos tienen relaciones entre ellos mismos lo que dificulta su entendimiento gráfico, es por esto que se implementan el uso de herramientas específicas para graficar ontologías, obteniendo así la *Figura 18. Gráfico de la ontología con herramienta de visualización especializada*, para acceder a estas figuras con más detalles de visualización puede dirigirse a el anexo 6.

## Figura 18

### Gráfico De La Ontología Con Herramientas Especializadas

<sup>127</sup> Se hace esta fracción de la representación del grafo completo debido a su magnitud y extensión. El grafo completo es añadido a este trabajo como el anexo número 6.

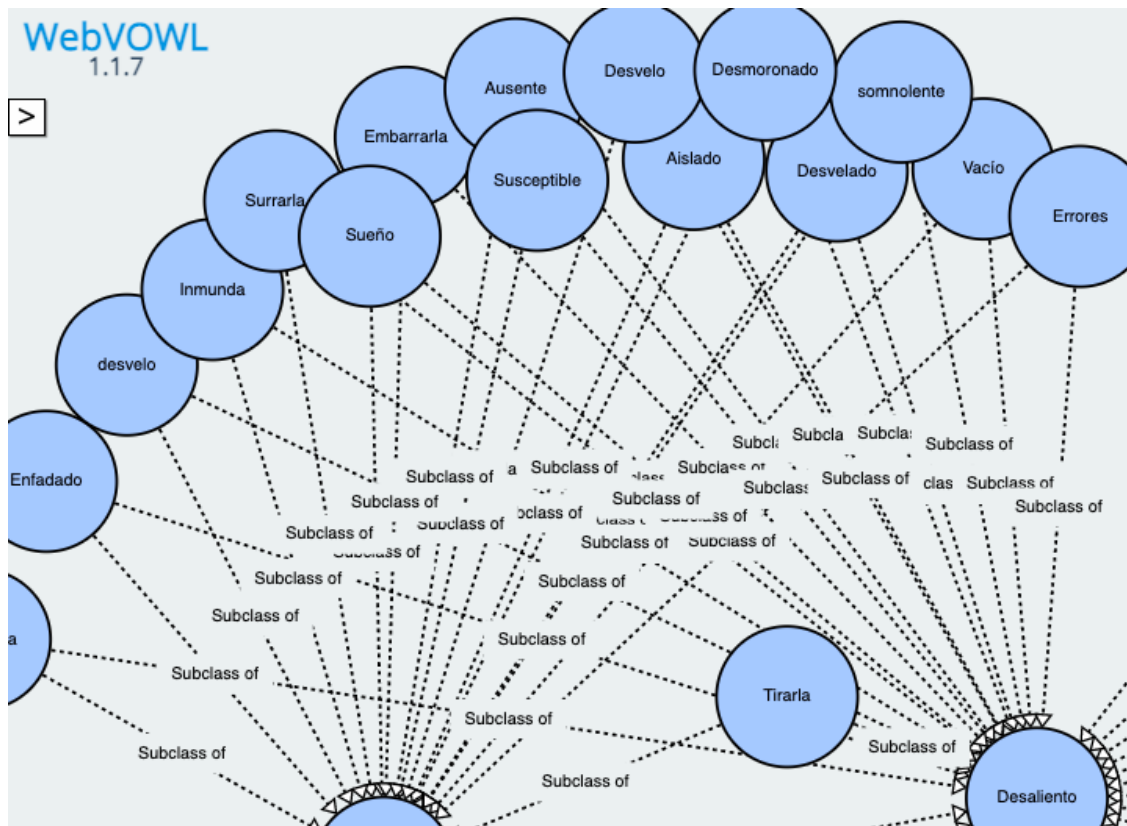


Figura 18. Gráfico de la ontología con herramienta de visualización especializada. Fuente: [El Autor]<sup>128</sup>

*Nota:* En el gráfico anterior se aprecia el gráfico de la ontología utilizando una herramienta que permite la visualización de esta de mejor manera.

Para poder tratar los conceptos y sus diferentes instancias en el entorno de desarrollo del método automático se realizan las actividades de validación, que en el ítem 2.7 son explicadas.

## 2.7 VALIDAR CON JUICIO DE EXPERTOS EN LAS ÁREAS DE PSICOLOGÍA E INGENIERÍA

### 2.7.1 VALIDACIÓN AUTÓNOMAS

Para llevar a cabo la validación de la ontología realizada en este objetivo se llevaron a cabo diferentes reuniones con personas del área psicología y de ingeniería que tendría una relación con la aplicabilidad de este proyecto. Más sin

<sup>128</sup> Se hace esta fracción de la representación del grafo completo debido a su magnitud y extensión. El grafo completo es añadido a este trabajo como el anexo número 6.

embargo el propio software de *Protégé* permite hacer una revisión interna de la ontología desarrollada en el mismo con base a las reglas internas del software y los resultados obtenidos de las validaciones automáticas.

- **Validaciones propias del software Protégé:**

Las validaciones propias del software *Protégé* son realizadas gracias a los *plugins*<sup>129</sup> que el software tiene integrado ya sean propios de él o añadidos por terceros. Estos son utilizados para que las ontologías sean verificadas bajo todas las reglas semánticas, sintácticas y de estructura propias de la documentación oficial “*ONTOLOGY DEVELOPMENT 101*”.

Cuando la validación interna es ejecutada se espera que se cumplan reglas propias para cada concepto y clase establecida, es decir que tengan un orden dentro del esquema y este responda a por lo menos una relación.

Se utilizó específicamente el *plugin Reasoner*<sup>130</sup>, el cual se encarga de la sincronización entre el modelo OWL<sup>131</sup> y el razonador interno de protección que supervisa la construcción del modelo OWL y sus cambios. Este *plugin* utiliza detectores de eventos del modelo OWL (por ejemplo, clase eliminada, creada, etc.) e informa al razonador que en la próxima solicitud es necesario llevar a cabo una sincronización completa o actualización parcial del desarrollo para que se valide su orden. Actualmente el *plugin Reasoner* es el principal validador que ofrece el *Software Protégé*, y su método de enseñar resultado es mediante la sección de *log*<sup>132</sup> del software.

- **Resultado obtenido de validación interna:**

Como se menciona en el ítem anterior la validación interna de la ontología desarrollada en *Protégé* se realiza con base en el trabajo del *plugin Reasoner*, el cual logra identificar si el modelo obtenido está reglamentado bajo los principios de

---

<sup>129</sup> “(En informática, un complemento o plug-in es una aplicación que se relaciona con otra para agregarle una función nueva y generalmente muy específica.)” Definición de plugin — Definición.de. (s. f.). Definición.de. {En línea}. {15 de Marzo de 2021} disponible en: <https://definicion.de/plugin/>

<sup>130</sup> (S/f). Recuperado el 20 de Marzo de 2021, de Stanford.edu website: {En línea}. {20 de Marzo de 2021} disponible en: <https://Protégéwiki.stanford.edu/wiki/ProtégéReasonerPlugin>

<sup>131</sup> Vista General del Lenguaje de Ontologías Web (OWL). (s. f.). w3.org {En línea}. {20 de Marzo de 2021} disponible en: <https://www.w3.org/2007/09/OWL-Overview-es.html>

<sup>132</sup> (En informática, se usa el término registro, log o historial de log para referirse a la grabación secuencial en un archivo o en una base de datos de todos los acontecimientos que afectan a un proceso particular. De esta forma constituye una evidencia del comportamiento del sistema.) Diaz, A. (2017, enero 19). ¿Qué son los Logs y por qué deben interesarte? Recuperado el 12 de abril de 2021, de Dbibyhavas.io website: {En línea}. {20 de Marzo de 2021} disponible en: <https://dbibyhavas.io/es/blog/que-son-los-logs/>

construcción de ontologías del mismo. La lista de resultados obtenidos son expuestos mediante un Log, este describe si es necesario hacer modificaciones a la estructura del modelo construido o si este es correcto para su implementación final.

En las Figuras 19. Captura del Log generado por Reasoner-Protégé y 20. Captura del mensaje de aprobación generado por OntoDebug se muestra el Log obtenido por la ontología desarrollada en este proyecto, este especifica que no hay necesidad de realizar modificaciones para la implementación de este modelo ontológico, en el campo de aplicación al que se desee destinar.

## Figura 19

### Captura Del Log Generado Por Reasoner-Protégé

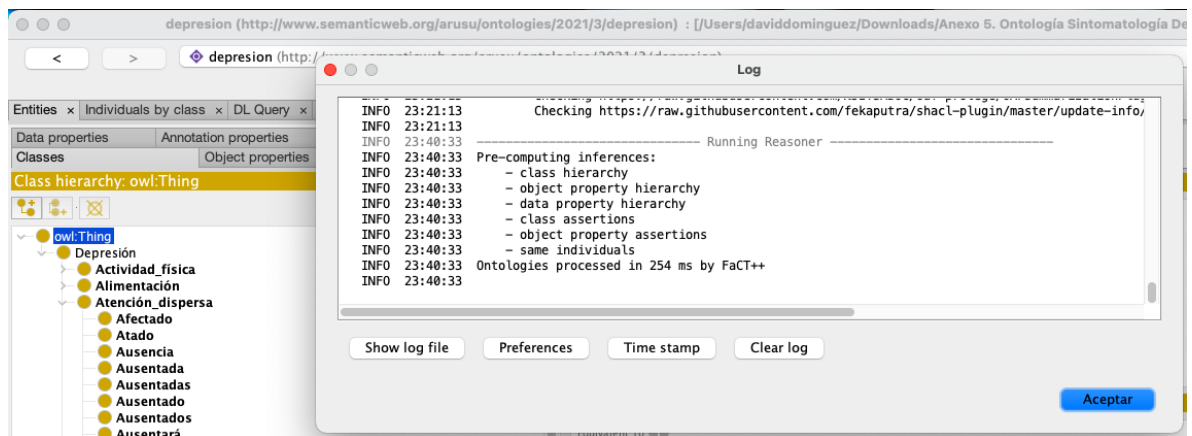


Figura 19. Captura del Log generado por Reasoner-Protégé<sup>133</sup>

Nota: En esta figura se muestra el log generado al construir la ontología.

## Figura 20

### Mensaje De Aprobación Generado por OntoDebug.

<sup>133</sup> El archivo “Log” completo generado por Protégé para la auto-validación de la ontología desarrollada, es añadido a este trabajo como el anexo número 7.



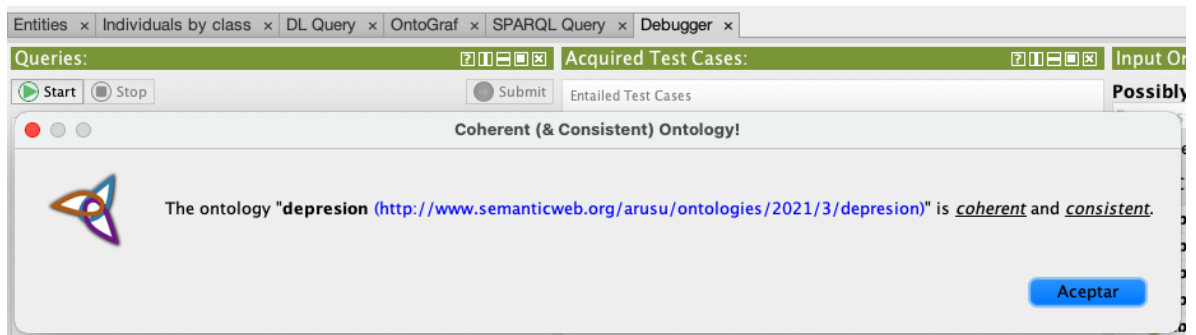


Figura 20. Captura del mensaje de aprobación generado por OntoDebug.<sup>134</sup>

*Nota:* En la anterior figura se puede apreciar el mensaje de aprobación el cual valida la coherencia y consistencia de la ontología creada.

## 2.7.2 VALIDACIÓN DE EXPERTOS

Para finalizar el desarrollo de este capítulo, se llevó a cabo una serie de reuniones, con las cuales se buscaría la validación con juicio de expertos tanto de áreas de salud mental de la facultad de psicología y de expertos de la facultad de ingeniería ambos de la Universidad Católica de Colombia, para que con estas validaciones y aprobaciones finalizara la construcción de la ontología y llevara cabo la exportación del contenido de esta al método automático, al igual que con esta ontología cumplir con el primero de los entregables de este trabajo de grado, su contenido tendría aplicabilidad en diferentes ontologías futuras y demás aporte a la biblioteca Universitaria.

Específicamente para este apartado se utiliza una métrica de validación para la ontología desarrollada, para así con esto construir un documento formal que valide y certifique la viable utilización de esta ontología tanto para el método automático desarrollado en este proyecto de grado, como para trabajos futuros y proyectos relacionados.

La métrica de validación para esta ontología es una adaptación a partir de la metodología encontrada del artículo científico "*Construcción de una ontología utilizando Protégé para la elicitación de requerimientos*"<sup>135</sup>

<sup>134</sup> OntoDebug - Protégé Wiki. (s. f.). Protégéwiki.Stanford.Edu. {En línea}. {20 de marzo de 2021} disponible en: <https://Protégéwiki.stanford.edu/wiki/OntoDebug>

<sup>135</sup> Marciszack, M. M. (s. f.). Construcción de una ontología para gramáticas formales y máquinas abstractas utilizando Protégé para la elicitación de requerimientos. sedici.unlp.edu.ar. {En línea}. {20 de marzo de 2021} disponible en: <http://sedici.unlp.edu.ar/handle/10915/19743>

**Documento:**

Marciszack, M., Pérez Cota, M., Antonelli, R., Cardenas, M. and Giandini, R., 2009. *Construcción De Una ontología Para Gramáticas Formales Y Máquinas Abstractas Utilizando Protégé Para La Elicitación De Requerimientos*. [ebook] Argentina: SEDICI, pp.296-300. Disponible en::

Esta métrica de validación para la ontología se encuentra en el apartado de resultados y puede ser consultada en la documentación del artículo. Con base a esto, se realizó una métrica de validación específica para este proyecto, agregando ítems de enfoque específicos tanto para los profesionales del área de la psicología y la salud mental, como para los profesionales de Ingeniería de Sistemas y computación, estos relacionados con su enfoque.

La métrica final utilizada para este proyecto lleva consigo parte de los campos y criterios únicos de la documentación “Construcción de una ontología para gramáticas formales y máquinas abstractas utilizando Protégé para la elicitación de requerimientos”, además de otras validaciones como lo son: preguntas, sugerencias, oportunidades, cambios, acuerdos, discusiones, además del apartado de las firmas, que corresponden a la validación del profesional en la presentación y seguimiento de la ontología desarrollada. La métrica de validación final es la siguiente, representada en la Tabla 7.

**Tabla 7**

*Métrica de validación ontología*

Métrica de evaluación para ontologías					
Sintoma	Categoría	Criterio evaluado	Especificación	Valoración	Observación
Generales	Suficiencia	Potencialidad para abstraer esencia del dominio	La información referida a las etapas de diseño e implementación de la ontología fue clara y suficiente.		
	Metodología	Replicación de metodología	La metodología utilizada para el modelado es fácil de replicar en otros proyectos.		
	Simplicidad	Entendimiento de la Ontología	La ontología de síntomas depresivos es entendible para el usuario		
Estado de ánimo deprimido	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras seleccionadas con las categorías sintomáticas correspondiente con el DSM IV.		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a los conceptos propuestos desde la psicología clínica		

<[https://www.profesores.frc.utn.edu.ar/sistemas/ssl/marciszack/Documentos/ontologías\\_CNIT2009\].pdf](https://www.profesores.frc.utn.edu.ar/sistemas/ssl/marciszack/Documentos/ontologías_CNIT2009].pdf)> [Consultado en 20 de marzo 2021].

	Coherencia	Número de conceptos	El número de palabras identificadas por cada categoría es suficiente para evaluar el dominio específico		
--	------------	---------------------	---	--	--

Tabla 7. Métrica de validación ontología<sup>136</sup>

*Nota:* En esta tabla se tienen las métricas que se utilizaron para la validación de la ontología.

Si la métrica respectiva es completa en su totalidad con un 100% de aprobación por el jurado, la tabla 8 presentada a continuación, la cual resumen todos los criterios de evaluación de ontología, se completa satisfactoriamente, dando así por terminado el capítulo 2, el cual correspondió a la construcción de la ontología que sea utilizada como insumo para el método automático desarrollado en este proyecto.

### Tabla 8

#### Métrica De Validación De Ontología Agrupada

Criterios de evaluación de ontologías						
Criterio > Autor	Suficiencia	Claridad	Coherencia	Relevancia	Metodología	Simplicidad
1						
2						
3						
4						
Autores:						

Tabla 8. Métrica de validación ontología agrupada<sup>137</sup>

*Nota:* En esta tabla se pueden observar los criterios de evaluación de las ontologías.

<sup>136</sup> La tabla 7 es una captura a la tabla original. Se hace captura de la métrica de validación debido a su extensión, la tabla original es añadida a este trabajo como el anexo número 8.

<sup>137</sup> Las correcciones y diferentes modificaciones obtenidas en la reunión y en las diferentes rúbricas fueron realizadas y se enviaron copias de las mismas a todos los jurados que participarán en esta validación de ontología para así dar finalidad al capítulo 2 de este proyecto.

**Fuente:** Ramos, E. (s. f.). Esquema para evaluar ontologías únicas para un dominio de conocimiento. scielo.org. {En línea}. {20 de marzo de 2021} disponible en: [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1690-75152009000100005](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1690-75152009000100005)

### **3. DISEÑO DEL MÉTODO AUTOMÁTICO**

Los diseños para el método automático son necesarios para obtener un buen desarrollo del mismo, permitiendo que este sea funcional y óptimo para los usuarios finales. En esta sección se mostrará el desarrollo que se llevó a cabo para obtener el diseño de este método automático.

Es necesario llevar a cabo una documentación específica referente a cada uno de los diferentes procedimientos con los cuales se trabajaron las fases respectivas: Análisis de información, estipulación de requerimientos y realización de diagramas que representan el diseño del método automático.

#### **3.1 ANÁLISIS DE INFORMACIÓN**

Esta sección reúne la investigación realizada para determinar cuáles serían las técnicas, tecnologías y librerías de desarrollo, con las cuales el método automático trabajara, para así lograr determinar las palabras que posiblemente señalan la presencia de sintomatología depresiva en las personas que responden a un formulario de preguntas el cual es destinado a señalar su estado de ánimo y mental.

Las diferentes técnicas, tecnologías y librerías que se investigan y clasifican en este apartado, son específicas del entorno de trabajo de Google Colab. Es importante resaltar que esta investigación y clasificación es distintiva del tercer capítulo de este proyecto de grado, esta investigación es importante en la fase de diseño, ya que permite dar a conocer cuál es la técnica o librería más ocionada y enfocada a lo esperado en el proyecto.

Para el presente proyecto se seleccionaron las herramientas presentadas en las tablas 9, 10, 11 a partir del criterio de viabilidad técnica que se describe en las mismas, todos los conceptos y términos acá descritos se encuentran sustentados en los marcos referenciales de este proyecto para más información y detalle.

Para el apartado de librería son señaladas las siguientes:

**Tabla 9**

*Selección De Técnicas y Tecnologías (Librerías)*

Selección de técnicas y tecnologías - Librerías	
Librería	Viabilidad
Spacy <sup>138</sup>	Óptima, centrado en el procesamiento de lenguaje natural
Requests <sup>139</sup>	Óptima, centrada en la extracción de información de páginas y documentación en línea.
BeautifulSoup <sup>140</sup>	Óptima, centrada en el buen uso y control de información de archivos XML, dado que este será el puente para conectar con el formulario completado por el usuario.
Deplacy <sup>141</sup>	Óptima, dado que se centra en la integración del usuario programado con el cuaderno de trabajo específico de Google Colab.
Es_core_news_sm <sup>142</sup>	Óptima, dado que centra todo el proceso de análisis de textos y del lenguaje, en la lengua española, y esta es la específica del país y la región en donde se realiza este proyecto.
pkg_resources <sup>143</sup>	Óptimo, dado que permite trabajar diferentes librerías y módulos en el mismo cuadernillo, ya que este método automático, pretende trabajar con diferentes librerías de Python dado que tiene uno o más procesos internos para la clasificación de palabras finales.

<sup>138</sup> spaCy · Industrial-strength Natural Language Processing in Python. (2021). spacy. [En línea] disponible en: {En línea}. {20 de marzo de 2021} disponible en: <https://spacy.io/>

<sup>139</sup> Requests. {En línea}. {25 de marzo de 2021} disponible en: disponible en: <https://pypi.org/project/requests/>

<sup>140</sup> Beautiful Soup Documentation — Beautiful Soup 4.9.0 documentation. (s. f.). Beautiful Soup Documentation. {En línea}. {25 de marzo de 2021} disponible en: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

<sup>141</sup> deplacy. PyPI. {En línea}. {25 de marzo de 2021} disponible en: <https://pypi.org/project/deplacy/>

<sup>142</sup> Spanish · spaCy Models Documentation. (s. f.). Spanish. {En línea}. {25 de marzo de 2021} disponible en: <https://spacy.io/models/es/>

<sup>143</sup> Package Discovery and Resource Access using pkg\_resources — setuptools 51.1.2 documentation. (s. f.). Package Discovery and Resource Access using pkg\_resources. {En línea}. {25 de marzo de 2021} disponible en: [https://setuptools.readthedocs.io/en/latest/pkg\\_resources.html](https://setuptools.readthedocs.io/en/latest/pkg_resources.html)

imp <sup>144</sup>	Óptimo, aunque no presenta una función en el entorno del análisis de lenguaje o de textos, el módulo IMP permitiría acceder a los archivos internos de las diferentes librerías importadas de una manera más específica.
--------------------	--

Tabla 9. Selección de técnicas y tecnologías - Librerías

Notas: En la tabla anterior se encuentran las librerías y tecnologías utilizadas.

Para el apartado de técnicas y tecnologías son señaladas las siguientes:

## Tabla 10

### Selección De Técnicas y Tecnologías (Funciones)

Selección de técnicas y tecnologías - Funciones	
Librería	Viabilidad
imp.reload <sup>145</sup>	Óptima, dado que este tipo de funciones permiten que el método automático genere uno o más análisis a diferentes textos en el mismo momento de ejecución y no sea necesario reiniciar todo el entorno donde se pueda perder información del archivo o tener que volver a cargarlo manualmente.
Counter <sup>146</sup>	Óptima, dado que nos permite hacer conteos básicos, para así saber el número de repeticiones de palabras específicas, esto se enfoca principalmente en análisis de textos relacionado con ontología de palabras.
import files <sup>147</sup>	Óptima, puesto que será viable al momento de necesitar importar o exportar archivos generados para el trabajo en el método automático, o los resultados del mismo.

<sup>144</sup> imp — Access the import internals — Python 3.9.1 documentation. (s. f.). imp. {En línea}. {25 de Marzo de 2021} disponible en: <https://docs.Python.org/3/library/imp.html>

<sup>145</sup> Python Examples of imp.reload. (s. f.). Python Examples of imp.reload. {En línea}. {25 de Marzo de 2021} disponible en: <https://www.programcreek.com/Python/example/6950/imp.reload>

<sup>146</sup> Hellmann, D. (s. f.). Counter - Python Module of the Week. Counter - Python Module of the Week. {En línea} {01 de Abril de 2021} disponible en: <https://pymotw.com/2/collections/counter.html>

<sup>147</sup> Masum, M. (2020, 27 octubre). How to Import and Export Datasets in Google Colab - Towards Data Science. Medium. {En línea} {01 de Abril de 2021} disponible en: <https://towardsdatascience.com/google-colab-import-and-export-datasets-eccf801e2971>

openpyxl <sup>148</sup>	Óptima, dado que será un aporte para el puente de conexión entre la información extraída de los formularios de Google con el entendimiento en el método automático de los mismos
read <sup>149</sup>	Óptimo, ya que esta función podrá ser utilizada para la administración de la información en los archivos generados o importados al método automático.
split <sup>150</sup>	Óptimo, dado que será implementado para la separación palabra a palabra en el en el método automático para poder así manejar las palabras del texto extraído por el paciente de una manera en índice de lista.
openpyxl.load_workbook <sup>151</sup>	Óptimo, permitiendo así la carga directa de hojas de cálculo, la cual es el método por el cual se extrae la información a las diferentes respuestas generadas por los formularios de Google.
iter_rows <sup>152</sup>	Óptimo, dado que funciona como complemento para los diferentes métodos, funciones y librerías que implementara este método automático controlando así todo el manejo de las hojas de cálculo de Excel en Python.
chunk.text <sup>153</sup>	Óptima, ya que con esta función podría ser implementada en el momento específico del análisis en el lenguaje natural de los diferentes textos.
most_common <sup>154</sup>	Óptima, ya que puede ser implementada para análisis finales, sobre el recuento de palabras más comunes repetidas y las más utilizadas que pueden significar o tener relación con la ontología de depresión.

<sup>148</sup> openpyxl - A Python library to read/write Excel 2010 xlsx/xlsm files — openpyxl 3.0.5 documentation. (s. f.). openpyxl A Python library to read/write Excel 2010 xlsx/xlsm files. {En línea} {01 de Abril de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/>

<sup>149</sup> GeeksforGeeks. (2019, 18 junio). Python | os.read() method. {En línea} {01 de Abril de 2021} disponible en: [https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read\(\)%20method%20in%20Python.bytes%20left%20to%20be%20read.](https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read()%20method%20in%20Python.bytes%20left%20to%20be%20read.)

<sup>150</sup> Python String split() Method. (s. f.). Python String split() Method. {En línea} {05 de Abril de 2021} disponible en: [https://www.w3schools.com/Python/ref\\_string\\_split.asp](https://www.w3schools.com/Python/ref_string_split.asp)

<sup>151</sup> Simple usage — openpyxl 3.0.5 documentation. (s. f.). Simple usage — openpyxl 3.0.5 documentation. {En línea} {05 de Abril de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/usage.html>

<sup>152</sup> openpyxl.worksheet.worksheet module — openpyxl 3.0.5 documentation. (s. f.). iter\_rows openpyxl.worksheet.worksheet module. {En línea} {05 de Abril de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/api/openpyxl.worksheet.worksheet.html>

<sup>153</sup> 7. Extracting Information from Text. (s. f.). Extracting Information from Text. {En línea} {05 de Abril de 2021} disponible en: <https://www.nltk.org/book/ch07.html>

<sup>154</sup> Code Faster with Line-of-Code Completions, Cloudless Processing. (s. f.). Code Faster with Line-of-Code Completions. {En línea} {15 de Abril de 2021} disponible en: [https://www.kite.com/Python/docs/collections.Counter.most\\_common](https://www.kite.com/Python/docs/collections.Counter.most_common)

re.finditer <sup>155</sup>	Óptima, ya que puede ser utilizada para la clasificación y ubicación de las palabras o conceptos más relevantes en las diferentes respuestas de las personas analizadas.
----------------------------	--

Tabla 10. Selección de técnicas y tecnologías - Funciones

Nota: En la tabla anterior se muestran las funciones que fueron utilizadas para la elaboración del proyecto.

Para el apartado de algoritmos y transformaciones de este proyecto seleccionamos:

**Tabla 11**

*Selección de Técnicas y Tecnologías (Algoritmos y transformaciones)*

Selección de técnicas y tecnologías - Algoritmos y transformaciones	
Técnica	Viabilidad
Procesamiento de lenguaje natural <sup>156</sup>	Óptimo, dado que como se menciona en la descripción anterior, entre los campos de aplicación del procesamiento lenguaje natural, se encuentra el análisis de sentimientos y este es el objetivo principal de este proyecto investigación.
Limpieza de textos <sup>157</sup>	Óptima, ya que la técnica de limpieza de textos, es primordial para cualquier proyecto en donde se utilice y se espere un manejo de datos, para que así de los mismos datos extraer información que se encuentre implícita en el mismo.

<sup>155</sup> Python Language - Iterando sobre los partidos usando `re.finditer` | Python Tutorial. (s. f.). Python Language - Iterando sobre los partidos usando `re.finditer`. {En línea} {15 de Abril de 2021} disponible en: <https://riptutorial.com/es/Python/example/12097/iterando-sobre-los-partidos-usando--re-finditer->

<sup>156</sup> N. (2019, 6 abril). Procesamiento del Lenguaje Natural (NLP). Aprende Machine Learning. {En línea} {15 de Abril de 2021} disponible en: <https://www.aprendemachinelarning.com/procesamiento-del-lenguaje-natural-nlp/>

<sup>157</sup> J., JayJay, V. A. P. B. (2018, 16 abril). Limpieza de texto utilizando Python. Data Science. [En línea] disponible en: {En línea} {21 de Abril de 2021} disponible en: <http://datascience.esy.es/limpieza-de-texto-utilizando-Python/>



Captura de información 158	Óptima, ya que este sería el proceso fundamental e inicial para método automático permitiendo así su usabilidad y aplicabilidad en el campo científico - mental.
Desarrollo de ontología 159	Óptimo, dado que el desarrollo una ontología genera el referente que permite entender las propiedades, restricciones, instancias y elementos identificables del concepto específico trabajado en el proyecto.
Métricas de desempeño 160	Óptimo, dado que este proyecto al manejar metodologías ágiles y un ítem específico para la refinación del proyecto, las métricas de desempeño son la base práctica y teórica que nos permite remodelar los objetivos del proyecto, generando así una mejora continua en los procesos de desarrollo.

Tabla 11. Selección de técnicas y tecnologías - Algoritmos y transformaciones

Nota: En la tabla anterior se observan los algoritmos utilizados para la elaboración del Método Automático.

<sup>158</sup> Captura de Información – Tandem. (2018, 10 julio). tandem. {En línea} {21 de Abril de 2021} disponible en:

<https://www.tandemweb.com/tandem/tag/captura-de-informacion/#:%7E:text=Eficacia%20y%20eficiencia%20en%20la.datos%20y%20aplicaciones%20de%20terceros>.

<sup>159</sup> X.D. (2019 - 2020). Desarrollo de ontologías. CS-FIB-UPC cbea. {En línea} {21 de Abril de 2021} disponible en: <https://www.cs.upc.edu/~bejar/ecsd/Teoria/ECSDI05b-Desarrollo.pdf>

<sup>160</sup> O.M. (2020, 27 agosto). Paso a paso para definir métricas de desempeño - Obed Medina. Medium {En línea} {21 de Abril de 2021} disponible en:

[.https://obedm.medium.com/paso-a-paso-para-definir-m%C3%A9tricas-de-desempe%C3%B1o-88a4207b5042#:~:text=Tener%20objetivos%20te%20permite%20enfocar.en%20el%20desempe%C3%B1o%20del%20equipo](https://obedm.medium.com/paso-a-paso-para-definir-m%C3%A9tricas-de-desempe%C3%B1o-88a4207b5042#:~:text=Tener%20objetivos%20te%20permite%20enfocar.en%20el%20desempe%C3%B1o%20del%20equipo).

## **3.2 ESPECIFICACIONES PARA EL MÉTODO AUTOMÁTICO ESPERADO**

### **3.2.1 DESCRIPCIÓN GENERAL**

Se tiene estimado que el método automático permita la identificación de sintomatología depresiva de una persona por medio de su narrativa escrita, está extraída de las respuestas a un formulario de preguntas psicológicas específicas, este método automático es construido con la tecnología de vanguardia, además de presentar un enfoque en garantizar el rendimiento óptimo en los navegadores web, dado que la instalación directa del método automático en un computador específico, limita el acceso al mismo.

Este método automático debe responder de manera responsiva en diferentes dispositivos, navegadores y sistemas operativos, siendo importante también la implementación de una interfaz clara, amigable e intuitiva con el usuario final.

Relatando el proceso específico que debe cumplir el método automático, se planea que esté responda al análisis de información extraída de las respuestas generadas por el paciente en su consulta por medio de un formulario de Google, está respuestas extraídas a una hoja de cálculo serán analizadas por el método automático situado en el entorno de Google Colab.

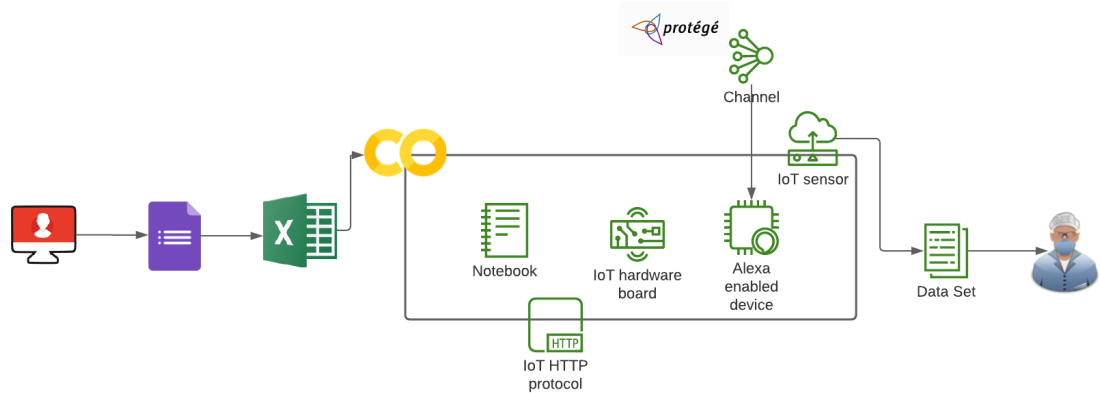
Una vez se extraigan estas respuestas a manera de listas y arreglos en el notebook de Google Colab, el método automático continuará con una fase de distintos procesamientos, como lo son limpieza del texto, análisis con librería de lenguaje natural y estadísticas de respuestas. Luego de esto con un proceso de identificación de palabras realizará una comparativa con la ontología desarrollada, que cumple con la clasificación de posibles términos y conceptos que identifiquen la sintomatología depresiva.

En el proceso siguiente, la correlación entre la ontología y las palabras más repetitivas de las respuestas del formulario, proceden a una clasificación de resultados y la generación de un informe final, que será entregado a expertos de salud mental y psicólogos, para ser analizado y utilizado como un insumo adicional en el momento de determinar si es posible que la persona que respondió al formulario en Google presenté alguno de los síntomas que identifican la presencia de depresión.

En la *Figura 21. Diagrama general de método automático esperado*, se grafica todo el proceso y la descripción general del mismo que se mencionó anteriormente.

**Figura 21**

*Diagrama General del Método Automático Esperado.*



*Figura 21. Diagrama general de método automático esperado. Fuente: [El Autor]*

*Nota: En la figura anterior se observa el diagrama general elaborado del método automático.*

### **3.2.2 REQUERIMIENTOS Y NECESIDADES DEL MÉTODO AUTOMÁTICO DADOS POR EL ÁREA DE SALUD MENTAL**

A continuación señalamos los respectivos requerimientos y necesidades con los que debe cumplir el método automático a desarrollar, estos requerimientos de uso fueron dados explícitamente por los profesionales del área de la salud mental, que serán aquellos que utilizarán este método automático en sus consultas. Estas estipulaciones son importantes para la fase de diseño del método, dado que con estos se describen las actividades que deben realizarse para que el comportamiento o función en particular del sistema sea el más óptimo posible.

Se espera señalar con los requerimientos las distintas condiciones que debe cumplir el método a desarrollar para que así presente un correcto desempeño y utilidad, tanto para profesionales del área de la salud mental, como para los desarrolladores y los usuarios finales.

Entre los requerimientos y necesidades descritas a continuación se expone una descripción detallada de la solicitud, al igual que una descripción del desarrollo previo y se complementa esta descripción con un flujo lógico del mismo.

Para la construcción de estos formatos de requerimientos se tomaron como referencia los modelos encontrados en dos *Formatos de Requerimientos de*

*Software*, el primero elaborado por Ofima s.a.s<sup>161</sup> y el segundo construido por el gobierno de Colombia, especialmente, para el *procedimiento: desarrollo de sistemas de información*<sup>162</sup> en la sección unidad para las víctimas.

La selección de estos anteriores modelos fue guiada por los parámetros para los requerimientos de software de la guía: *SWEBOK Guide to the Software Engineering Body of Knowledge v3.0*<sup>163</sup> (SWEBOK Guía del Cuerpo de conocimientos de ingeniería de software Versión 3.0), además de esto porque los documentos anteriormente señalados presentan un caso de éxito en sus aplicaciones.

Los requerimientos y necesidades descritas para el diseño y posterior construcción de este método automático con descritas en la *Tabla 12. Tabla de agrupamiento de requerimientos y necesidades generadas*.

**Tabla 12**

*Agrupamiento de Requerimientos y Necesidades Generadas.*

Requerimientos y necesidades especificadas	
Identificador	Título de la descripción
01	Ejecución práctica del método automático
02	Integrar librerías específicas de análisis de sentimientos
03	Identificar información del consultante mediante un código único (ID)
04	Identificar fuentes de clasificación para la sintomatología depresiva
05	Asegurar la privacidad de la información

*Tabla 12. Tabla de agrupamiento de requerimientos y necesidades generadas. Fuente:[El Autor]*

*Nota:* En la tabla anterior se tienen los requerimientos y las necesidades que se tuvieron en cuenta en la elaboración del proyecto.

<sup>161</sup> Ofima s.a.s, “Ejemplo\_Formato\_Requerimiento\_SoftwareVersion2.0.doc” {En línea} {21 de Abril de 2021} disponible en:

[https://www.ofima.com/wp-content/uploads/2018/05/Ejemplo\\_Formato\\_Requerimiento\\_SoftwareVersion2.0.doc](https://www.ofima.com/wp-content/uploads/2018/05/Ejemplo_Formato_Requerimiento_SoftwareVersion2.0.doc)

<sup>162</sup>(S/f-b). Recuperado el 12 de abril de 2021, de Gov.co website: {En línea} {25 de Abril de 2021} disponible en:

<https://www.unidadvictimas.gov.co/sites/default/files/documentosbiblioteca/formatodeespecificacionderequerimientosdesoftwarev1.doc>

<sup>163</sup> Pierre Bourque, Richard E. (Dick) Fairley, “SWEBOK Guide to the Software Engineering Body of Knowledge v3.0” {En línea} {25 de Abril de 2021} disponible en:

<https://ieeecs-media.computer.org/media/education/swebok/swebok-v3.pdf>

En la *Tabla 13. Tabla de requerimiento, Identificar información del consultante mediante un código único (ID)* se expone uno de los requerimientos y necesidades que fueron estipulados dentro del diseño de este método automático. Se pueden consultar los demás requerimientos y necesidades en el anexo 9.

**Tabla 13**

*Identificar información del consultante mediante un código único (ID)*

<b>PROYECTO</b>	Método automático para la identificación de sintomatología depresiva por medio de la narrativa escrita	
<b>Código:</b>	03	
Nombre Requerimiento:	Identificar información del consultante mediante un código único (ID)	
Responsable(s) Solicitud:	Profesionales del área de salud mental	
Responsable Funcional designado por el equipo de desarrollo:	Andrés David Domínguez Rozo Cristian Andrés Robayo Merchan	
<b>Descripción de la solicitud</b>		
<b>Usuario</b>		
Dado que este formulario de preguntas va a ser entregado y resuelto por más de un paciente, es necesario que esta información cuente con un código único (ID), que permita identificar tanto la información original del consultante como los resultados del método automático.		
<b>Descripción del desarrollo estipulado</b>		
<b>Descripción:</b> Generar un código único (ID) que permita la identificación de todas las respuestas dadas por un consultante específico, para que el experto de salud mental pueda individualizar el análisis generado por el método automático.		
<b>Flujo para el desarrollo estipulado</b>		
<b>Flujo lógico</b>	<b>Paso 1:</b>	El usuario se encuentra en la cuenta de Google que tiene el método automático en los cuadernillos de Google Colab almacenado.
	<b>Paso 2:</b>	El usuario abre el entorno de Google Colab, seguido del <i>notebook</i> de desarrollo del método automático.
	<b>Paso 3:</b>	El usuario ejecuta el botón correspondiente a la sección de carga del archivo de respuestas dadas por el consultante a las preguntas.
	<b>Paso 4:</b>	Internamente el archivo que almacena todas las respuestas dadas por los consultantes, genera

		automáticamente un código único (ID) específico para cada paciente y sus respuestas
	<b>Paso 5:</b>	Internamente el <i>notebook</i> de desarrollo correlaciona la hoja de cálculo con las respuestas al entorno de análisis del método automático, señalando como “llave” la columna de los códigos únicos.
	<b>Paso 6:</b>	El usuario digita el código único (ID) del consultante que desee tratar, este lo digita en el recuadro específico para esta información.
	<b>Paso 7:</b>	En adelante el <i>notebook</i> de desarrollo realiza toda la serie de análisis y demás clasificaciones de las respuestas generadas por el consultante pertenecientes específicamente a el ID señalado.
<b>Flujo alternativo para el desarrollo estipulado</b>		
	<b>Paso 1:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar los videos guías de ejecución.
	<b>Paso 2:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar la documentación “Manual de uso de Método automático para la identificación de síntomas depresivos”.

Tabla 13. Tabla de requerimiento, Identificar información del consultante mediante un código único (ID).  
Fuente:[El Autor]

*Nota:* En la tabla anterior se aprecia el requerimiento generado para la identificación de aquella información fundamental del consultante.

### 3.2.3 PERSPECTIVA DEL PRODUCTO

El método automático que permite la identificación de sintomatología depresiva de una persona por medio de su narrativa escrita, extraída de las respuestas a un formulario de preguntas psicológicas específico, está planteado como un método automático independiente, cuyo funcionamiento no está condicionado por ningún otro aplicativo. La interacción del mismo está directamente relacionado con el manejo de información a un análisis psicológico utilizando diferente implementación de algoritmos.

La arquitectura y las subsecciones de este método automático, van encaminados con los procesos fundamentales del mismo que son académicos y de libre acceso,

estos procesos son descritos en el apartado 3.3 de este capítulo, al igual que el diseño descriptivo (pseudocódigo) del método automático a realizar en este proyecto de grado el cual se señala en el apartado 3.4.1 del mismo.

### 3.3 PROCESOS FUNDAMENTALES

Los diseños descritos a continuación son trabajados bajo la metodología de lenguaje unificado de modelo, dado que se espera que sean entendidos para su implementación por cualquier persona relacionada con el área de software y que siguen un lineamiento común del mismo. En la actualidad este modelo de diseño está respaldado por el Object Management Group.<sup>164</sup>

#### 3.3.1 EXTRACCIÓN DE LA INFORMACIÓN (ETL)

Uno de los principales procesos fundamentales para este proyecto, es la extracción de información (ETL), este empieza cuando el usuario o paciente específico acepta llevar a cabo la consulta médica para la identificación a posibles síntomas de depresión que presente. El paciente realiza un formulario de Google de manera virtual, en donde se puede tomar la opción de aceptar o no el manejo de sus datos con la ley de habeas data, si este acepta logra acceder a el proceso de desarrollar el formulario construido con preguntas que esperan determinar si la persona puede o no experimentar sintomatología depresiva, si no, la consulta es finalizada.

Los formularios de Google como las hojas de cálculo son herramientas de trabajo para Google Drive, tanto el formulario como las diferentes respuestas están vinculadas con una hoja de cálculo respectiva, y al ser estas herramientas de Google Drive la transparencia y manera de trasladar la información es práctica y segura. Una vez los formularios son desarrollados las respuestas pasan a una hoja de cálculo, en donde también podrán ser vinculadas con el entorno de Google Colab, además de esto en el diseño se destaca que toda la información será mantenida y procesada, bajo el entorno de Google Drive.

La *Figura 22. Diagrama extracción de la información (ETL)* presenta el diseño estipulado para cumplir con el proceso de extracción de la información necesaria para la operación del método automático en las consultas.

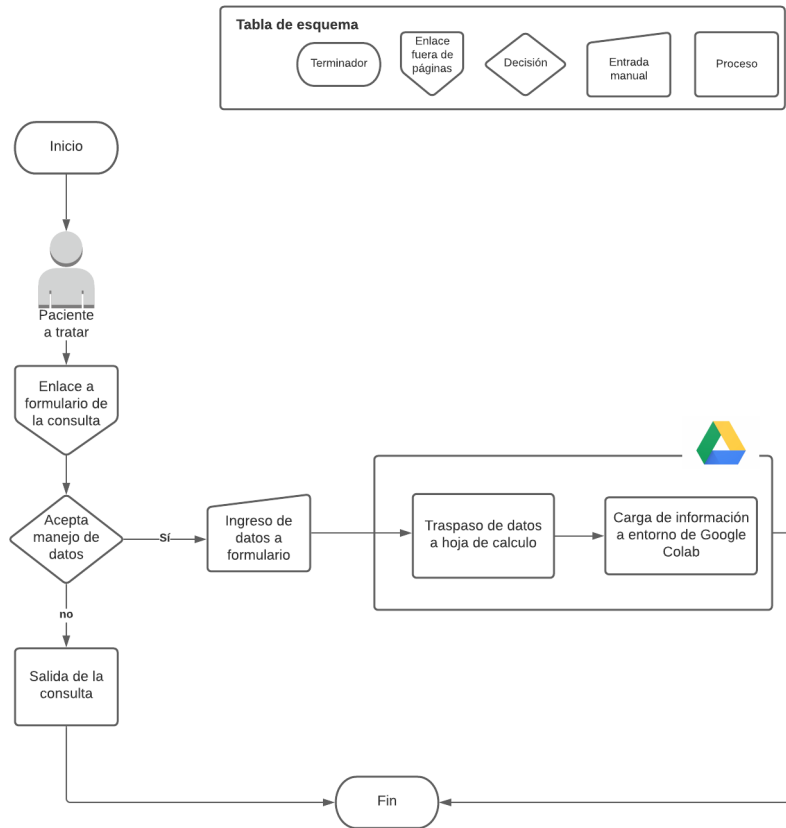
---

<sup>164</sup> Qué es el lenguaje unificado de modelado (UML). (s. f.). Lucidchart. {En línea} {25 de Abril de 2021} disponible en:

<https://www.lucidchart.com/pages/es/que-es-el-lenguaje-unificado-de-modelado-uml>

**Figura 22**

*Diagrama De Extracción De La Información*



*Figura 22. Diagrama extracción de la información (ETL). Fuente:[El Autor]*

*Nota:* En este diagrama se muestra el proceso de la extracción de la información.

### **3.3.2 LIMPIEZA DE INFORMACION**

Para el proceso de la limpieza de archivos es importante destacar que tanto el formulario de Google, como las hojas de cálculo y el entorno de Google Colab son herramientas de la nube Google Drive, es decir que la información trabajada en estas, puede ser transmitida de una a la otra, es por eso que el proceso de extracción de información fue realizado desde los formularios de Google, pasando por las hojas de cálculo y con estas generar un puente directo al entorno de Google Colab.

La limpieza archivos es fundamental para que el entendimiento computacional del texto sea el más óptimo, ya que se eliminan caracteres que no aportan a la comprensión artificial del texto, esto por parte de los diferentes algoritmos que



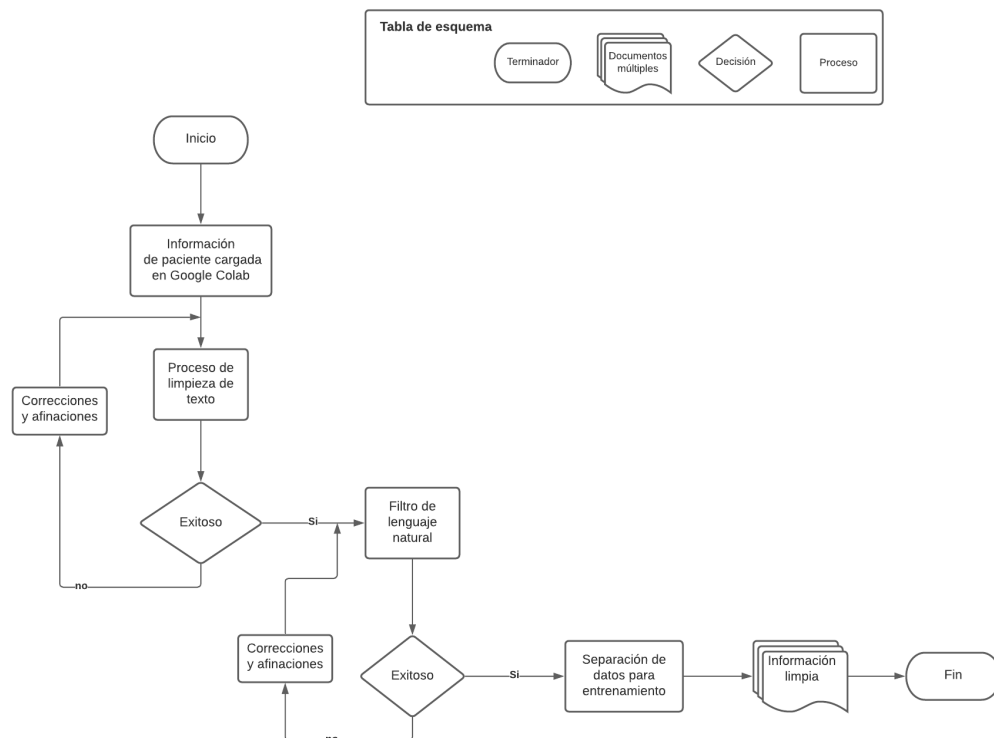
están presentes en el método automático como lo es el algoritmo de lenguaje natural.

Una vez las respuestas específicas por el paciente a tratar fueron importadas al entorno de Google Colab estas son transformadas en arreglos gracias a funciones de Python, estas respuestas ahora en arreglos son tratadas por diferentes procesamientos como lo son el de reemplazar caracteres incomprensidos, y otros diferentes de la rama de lenguaje natural. Llevar a cabo el proceso de limpiar el texto permite generar un nuevo arreglo con la información limpia y coherente para la , esta información está lista para el tratamiento por los diferentes algoritmos de clasificación y el proceso de entrenamiento y pruebas para este método automático.

La *Figura 23. Diagrama de limpieza de información*, presenta el diseño estipulado para cumplir con el proceso de limpieza de información necesaria para la operación del método automático en las consultas.

**Figura 23**

*Diagrama de Limpieza De Los Archivos*



*Figura 23. Diagrama de limpieza de los archivos. Fuente:[El Autor]*

*Notas:* El diagrama anterior muestra el proceso para la limpieza de archivos.

### 3.3.3 ANÁLISIS DEL TEXTO

La información original extraída de las respuestas directas del formulario, es procesada por diferentes funciones secundarias, generando con esto un nuevo arreglo de palabras “limpias”, con el cual se pueden realizar tratamientos en diferentes algoritmos como el de lenguaje natural y el de clasificación.

Diferentes algoritmos y procedimientos como lo son clasificación de sentimientos según un entrenamiento de máquina y correlación con la ontología de sintomatología depresiva son los que permiten determinar el sentimiento encontrado en las palabras de manera artificial con base a un entrenamiento y metodología propia en cada uno.

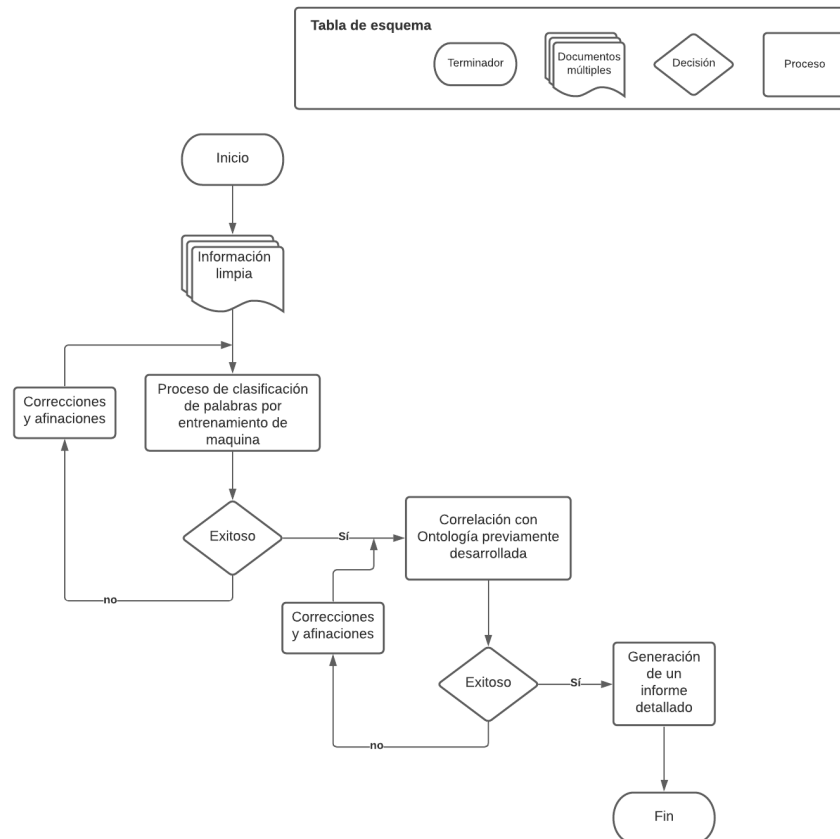
Trabajando con el arreglo de información “limpia”, se genera el informe de clasificación y correlación dictaminado por la relación que se genera entre las palabras de la ontología y las palabras descritas por el paciente al llenar el formulario.

Es importante destacar que la ontología que se utiliza es el referente teórico fundamental de este trabajo, esta fue desarrollada bajo la documentación de “Ontology development 101”, y también fue validada por profesionales tanto de la Salud Mental como de Ingenieros de Sistemas de la Universidad Católica de Colombia. (véase capítulo 2)

La Figura 24. Diagrama de Análisis del texto, presenta el diseño estipulado para cumplir con el proceso de análisis de información y texto, necesaria para la operación del método automático en las consultas.

**Figura 24**

*Diagrama de Análisis De Texto*



*Figura 24. Diagrama de Análisis del texto. Fuente: [El Autor]*

*Nota:* En el diagrama anterior se observa el proceso del análisis del texto.

### **3.3.4 MÉTRICAS DE DESEMPEÑO**

Las métricas de desempeño permiten llevar a cabo el proceso de calificación en los diferentes resultados obtenidos de este proyecto, para así poder validar el éxito del método automático. Estas métricas de desempeño hacen parte de la metodología estipulada, dado que con los resultados obtenidos de estas se permite clasificar si el desempeño técnico de los desarrollos fue el suficiente para cumplir con los objetivos propuestos.

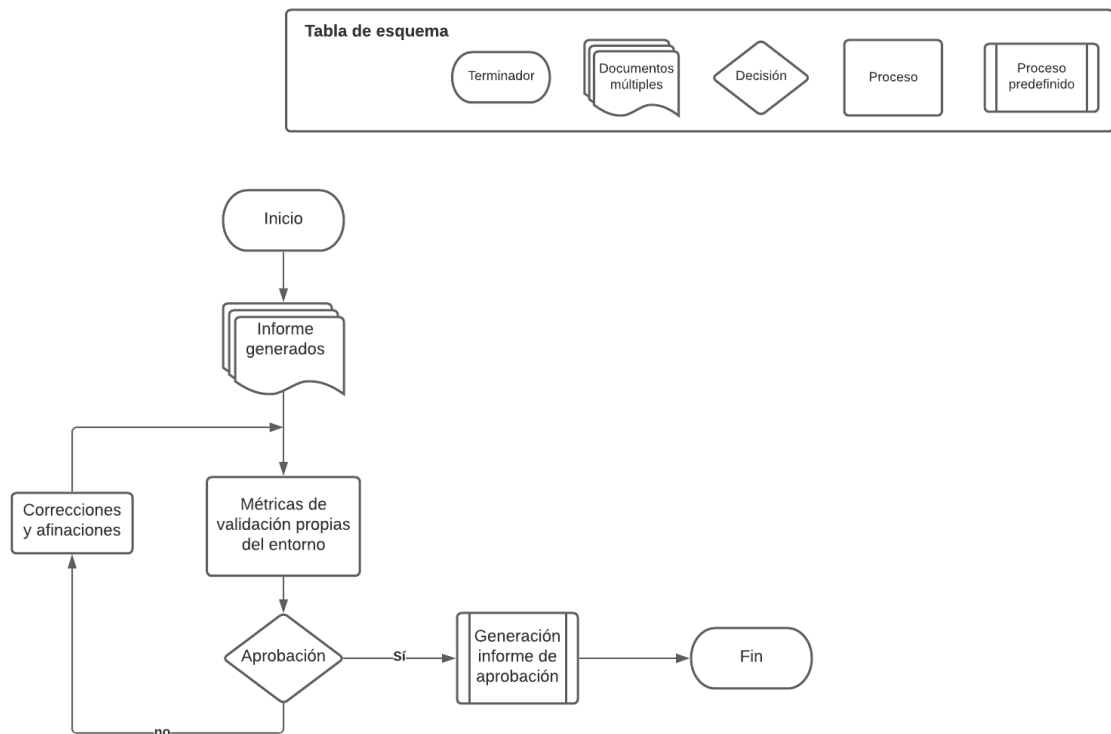
Este diseño muestra como diferentes métricas de desempeño de software cuantificable fueron evaluadas, tanto en el método automático desarrollado, como en la ontología, las métricas de desempeño generan así un informe tanto con características destacables, como con características por mejorar, estos procesos

son circulares dado que nos permiten realimentar el proyecto y afinar en cada momento este método automático, el resultado final de estas métricas de desempeño son la certificación y validación por diferentes profesionales en las áreas de interés del mismo.

La Figura 25. Diagrama Métricas de desempeño, presenta el diseño estipulado para cumplir con la implementación de las métricas de desempeño, necesarias para la validación del método automático estipulado.

**Figura 25**

*Diagrama Métricas De Desempeño*



*Figura 25. Diagrama Métricas de desempeño Fuente: [El Autor]*

*Nota:* En el diagrama anterior se observa el procedimiento para desarrollar las métricas de desempeño.

### 3.4 PSEUDOCÓDIGOS DEL MÉTODO AUTOMÁTICO

Para el diseño teórico de cada fase procedimental de este método automático, se utilizan pseudocódigos dado que estos permiten describir a alto nivel y de manera

compacta e informal, el principio para llevar a cabo la programación y escritura de código formal.

En ciencias de la computación y mayormente utilizado por programadores, los pseudocódigos son implementados para omitir secciones del código complejas y poder permitir dar una explicación al paradigma que se tomó para poder realizar el algoritmo final implementado. Este además es una herramienta esencial del diseño de software, para generar un preliminar al momento de empezar a escribir el código que cumpla con la función requerida para cada apartado o proceso de este método automático.

Con base a lo anterior, se presentan los siguientes pseudocódigos más esenciales para este método automático y que generan importancia significativa en la fase de diseño, antes de empezar a codificar.<sup>165</sup>, los demás realizados para este método automático, se encuentran en el anexo 9.

### 3.4.1 PROCESO DE LIMPIEZA DE TEXTOS

1. PROCESO Limpieza\_informacion\_respuestas;
2. IMPORTAR (librerías de lectura de archivos Excel);
3. GENERAR (variables que señalan hoja de cálculo);
4. GENERAR (arreglos vacíos donde se almacena información extraída);
5. **INICIO CICLO REPETITIVO**
6. ESCRIBIR (ciclo repetitivo que recorra la hoja de cálculo);
7. ASIGNAR(variable que almacena la información en cada iteración con función append);
8. **SALIR CICLO REPETITIVO**
9. REEMPLAZAR (Caracteres \_\_\_\_\_ al igual que espacios entre líneas);
10. COMPLETAR (Proceso de completación de palabras automáticamente);
11. ALMACENAR (Información resultante en nueva variable);
12. FIN

### 3.4.2 PROCESO DE IMPLEMENTACIÓN DE ALGORITMO DE LENGUAJE NATURAL

1. PROCESO Implementacion\_lenguaje\_natural;
2. IMPORTAR (librería de lenguaje natural);
3. GENERAR (variables que llama funciones de la librería de lenguaje natural);

---

<sup>165</sup> (S/f-c). Recuperado el 25 de abril de 2021, de Edu.mx website: {En línea} {25 de Abril de 2021} disponible en:

[https://www.uaeh.edu.mx/docencia/P\\_Presentaciones/tizayuca/ing\\_computacion/Fundamentos%20de%20Programacion.pdf](https://www.uaeh.edu.mx/docencia/P_Presentaciones/tizayuca/ing_computacion/Fundamentos%20de%20Programacion.pdf)

4. OPERAR (funciones de lenguaje natural con los arreglos de información limpia);
5. ALMACENAR (listas generadas con palabras claves seleccionadas por el algoritmo de lenguaje natural);
6. FIN

### 3.4.3 CORRELACIÓN DE PALABRAS DESTACADAS CON ONTOLOGÍA DESARROLLADA

1. PROCESO Correlacion\_Palabras\_ontología
2. IMPORTAR (librería de lectura de ontologías);
3. CARGAR (documento de ontología);
4. GENERAR (variable que almacena información de la ontología)
5. IMPORTACIÓN (Del documento de ontología a la variable que almacena esta información, en tipo de arreglo);
6. **INICIO CICLO REPETITIVO**
7. DECLARACIÓN (Recorrer por medio de una variable arreglos con información de ontología y de las respuestas)
8. **INICIO CONDICIONAL**
9. VALIDACIÓN (COINCIDEN TÉRMINOS)
10. SI CUMPLE (Adicionar elemento coincidente en otro arreglo)
11. NO CUMPLE (Continuar con validación en las demás palabras)
12. **FIN CONDICIONAL**
13. **FIN CICLO REPETITIVO**
14. IMPRIMIR (Arreglo con palabras coincidentes entre respuestas y ontología)
15. FIN

El diagrama general del método automático es la agrupación de los diagramas específicos anteriormente presentados, por esto presenta dificultades para su lectura en este tipo de documentos debido a su tamaño. Es por esto que este diagrama general del método automático se encuentra en formato pdf, para así detallar sus particularidades y el texto en el anexo 10.

## 4. IMPLEMENTACIÓN Y DESARROLLO DEL MÉTODO AUTOMÁTICO

Este apartado responde al desarrollo e implementación del método automático, trabajando con base en el diseño previamente establecido en el capítulo anterior (*Capítulo 3. DISEÑO DEL MÉTODO AUTOMÁTICO*).

Para llevar a cabo el proceso de implementación y desarrollo del método automático, fueron llevadas a cabo las siguientes fases de desarrollo donde cada una de estas cumple con una orientación, una finalidad y una metodología de trabajo. Estas fases son descritas a continuación:

#### **4.1 FASE 1. DESARROLLO DEL FORMULARIO DE PREGUNTAS**

El desarrollo del formulario de preguntas que pretende señalar la presencia o no de sintomatología depresiva, es la primera fase para el completo desarrollo y aplicación del método automático.

En la *Figura 26. Vista del usuario formulario*, se muestra una representación del formulario generado para identificar los posibles síntomas de depresión en la persona que lo responda, este fue debidamente estructurado por los profesionales del área de la Salud Mental, en el entorno de trabajo *Google Forms* dado que este es sencillo de compartir y accesible para trabajar con las respuestas obtenidas.

#### **Figura 26**

*Vista Del Usuario (Formulario).*

The image shows a screenshot of a Google Forms questionnaire. At the top, there is a header image with abstract, colorful, and somewhat blurry shapes. Below the image, the title of the form is "Preguntas para identificar tu estado de ánimo" (Questions to identify your mood state). Underneath the title, there is a red asterisk followed by the word "Obligatorio" (Mandatory). Below that, there is a checkbox for "Acepta el tratamiento de sus datos bajo la ley de Habeas Data: \*" (Accept the treatment of your data under the Habeas Data law: \*). At the bottom of the form, the text "LEY HABEAS DATA" is visible. The entire form is displayed within a browser window, with the URL "docs.google.com/forms/d/e/1FAIpQLSc7ia-b9MxaN6DmTTIn3WHgP6nv\_nGvuadQZID3HP\_VCFWg/viewform" visible in the address bar.

*Figura 26. Vista del usuario formulario. Fuente: [El Autor]*

*Nota: En la anterior figura se muestra como es la vista del usuario en el formulario*

Para la implementación y estructura de estas preguntas en los procesos que se

llevarán a cabo en método automático, se realizaron una serie de reuniones entre profesionales del área de la salud mental e ingenieros de sistemas, también en estas reuniones se trabajaron los procesos de validar si las posibles respuestas a estas preguntas logran por sí mismas señalar si existe o no la posibilidad de que la persona presentase algún síntoma de depresión y que internamente el método automático pudiera clasificarlo.<sup>166</sup>

Este formulario de preguntas genera una hoja de cálculo en el entorno de *Google Sheets*, la cual permite almacenar las diferentes respuestas generadas. En esta hoja de cálculo se realizó una Macro interna, la cual genera una serie de códigos únicos (ID) específicos para cada respuesta de cada paciente a tratar que resolverá el cuestionario. La macro realizada y el script de la misma pueden detallarse en las figuras 27 y 28 respectivamente.

## Figura 27

### Vista Del Usuario y Macro Realizada

The screenshot displays the Google Sheets interface on the left and the Apps Script editor on the right. The spreadsheet contains a table with the following data:

ID	Marca temporal	Acepta el tratamiento d	¿Como has estado esti
1			
2	13/02/2021 20:39:51	Si	Bien
3	13/02/2021 20:42:26	Si	Bien
4	13/02/2021 20:44:26	Si	Mal
5	13/02/2021 20:45:34	Si	Bien
6	13/02/2021 20:50:24	Si	Bien
7	13/02/2021 20:51:31	Si	Bien
8	13/02/2021 20:54:42	Si	Bien
9	13/02/2021 20:56:24	Si	Mal
10	13/02/2021 20:56:36	Si	Mal
11	13/02/2021 20:57:51	Si	Bien
12	13/02/2021 21:16:58	Si	Bien
13	13/02/2021 21:18:28	Si	Bien
14	13/02/2021 21:19:27	Si	Bien
15	13/02/2021 21:24:57	Si	Bien

The Apps Script editor on the right shows the following JavaScript code:

```

function generarID() {
  //Hoja sobre la que trabaja
  var hoja = SpreadsheetApp.
  getActiveSheet().getSheetByName
  ('Respuestas de formulario 1');
  //Empieza desde el primer registro
  var ultimaFila = hoja.getLastRow();
  var idConsecutivo = ultimaFila - 1;
  //Colocar el ID
  hoja.getRange(ultimaFila, 1).setValue
  (idConsecutivo);
}
  
```

Figura 27. Vista del usuario y macro realizada. Fuente:[El Autor]

Nota: En la anterior figura se muestra la vista del usuario y la macro realizada.

## Figura 28

<sup>166</sup> Si desea profundizar en el diagrama procedimental para el desarrollo de este formulario de preguntas, que es uno de los trabajos consolidados gracias a la interdisciplinariedad de este proyecto de grado, diríjase a la al anexo 9, ítem 8.



## Script De Macro En Google Sheets

```
function generarID() {  
  //Hoja sobre la que trabaja  
  var hoja = SpreadsheetApp.getActiveSpreadsheet().getSheetByName('Respuestas de  
formulario 1');  
  //Empieza desde el primer registro  
  var ultimaFila = hoja.getLastRow();  
  var idConsecutivo = ultimaFila - 1;  
  //Colocar el ID  
  hoja.getRange(ultimaFila, 1).setValue(idConsecutivo); }  
}
```

Figura 28. Script de macro en Google Sheets. Fuente: [El Autor]

*Nota:* En la figura anterior se observa el script de la macro en Google Sheets

### Descripción:

Este script muestra como es la operación entre variables permitiendo que el la hoja de cálculo asigne un Código único ID, a cada respuesta que se genere proveniente del formulario de preguntas. Está a su vez mostrándola en una columna independiente al costado derecho de toda la hoja de cálculo con las respuestas.

Así tanto de una manera visual como interna se le asignó a cada respuesta de los consultantes un código único ID que será consolidado en un formulario de preguntas cuyas respuestas serán analizadas por el método automático, para generar un posible informe que señala la presencia o no de sintomatología depresiva en la persona que lo responde.

Para consultar en detalle y de manera completa este formulario de preguntas y la macro realizada, diríjase al anexo 12. en la sección *Herramientas> Editor de secuencia de comandos*.

## 4.2 FASE 2. PROCESO DE ENTENDIMIENTO Y ANÁLISIS DE INFORMACIÓN

### 4.2.1 LIBRERÍAS

Para la fase del desarrollo que cubre el proceso de entendimiento y análisis de la información, se empieza con la selección y carga librerías, las librerías que fueron seleccionadas se tomaron bajo criterios técnicos y de usabilidad directamente relacionados con las necesidades que solicitaban los profesionales del área de la Salud Mental, estas solicitudes descritas más al detalle se encuentran en el tercer requerimiento capítulo tres. Para los procesos de carga librerías en Google colab se realizó bajo el siguiente Script.

La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 29 y 30 respectivamente.

## Figura 29

*Vista Del Usuario(Colab).*



Figura 29. Vista del usuario. Fuente:[El Autor]

*Nota:* En la figura 29 se muestra la vista del usuario en colab.

## Figura 30

*Script De La Función Activar Librería*

```
#@title <- Activa las librerías
import spacy
import requests
from bs4 import BeautifulSoup

!pip install deplacy
!Python -m spacy download es_core_news_sm
import pkg_resources,imp
imp.reload(pkg_resources)
nlp=spacy.load("es_core_news_sm")
```

```
from collections import Counter
```

Figura 30. Script de la función "Activar librerías". Fuente:[El Autor]

Nota: En la figura anterior se muestra el Script de la función de Activar Librerías.

### **Descripción:**

Se importan las librerías, módulos y herramientas que permiten el uso de herramientas e información tanto externas como internas, para las distintas operaciones que se llevarán a cabo en el método automático. La descripción de las librerías acá presentadas se encuentran detalladas en el (Capítulo 3. DISEÑO DEL MÉTODO AUTOMÁTICO ítem 3.1)

### **4.2.2 EXTRACCIÓN Y CARGA DEL ARCHIVO DE RESPUESTAS**

La segunda fase del desarrollo para los procesos de entendimiento y análisis de la información, se basan en realizar una conexión con el contenido de Google Drive en el cual se encuentra la hoja de cálculo que almacena todas las respuestas del formulario de preguntas para la identificación de síntomas de depresión.

### **EXTRACCIÓN DE LA HOJA DE CÁLCULO**

Este proceso ejecuta un Script el cual busca conectar el método automático al contenido de Google Drive, para así hacer descarga de archivos, específicamente el archivo de respuestas del formulario de preguntas. La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 31 y 32 respectivamente.

**Figura 31**

## Vista Del Usuario (Proceso Extracción De Datos).



Figura 31. Vista del usuario proceso extracción de datos. Fuente:[El Autor]

Nota: En la Figura Anterior se muestra la vista del usuario en el proceso de extracción de datos.

## Figura 32

### Script Proceso Extracción De Datos

```
#@title <- Conecta con Drive
from google.colab import drive
drive.mount('/content/drive')

#@title <- Descarga el archivo más reciente desde Google Drive
from google.colab import files
files.download('/content/drive/MyDrive/pruebas tesis/Preguntas para
identificar tu estado de ánimo (respuestas).xlsx')
```

Figura 32. Script proceso extracción de datos. Fuente:[El Autor]

Nota: En la anterior figura se muestra el script con el cual se extrajeron los datos.

### Descripción:

Estos Scripts importan una librería interna del entorno de Google Colab la cual le permite conectarse con el entorno en Google Drive y a la vez descargar archivos que estén almacenados en este. Para activar la librería de conexión con las diferentes carpetas de Google Drive es necesario autenticar la cuenta, es decir, la primera vez que se ejecuta se solicitará iniciar sesión de la cuenta respectiva para permitir la conexión de entornos.

## CARGA DEL ARCHIVO DE RESPUESTAS

Este proceso ejecuta un Script y abre una ventana para cargar el archivo que se descargó en el proceso anterior, esto se hace para que el método automático trabaje con un archivo xlsx , dado que con un archivo extraído directamente desde Google Sheets algunas funciones del lenguaje natural y de limpiezas de texto no serían posibles. Las vistas que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 33, 34 y 35 respectivamente.

### Figura 33

#### Vista Del Usuario (Carga De Archivo De Respuestas)



Figura 33. Vista del usuario carga de archivo de respuestas. Fuente:[El Autor]

Nota: En la figura anterior se muestra la vista del usuario de cargar el archivo de respuestas.

La celda de carga de archivos, se abre la ventana de navegador de archivos de Windows o iOS, esta le permite seleccionar al profesional del área Salud Mental, buscar el archivo respuestas que se descargó en el procedimiento anterior, seleccionar y darle clic en “abrir”.

### Figura 34

#### Vista Del Usuario (Validación De Archivo De Respuestas)

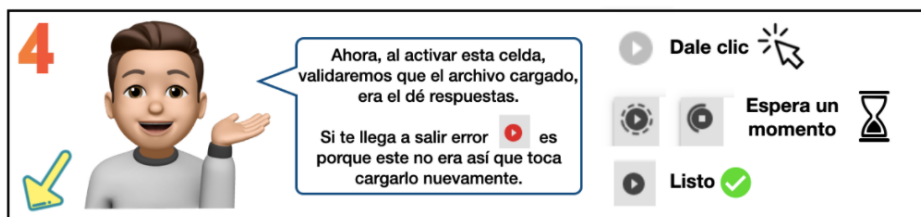


Figura 34. Vista del usuario validación carga de archivo de respuestas. Fuente:[El Autor]

Nota: En la figura se muestra la vista del usuario en la parte de la validación del archivo de respuestas.

La celda de validación nos permite verificar y enviar el archivo específico de

respuestas al proceso de análisis del método automático, es importante que esta validación sea correcta para proseguir. Internamente dentro de las secciones de carga y validación de archivo encontramos los siguientes *Scripts*:

### Figura 35

#### *Script Carga De Archivo De Respuestas*

```
#@title <- Valida que el archivo que cargaste se llama "Preguntas para identificar tu estado de ánimo (respuestas).xlsx"
import openpyxl
doc = openpyxl.load_workbook ("Preguntas para identificar tu estado de ánimo (respuestas).xlsx")
hoja = doc.get_sheet_by_name("Respuestas de formulario 1")
preguntas = hoja.delete_rows(1)
```

*Figura 35. Script carga de archivo de respuestas. Fuente:[El Autor]*

*Nota:* En esta figura se observa el script de la carga de los archivos de respuestas.

#### **Descripción:**

Se importa la librería especializada en el manejo de hojas de cálculo y archivos xlsx, también con la asignación una variable la cual señala el archivo de excel, se utilizará para las diferentes operaciones internas que el método automático realizará basándose en la información, que esta tenga almacenada.

### **4.2.3 ANÁLISIS DE INFORMACIÓN**

#### **SELECCIÓN DEL CÓDIGO ÚNICO (ID) DEL CONSULTANTE**

Esta celda le permite ingresar al profesional de la Salud Mental el código único (ID) correspondiente a la información del paciente y las respuestas generadas por el mismo, los ID 's se encuentran establecidos automáticamente en las hojas de cálculo que almacenan las respuestas de los consultantes. La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 36 y 37 respectivamente.

### Figura 36

## Vista Del Usuario (Selección De Código Único ID)

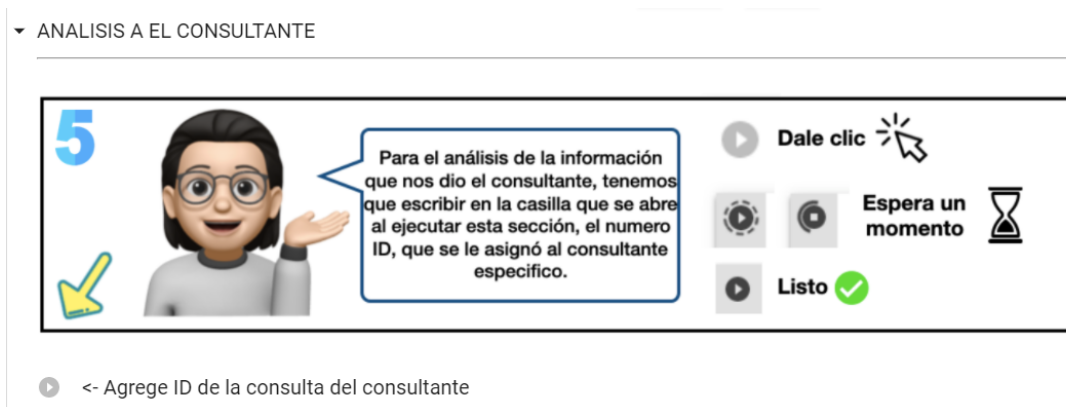


Figura 36. Vista del usuario selección de código único ID. Fuente:[El Autor]

Nota: Esta figura muestra la vista del usuario en la selección del código único ID

### Figura 37

#### Script Selección De Código Id

```
id = int(input("Ingrese el ID de la consulta del paciente = "))
```

Figura 37. Script selección de código único ID. Fuente:[El Autor]

Nota: La figura muestra el código donde se recibe el Id.

#### Descripción:

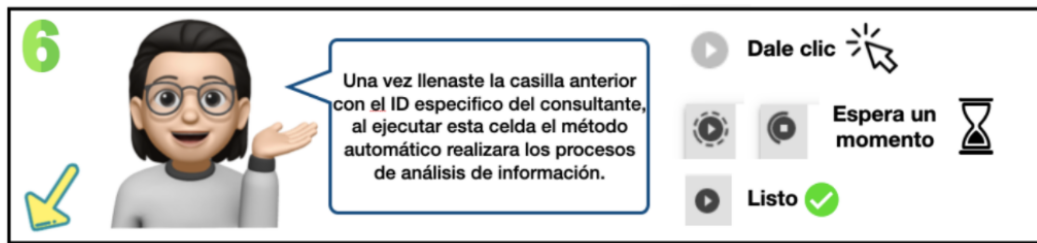
Con funciones internas del lenguaje Python se le solicita a el profesional de Salud Mental que ingrese el ID de las respuestas del consultante , que desea trabajar para el análisis de respuestas.

### PREPARACIÓN DE INFORMACIÓN PARA SU ANÁLISIS

Esta sección de código selecciona la información específica de las respuestas generadas por el consultante, y se realiza a esta una serie de procesos como lo son el de limpieza y conteo de palabras. Esto se da para generar informes preliminares y así abrir paso a el proceso de análisis por medio del lenguaje natural. La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 38 y 39 respectivamente.

### Figura 38

## Vista Del Usuario (Preparación De Información Para Su Análisis)



▶ <-Análisis de respuestas con lenguaje natural y repeticion de palabras

Figura 38. Vista del usuario preparación de información para su análisis. Fuente:[El Autor]

Nota: En la figura se muestra la vista del usuario, de la preparación de la información para el respectivo análisis

## Figura 39

### Script Preparación De Información Para Su Análisis

```
print("Análisis de respuestas, del paciente ID ",id,"\n")
##### Librerias especificas #####
import spacy
from spacy.lang.es.examples import sentences
import unicodedata
import re
nlp=spacy.load("es_core_news_sm")
##### Arreglos #####
respuestas=[]
palabras=[]
respuestasnone=[]
##### Información correspondiente al usuario desde hoja de calculo #####
celdainicial = "C{0}"
celdainicialF = celdainicial.format(id)
celdafinal = "AC{0}"
celdafinalF = celdafinal.format(id)
multiple_cells = hoja[celdainicialF:celdafinalF]
for row in multiple_cells:
    for cell in row:
        respuestasnone.append(cell.value)
##### Limpieza de texto #####
respuestas = list(filter(None, respuestasnone))
Respuestascompletas = " ".join(respuestas)
output_clean = Respuestascompletas.replace('\n', ' ')
output_clean2 = output_clean.replace('.', ' ')
```



```

output_clean3 = output_clean2.replace(',',' ')
trans_tab = dict.fromkeys(map(ord, u'\u0301\u0308'), None)#Tildes
output_clean4 = unicodedata.normalize('NFKC', unicodedata.normalize('NFKD',
output_clean3).translate(trans_tab))#Formato de texto
output_clean5 = output_clean4.title()
doc = nlp(output_clean5)
palabras = output_clean5.split()
##### Conteo de palabras #####
contador = Counter(palabras)
buscar = 'depression'
regex="("+".*?".join(buscar) + ")"
lista1Cadena=" ".join(palabras)
encontrados = re.finditer(regex, lista1Cadena, re.MULTILINE)
cantidad=len([matchNum for matchNum, match in enumerate(encontrados, start=1)])
print("_____")
print("_____YA SE PUEDEN EJECUTAR_____")
print("_____LOS SIGUIENTES ANÁLISIS_____")
print("_____")

```

Figura 39. Script preparación de información para su análisis. Fuente:[El Autor]

*Nota:* En la figura anterior se muestra el script de preparación de información para realizar el respectivo análisis.

### Descripción: (De arriba hacia abajo)

- Se le enseña al usuario cuál será el ID del consultante al cual se le realiza el análisis de respuestas.
- Se importan librerías específicas para los procesos de análisis y limpieza que serán llevados a cabo.
- Se generan los arreglos que almacenarán la información extraída y analizada de las respuestas.
- Se carga la información correspondiente al usuario desde la hoja de cálculo, sesgando únicamente la fila a la cual hace correspondencia el ID de respuestas.
- Se genera una limpieza de texto, eliminando vacíos, puntos, comas saldos de línea y demás caracteres que entorpecen a los modelos de lenguaje natural. También se eliminan las tildes de las palabras, esto para generar un arreglo de palabras finales que no integre caracteres especiales, ya que no son sencillos de entender para el lenguaje de máquina.
- Se modifica el formato de texto generando únicamente la primera letra en mayúscula para todas las palabras, esto para que cuando se relacionen con la Ontología las relaciones sean directas, ya que la ontología tiene almacenados los conceptos de esta manera.
- Se realiza un conteo de palabras en busca del concepto específico de

depresión, esto ya que es el concepto central de estos análisis y la investigación.

- Se presenta un mensaje el cual notifica al profesional del área la Salud Mental que ya finalizó el análisis y puede empezar a ejecutar las celdas que muestran gráficamente los análisis generados en este Script.

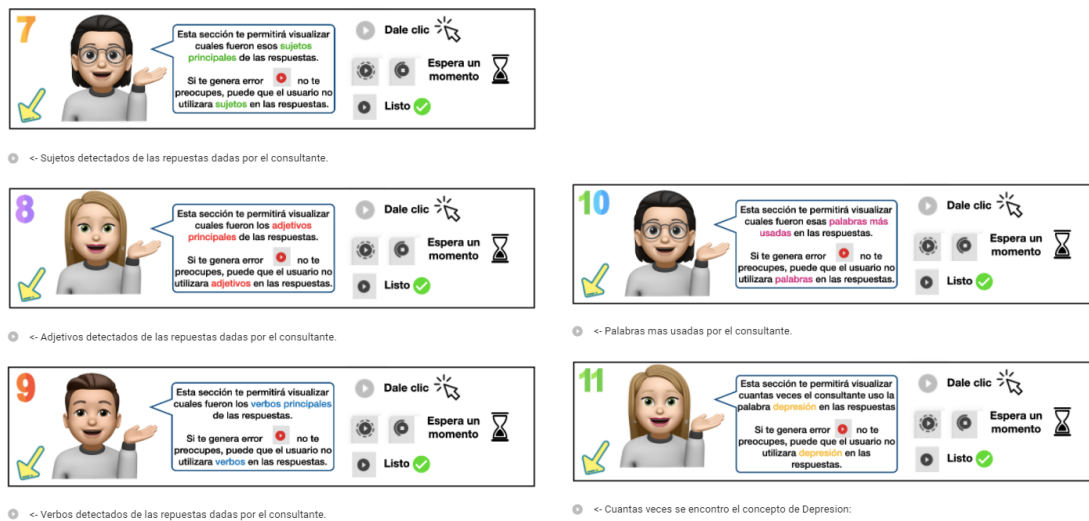
## ANÁLISIS GENERADOS

Los análisis que se presentan a continuación, son análisis específicos generados con los conceptos fundamentales de las respuestas dadas por el consultante, esos conceptos fueron extraídos por medio de modelos de lenguaje natural.

Los análisis aquí presentados muestran sus resultados de manera gráfica y todos llevan el mismo esquema de desarrollo, es por eso que en este apartado solo se describe el desarrollo de uno, al igual el script que se presenta es un Script genérico para estos análisis. Las vistas que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 40 y 41 respectivamente.

**Figura 40**

*Vista Del Usuario (Análisis Generados)*



*Figura 40. Vista del usuario análisis generados. Fuente:[El Autor]*

*Nota:* En la figura anterior se muestra la vista del usuario con los análisis generados.

**Figura 41**

```

##### Identificación categoría de análisis de lenguaje
POS_F = [] ## Etiquetas POS universales
for token in doc:
    if token.pos_ == "POS": ## Etiquetas POS universales
        POS_F.append(token.text) ## Etiquetas POS universales
contador = Counter(POS_F)
labels, values = zip(*contador.most_common(20))
##### Generación gráfica de presentación de resultados
indexes = np.arange(len(labels))
width = 1
fig = plt.figure(figsize=(10,5))
ax = fig.gca()
ax.bar(indexes, values, 0.5)
ax.set_xticks(indexes)
ax.set_xticklabels(labels, rotation=90, fontsize=20)
ax.set_title("Etiquetas POS universales detectados de las respuestas dadas por el
paciente en la consulta.", fontsize=20)
plt.show()

```

Figura 41. Script genérico análisis generados. Fuente:[El Autor]

*Nota:* En la figura se muestra el script realizado para los análisis que se realizaron.

### Descripción:

- La primera sección de este Script genera un arreglo vacío el cual va a ser llenado por un ciclo, el cual busca entre la lista de palabras o token generados en el script anterior, aquellos que tengan relación con las *Etiquetas POS universales*, estas van a ser contadas y se presentarán solamente las 20 más comunes en la gráfica.
- En la segunda sección del Script se generan los parámetros fundamentales para la realización de una gráfica, la cual presenta los conceptos obtenidos.

Se refiere a *Etiquetas POS universal* a aquellas etiquetas que marcan las categorías principales de la parte del discurso. Estas son implementadas para distinguir propiedades léxicas y gramaticales adicionales de las palabras, en el Script se menciona como (*POS\_F*).

La tabla 14, reúne las *Etiquetas POS universales* que pueden ser analizadas en

este método automático gracias al modelo de lenguaje natural.<sup>167</sup>

## Tabla 14

*Tabla De Etiquetas POS*

Etiquetas POS universales			Listado	
Palabras de clase abierta	Palabras de clase cerrada	Otro	Etiqueta POS universal	Significado
<a href="#">ADJ</a>	<a href="#">ADP</a>	<a href="#">PUNCT</a>	ADJ	Adjetivo
<a href="#">ADV</a>	<a href="#">AUX</a>	<a href="#">SYM</a>	ADP	Aposición
<a href="#">INTJ</a>	<a href="#">CONJ</a>	<a href="#">X</a>	ADV	Adverbio
<a href="#">SUSTANTIVO</a>	<a href="#">DET</a>		AUX	Verbo auxiliar
<a href="#">PROPN</a>	<a href="#">NUM</a>		CONJ	Conjunción coordinadora
<a href="#">VERBO</a>	<a href="#">PARTE</a>		DET	Determinante
	<a href="#">PRON</a>		INTJ	Interjección
	<a href="#">SCONJ</a>		NOUN	Sustantivo
			NUM	Numeral
			PART	Partícula
			PRON	Pronombre
			PROPN	Nombre propio
			PUNCT	Puntuación
			SCONJ	Conjunción subordinante
			SYM	Símbolo
			VERB	Verbo
			X	Otro

*Tabla 14. Tabla de Etiquetas POS universales para el modelo de lenguaje natural*

*Nota:* En la tabla anterior se muestran las etiquetas POS las cuales son utilizadas para el modelo del lenguaje natural.

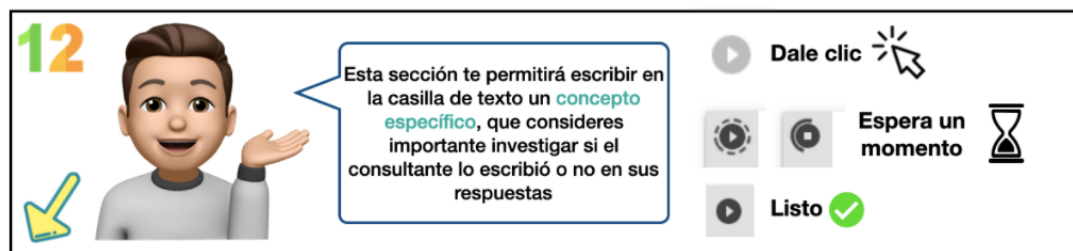
## REPETICIÓN DE UN CONCEPTO ESPECÍFICO

<sup>167</sup> Universal POS tags. (s. f.). Universaldependencies.Org. {En línea} {25 de Abril de 2021} disponible en: <https://universaldependencies.org/docs/u/pos/>

Esta sección del análisis permite por medio de funciones internas de lenguaje Python, ofrecer la búsqueda de repeticiones de un concepto específico, que el profesional del área de la salud desee investigar. La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 42 y 43 respectivamente.

## Figura 42

### Vista Del Usuario (Repetición De Un Concepto Específico)



◉ <- Veces se encontro el concepto de:

Figura 42. Vista del usuario repetición de un concepto específico. Fuente:[El Autor]

Nota: En la figura se muestra la vista del usuario en la repetición de un concepto específico

## Figura 43

### Script Repetición De Un Concepto Específico

```
#@title <- Veces se encontró el concepto de:
buscar = str(input("Que concepto desea buscar su frecuencia de uso = "))
veces = 0
buscarM = buscar.title()
veces = palabras.count(buscarM)
if buscarM in palabras:
    print("El concepto de", buscar, "se encuentra en las respuestas dada por el
usuario,", veces, "veces.")
else:
    print("El concepto de", buscar, "NO se encuentra en las respuestas dada por el
usuario.")
```

Figura 43. Script repetición de un concepto específico. Fuente:[El Autor]

Nota: En la figura se muestra el script de repetición de un concepto específico.

### Descripción:

- La primera sección de este Script abre un cuadro de texto por el cual el profesional del área de Salud Mental ingresa la palabra que desea investigar, entre las respuestas que generó el consultante.
- La segunda sección del Script realiza todo el proceso de recorrer el arreglo de palabras de las respuestas, contando la frecuencia de uso de la palabra específica que se desea investigar.
- La tercera sección del Script muestra un mensaje por el cual se le da a saber al profesional del área de salud mental, la palabra que se investigó y el número de repeticiones que se utilizó la misma en las respuestas generadas.

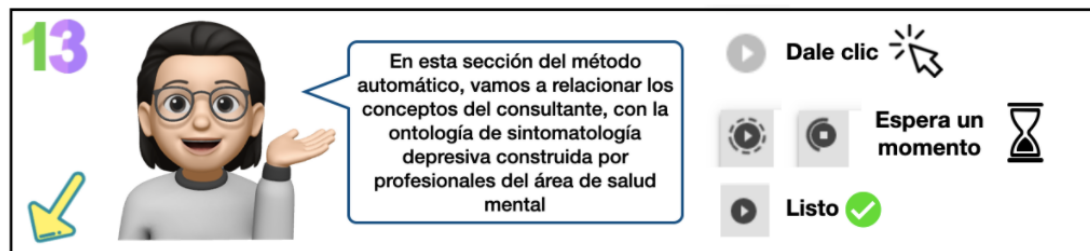
### 4.3 FASE 3. CORRELACIÓN DE INFORMACIÓN CON ONTOLOGÍA DESARROLLADA

Este es el proceso por el cual se relacionan los conceptos escritos por el consultante con los de la ontología desarrollada, esta sección se encarga de describir la lista ordenada de conceptos y palabras que presentaron una correlación con alguno de los nueve síntomas principales de la depresión. La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 44 y 45 respectivamente.

Para llevar a cabo este proceso del método automático se realiza el siguiente desarrollo:

**Figura 44**

*Vista De Usuario (Carga De Ontologia)*



<- Carga la Ontologia

*Figura 44. Vista del usuario Carga de Ontología. Fuente: [El Autor]*

*Nota: En la figura anterior se muestra la vista del usuario en la carga de la ontología.*

### FIGURA 45

## Script Carga De Ontología

```
#@title <- Carga la Ontología
Animo=["Perezoso","Desalentada","Desalentadas","Decaído","Desmotivado","Desalentado","Desalentados","Desalentó","No Tengo Ganas","No Quiero","Estoy Cansado","Me Canso","No Tengo Alientos","Pena..."] #Más conceptos
Apatia=["No Tiro","No Me Dan Ganas","Pasivo","No Siento ","No Quiero Nada","Dejadedez","Dejado","Vacío","Abandono","Desinteresado","Desánimo","Desganado","Desaliento","Descuidado..."] #Más conceptos
Alimentación=["Indiferente","Yo","Apetito","Desinterés","Emoción","Abatimiento","Dejadedez","Dejado","Vacío","Abandono","Desinteresado","Desánimo","Deganado","Desaliento","Descuidado","Todo Yo",...] #Más conceptos
Descanso=["Débil","Débiles","Debilité","Raquítico","Debilitado","Enclenque","Flojo","Apagado","Debilitada","Mamado","Debilitados","Floja","Debilitadas","Desalentada","Desalentadas","Decaído",...] #Más conceptos
Actividad_fisica=["Desaliento","Frustración","Trauma","Abatimiento","Extenuación","Incapacidad","Agitado","Débil","Débiles","Debilité","Raquítico","Debilitado","Enclenque","Flojo","Apagado","Debilitada",...] #Más conceptos
Desaliento=["Ausente","Ausentes","Ausentado","Ausencia","Alejado","Lejano","Ausentada","Ausentó","Ausentará","Ausente","Débil","Débiles","Debilité","Raquítico","Debilitado","Enclenque","Flojo","Apagado",...] #Más conceptos
Cargo_de_conciencia=["Culpable","Culpado","Culpada","Culpa","Culpar","Culpar","Culpables","Errores","Embarrarla","Cagarla","Tirarla","Cerrarla","Frustrado","Fracasado","Fiasco","Chasco","Apenado","Penoso",...] #Más conceptos
Atencion_dispersa=["Ausente","Ausentes","Ausentado","Asentados","Ausentada","Ausentadas","Ausentó","Ausentará","Ausente","Ausencia","Errores","Embarrarla","Cagarla","Tirarla","Cerrarla","Frustrado",...] #Más conceptos
Ideas_suicidas=["Desvelo","Sensible","Frágil","Miedo","Intranquilo","Desilusionado","Decepcionado","Susceptible","Aislado","Insatisfecho","Solo","Vacío","Ausente","Melancólico","Negativo","Aislamiento",...] #Más conceptos
```

Figura 45. Script Carga de Ontología. Fuente:[El Autor]

Nota: En esta figura se observa el script de la Carga de Ontología.

### Descripción:

- Este Script carga la ontología de sintomatología depresiva previamente construida por profesionales del área de salud mental e ingenios de sistemas, a manera de lista en Python, esto para que después se pueda operar esta ontología por funciones de correlación y búsqueda de conceptos específicos.

## RELACIÓN DE LA ONTOLOGÍA CON PALABRAS GENERADAS POR EL

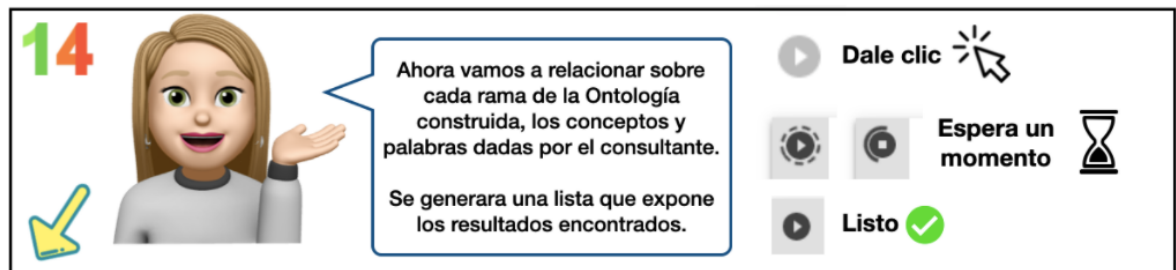
## CONSULTANTE

Este proceso es el encargado de generar un análisis de relaciones semánticas entre los conceptos que escribió el consultante en respuesta a el formulario de preguntas y los conceptos finales establecidos por la ontología de palabras de sintomatología depresiva.

Este proceso de relaciones es uno de los entregables finales de este método automático y es uno de los que tienen mayor valor, dado que con esta relación el profesional del área de la Salud Mental logra acceder a una posible predicción de cuál es el síntoma depresivo que su paciente está presentando, y a su vez esta relación es válida, ya que se sustenta con una base referencial fija, que es la ontología construida. La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 46 y 47 respectivamente.

### Figura 46

*Vista Del Usuario (Relación De La Ontología Con Palabras Generadas Por El Consultante)*



- ▶ <- Relaciona la Ontología de sintomatología depresiva, con las diferentes respuestas generadas por el paciente.

*Figura 46. Vista del usuario relación de la ontología con palabras generadas por el consultante. Fuente:[El Autor]*

*Nota:* En la anterior figura se muestra la vista del usuario con relación a la ontología.

### Figura 47



## Script Relación De La Ontología Con Palabras Generadas Por El Consultante

```
print("Palabras relacionadas con la sintomatología: @sintoma")
##### Búsqueda de relaciones
rama = []
palabras_estudiar=set(@conceptos)
alegria=set(@sintoma)
estudio= palabras_estudiar & @sintoma
##### Clasificación de relaciones
if len(estudio) > 0 :
    print("Hay {} elementos coincidentes".format(len(estudio)))
    print(estudio)
    rama.append("@concepto")
else:
    print("No hay repeticiones")
```

*Figura 47. Script relación de la ontología con palabras generadas por el consultante. Fuente:[El Autor]*

*Nota:* En la figura anterior se muestra un script de la relación de las palabras que se generaron por el consultante con la ontología.

### Descripción:

- La primera parte de Script busca relaciones entre la lista de palabras extraídas de las respuestas del consultante con relación a las palabras de cada una de las ramificaciones de sintomatología depresiva. Este proceso se hace para cada una de las nuevas ramas que componen a la Ontología.
- La segunda sección de este Script nos permite identificar a qué ramificación de la ontología hace parte del concepto que se identificó en el proceso anterior, esta información es almacenada tanto en un arreglo como en una variable para luego ser presentada en un informe de resultados.

## 4.4 FASE 4. GENERACIÓN DE INFORME

Es en este proceso en donde se genera el informe que será dado a los profesionales del área de la Salud Mental, el cual reúne toda la información analizada y destacada del proceso de correlación de conceptos con la Ontología desarrollada para este método automático.

La vista que tendrá el usuario de esta sección y el script interno de la misma pueden detallarse en las figuras 48 y 49 respectivamente.

### 4.4.1 INFORME “RELACIÓN DE CONCEPTOS CON LA ONTOLOGÍA DE

## SINTOMATOLOGÍA DEPRESIVA”

### Figura 48

Vista Del Usuario(Generación Informe De Resultados)

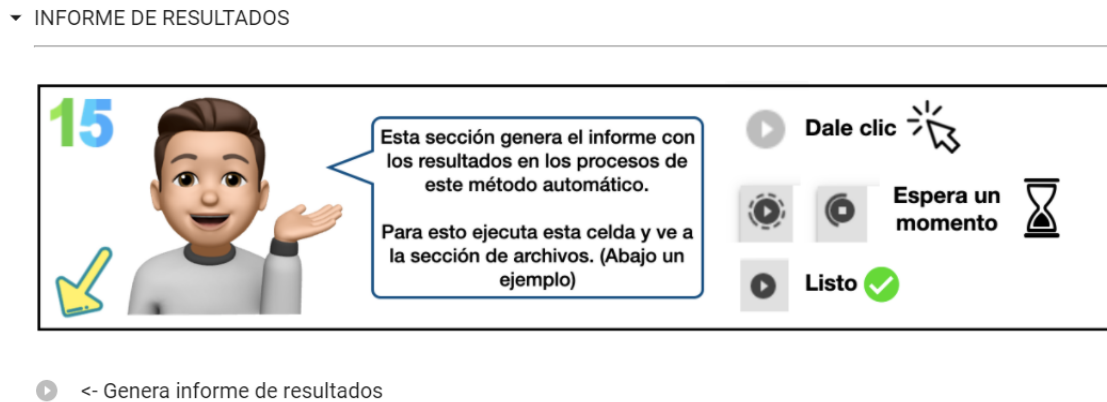


Figura 48. Vista del usuario Generación informe de resultados. Fuente:[Autor]

Nota: En la figura se puede apreciar la vista del usuario en la que se genera el informe de resultados.

### Figura 49

Script Generación Informe Resultados

```
#@title <- Genera informe de resultados
#cargar informe

contador = Counter(palabras)
mas = contador.most_common(20)
Generar_informe=open("Informe.txt","w")
Generar_informe.close()
informe = ("Este es el informe generado para el paciente con el ID : {0} \n"
"\n" "Él escribió estas palabras en sus respuestas: {1} \n"
"\n" "Al relacionar estas palabras con la Ontología encontramos que
presenta correlación con: \n""La sintomatología de : {2}""\n"
"Si anteriormente no se veía ninguna relación es porque no se
..... encontraron correlaciones.\n""\n""Las palabras más significativas dadas
..... por el consultante son: \n""{3}""\n""\nLos adjetivos que este utilizó
..... fueron: \n""{4}""\n""\nLos verbos que este utilizó fueron: \n""{5}""\n"
"\nLas palabras más usadas por el consultante son: \n""{6}""\n"
"\n""\nSi en alguno de los anteriores, te genero vacío, es porque el
..... consultante no utilizo ningún concepto\n""relacionado a esa área
..... específica.\n")
```

```

    )
informeP = informe.format(id,palabras,ramaf,Esencial,ADJ_F,VERB_F,mas)
informeF = "".join(informeP)
Cargar_informe=open("Informe.txt","a")
Cargar_informe.write(informeF)
Cargar_informe.close()

```

Figura 49. Script Generación informe de resultados. Fuente: [El Autor]

*Nota:* En la figura se muestra el script que genera un informe de resultados.

### Descripción:

- Este proceso genera un archivo de texto plano, este se escribe por medio de un Script que reúne las variables que almacena la información más significativa y destacada del proceso de análisis de información, al igual que el proceso de análisis y correlación de información entre la ontología y las respuestas dadas por el consultante.

Este informe es primordial para las consultas que llevará a cabo el profesional del área de la Salud Mental, se especifica en este informe toda la información detallada de los análisis realizados a las respuestas generadas por el consultante en el desarrollo del formulario. Se destaca que para cada consultante, se genera un informe único y diferente, dado que este es construido por medio de variables. (Para ver el archivo informe.txt diríjase al anexo 13)

Para acceder al método automático desarrollado en el entorno de Google Colaboratory, en su versión entregada puede dirigirse al siguiente enlace: [Método automático Sintomatología depresiva \(escrita\)](#)

En este podrá apreciar el desarrollo *front end*, el cual cubre las necesidades y requerimientos dados por el profesional del área de salud, en donde solicitaba que con ayuda gráfica o visual se diera una explicación continua y sencilla de como llevar a cabo estos procesos.

## 5. VALIDACIÓN DEL MÉTODO AUTOMÁTICO

En este capítulo se evalúa el rendimiento del método automático a partir de las siguientes métricas de desempeño: precisión, Recall (“Recordar”), F1-score (“Puntuación F1”) y exactitud.

Para lograr validar el método automático fue necesario entrenar un modelo implementado a través de un dataset que fue generado con base a la ontología construida. Además se muestran las herramientas utilizadas, como también, el proceso realizado para lograr el entrenamiento y validación del modelo con las métricas de rendimiento.

## 5.1 HERRAMIENTAS UTILIZADAS

En el proceso de validación del método automático se implementa la herramienta de *Monkeylearn*, esta herramienta ofrece una inteligencia artificial especializada en textos, ofreciendo distintos tipos de modelos. Adicionalmente, maneja sus propios datasets integrados, a su vez, posee una característica particular la cual consiste en entrenar un modelo propio mediante un dataset que puede ser cargado en los formatos *CSV* y *Excel*.

Este tipo de dataset puede ser entrenado con 2 tipos de algoritmos *SVM* y *Naive Bayes*, sin embargo, para este proyecto se escogió *SVM* debido a la facilidad que este tiene al momento de clasificar textos y, adicionalmente, por su buen rendimiento en distintos tipos de datasets.<sup>168</sup> Para poder implementar *monkeylearn* en el proceso de validación del método automático, se integró una sección de código en el desarrollo del Script, ya que esta herramienta tiene su propia librería la cual se puede utilizar en distintos tipos de lenguajes de programación, en este caso se utiliza para *Python* y la librería que se destaca en la Figura 50. Script Implementación de librería *MonkeyLearn*.

### Figura 50

*Script Implementación Librería Monkeylearn*

```
from monkeylearn import MonkeyLearn
```

*Figura 50. Script Implementación de librería MonkeyLearn. Fuente:[El Autor]*

*Nota:* En la figura anterior se muestra el script que permite la implementación de *MonkeyLearn*

### Figura 51

---

<sup>168</sup> *MonkeyLearn API Reference.* (s. f.). monkeylearn. [En línea] {27 de Abril de 2021} disponible en: <https://monkeylearn.com/api/v3/>

## Implementación De Librerías En Python 1

Se instala la librería necesaria para que el algoritmo funcione, presionando el boton en el cuadro rojo:

```
!pip install monkeylearn
```

```
[ ] !pip install monkeylearn
```

```
Collecting monkeylearn
  Downloading https://files.pythonhosted.org/packages/85/71/402de0a734641f015facd3d
Requirement already satisfied: requests>=2.8.1 in /usr/local/lib/python3.7/dist-packa
Requirement already satisfied: six>=1.10.0 in /usr/local/lib/python3.7/dist-package:
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/loca
```

Figura 51. Implementación de librerías en cuadernillo Google Colab 1. Fuente:[El Autor]

*Nota:* La implementación de *Python* se realizó como se aprecia en la Figura 51. Implementación de librerías en cuadernillo Google Colab

En este punto se instala la librería necesaria para lograr implementar el modelo en *google colab*, tal como se aprecia en la Figura 52. Implementación de librerías en cuadernillo Google Colab 2.

## Figura 52

### Implementación De Librerías En Python 2

```
from monkeylearn import MonkeyLearn
print("Texto a clasificar: ")
texto=input()
ml = MonkeyLearn('e0deb7410556d3271c8f53f54f7ef0eae9e24ef2')
data = [texto]
model_id = 'cl_39uqPrW6'
result = ml.classifiers.classify(model_id, data)
print(result.body)
```

Se escribe el texto a analizar:

Texto a clasificar:

Figura 52. Implementación de librerías en cuadernillo Google Colab 2. Fuente:[El Autor]

*Nota:* En esta figura se muestra cómo se implementa la librería de MonkeyLearn en Python. La Figura 53. Implementación de librerías en cuadernillo Google Colab 3, representa el proceso de implementar el código con la librería anteriormente

instalada, adicionalmente se almacena una variable texto, la cual llena el usuario, y se clasifica por medio del modelo entrenado en *Monkeylearn*

## Figura 53

### Implementación De Librerías En Python 3

Despues muestra el texto y el tipo de sintoma que presenta, como se muestra en los cuadros rojos:

```
Texto a clasificar:  
Me quedo grande  
[{"text": "Me quedo grande", "external_id": None, "error": False, "classifications": [{"tag_name": "Disminución de interés o placer"}]
```

```
from monkeylearn import MonkeyLearn  
print("Texto a clasificar: ")  
texto=input()  
ml = MonkeyLearn('e0deb7410556d3271c8f53f54f7ef0eae9e24ef2')  
data = [texto]  
model_id = 'cl_39uqPrW6'  
result = ml.classifiers.classify(model_id, data)  
print(result.body)
```

Figura 53. Implementación de librerías en cuadernillo Google Colab 3.Fuente:[El Autor]

*Nota:*En la figura anterior se aprecia como es la clasificación que devuelve el modelo.

El modelo nos devuelve un objeto tipo *Json* en el cual se almacena información importante como el texto que fue escrito y la etiqueta en la que se clasifica ese texto según el modelo entrenado.

## 5.2 PROCESO DE LA HERRAMIENTA.

El proceso que fue utilizado, en primera medida, para desarrollar el modelo en la herramienta *Monkeylearn*, inició con la selección de la opción de creación de un modelo, desde su [aplicación web](#).<sup>169</sup>

Después, la herramienta permite la elección del tipo de clasificación que se va a realizar, en este caso, debido a la naturaleza de los datos a trabajar, se selecciona una clasificación por tópicos. Así mismo, *MonkeyLearn* permite seleccionar qué columna va a ser utilizada para usarse como etiqueta y como texto, en este caso, de acuerdo a los síntomas encontrados en el análisis realizado mediante la ontología, tenemos como texto aquellas oraciones relacionadas de acuerdo al tipo de sintomatología encontrada, tal como se puede observar en la Figura 54. Proceso de construcción de métricas.

## Figura 54

<sup>169</sup> MonkeyLearn API Reference. (s. f.). monkeylearn.[En línea] {27 de Abril de 2021} disponible en: <https://monkeylearn.com/api/v3/>

## Proceso De Construcción De Métricas

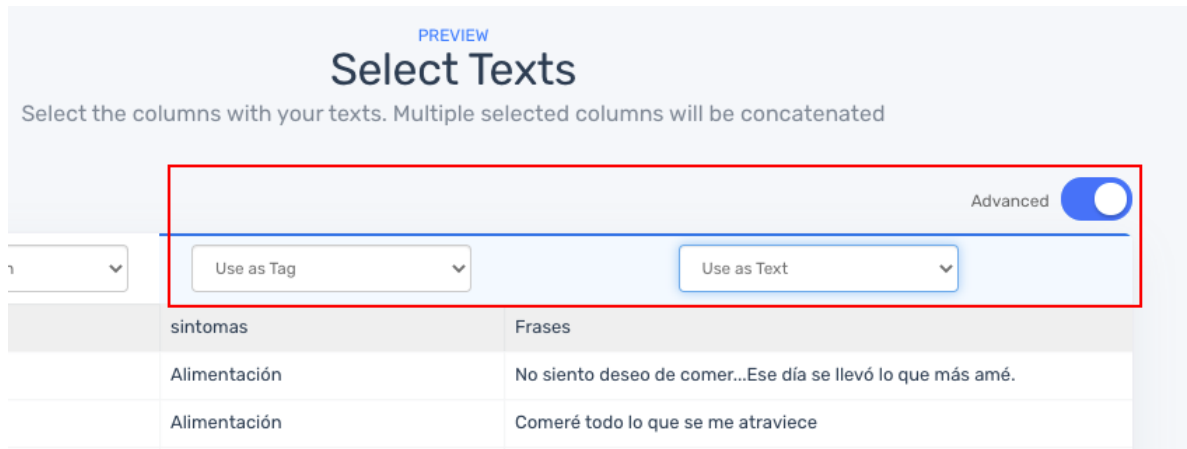


Figura 54. Proceso de construcción de métricas 1. Fuente:[El Autor]

*Nota:* En la figura anterior se muestra la primera parte del proceso con el cual se construye la métrica.

Adicionalmente, *MonkeyLearn* permite probar el modelo entrenado digitando textos de prueba manualmente, siendo clasificados en la parte de TAG lo cual se puede apreciar en el cuadro rojo de la figura siguiente, en esta parte se utiliza el modelo y se tienen en cuenta las etiquetas utilizadas para la clasificación. Esto se aprecia en la Figura 55. Proceso de construcción de métricas 2.

## Figura 55

### Proceso De Construcción De Métricas

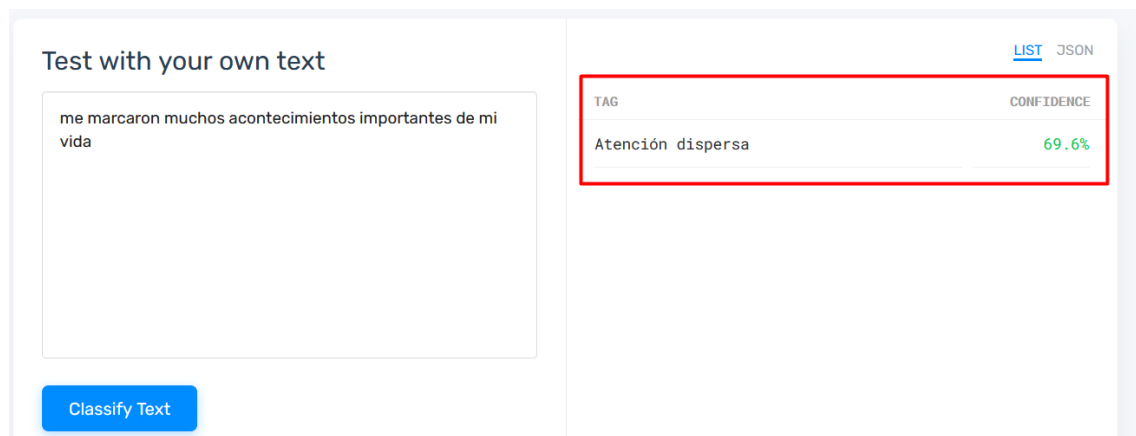


Figura 55, Proceso de construcción de métricas 2. Fuente:[El Autor]

*Nota:* Se aprecia el último paso para la construcción de la métrica.

### 5.3 MEDIDAS DE RENDIMIENTO OBTENIDAS

Las métricas que se dieron en el modelo entrenado en *MonkeyLearn* se encuentran en cada etiqueta mostrada, así mismo, el rendimiento se puede hallar a través de las métricas mostradas por la herramienta, las cuales son: precisión y recall, que miden la calidad y la cantidad respectivamente.

De manera adicional, se evalúa el modelo con todas las etiquetas que han sido utilizadas, dentro de las métricas de *exactitud* y *F1 Score*. En este sentido, *exactitud* permite saber el porcentaje que el modelo ha acertado correctamente y *F1 Score* es una métrica que utiliza *precisión* y *recall* en un solo valor, permite comparar el rendimiento de estas.

#### 5.3.1 Etiqueta Actividad Física

En esta etiqueta la métrica de *precisión* dio un resultado del 68%, esto indica el porcentaje de los textos que se pueden estar clasificar en esta etiqueta, la métrica *recall* da un valor del 76%, significando el porcentaje que se puede clasificar con los textos que realmente tengan este síntoma, adicionalmente se tienen en cuenta los siguientes valores: 19 *Verdaderos Positivos* textos que fueron clasificados en esta etiqueta de manera correcta, 260 *Verdaderos negativos* textos que no están clasificados en esta etiqueta, 9 *Falsos Positivos* textos que fueron clasificados en esta etiqueta de manera incorrecta y 6 *Falsos Negativos* textos que no fueron clasificados en esta etiqueta, cuando realmente deberían estar en esta.

Esto significa que el rendimiento del modelo al predecir esta etiqueta aproximadamente de 10 textos 7 los clasificaría de manera correcta al presentar las características de poder ser clasificados en la *etiqueta de actividad física*, las oraciones utilizadas en el dataset estaban correctamente balanceadas y cada etiqueta tiene una cantidad similar.

Además, como se puede observar en la imagen la palabra clave de esta etiqueta es *pereza*. El resultado gráfico obtenido para la etiqueta de Actividad Física, puede observarse en la Figura 56. Etiqueta Actividad Física.



## Figura 56

### Etiqueta Actividad Física

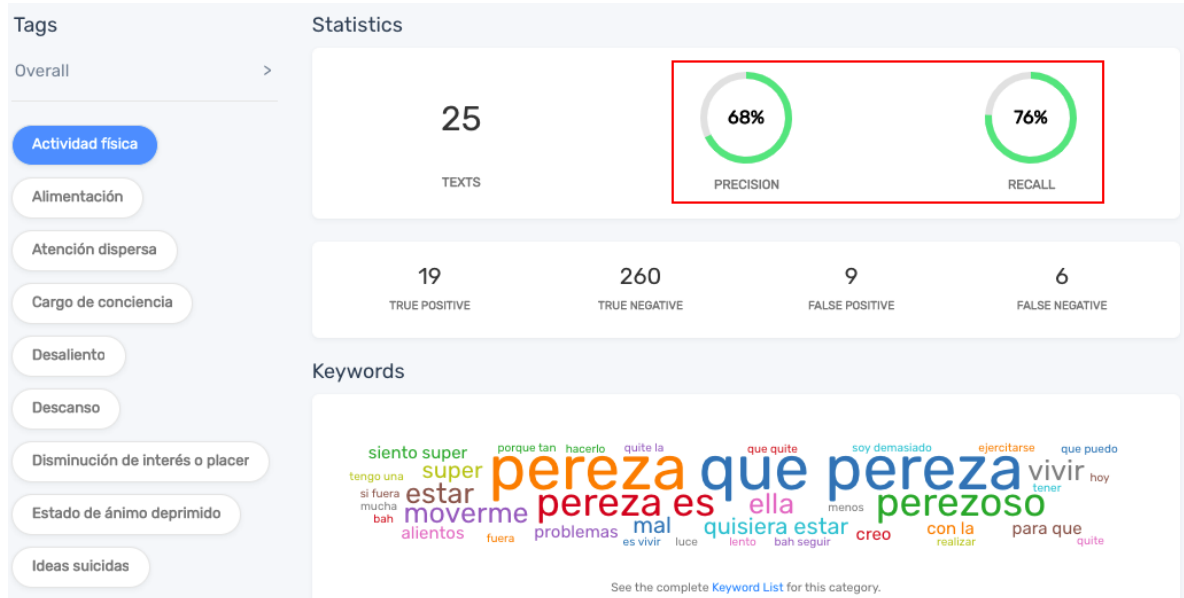


Figura 56. Etiqueta Actividad Física. Fuente: [El Autor]

Nota: Se muestra la figura de la etiqueta de actividad física junto con sus métricas.

### 5.3.2 Etiqueta Alimentación

En la etiqueta de alimentación se obtuvieron mejores resultados, los cuales fueron los siguientes una *precisión* del 83% este porcentaje quiere decir la calidad que tiene esta etiqueta al momento de clasificar, y la métrica *recall* de 83%, esto significa la cantidad que es capaz de clasificar, con 24 *verdaderos positivos* textos clasificados en esta etiqueta de manera correcta, 260 *verdaderos negativos* cantidad de textos que no pertenecen a esta etiqueta y fueron clasificados de manera acertada, 5 *falsos positivos* cantidad de textos que no pertenecen a esta etiqueta y que fueron clasificados como si hicieran parte, y 5 *falsos negativos* textos que debieron ser clasificados de manera positiva pero se clasificaron como si no pertenecieran a esta etiqueta.

El rendimiento de esta etiqueta fue muy bueno debido al porcentaje alto en *precisión* y *recall*, lo cual quiere decir que de 10 textos que deban ser clasificados en esta etiqueta 8 serán correctos, la palabras clave de esta etiqueta es *comer*. El resultado gráfico obtenido para la etiqueta de Alimentación, puede observarse en la Figura 57. Etiqueta Alimentación

## Figura 57

### Etiqueta Alimentación

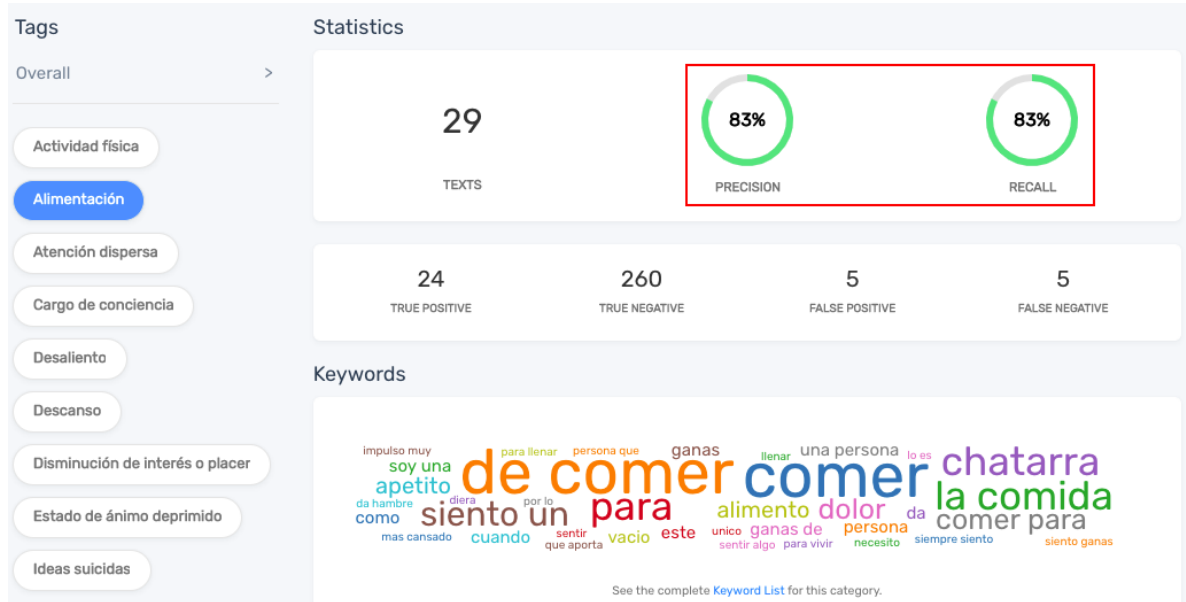


Figura 57. Etiqueta Alimentación. Fuente:[El Autor]

Nota: En esta figura se muestra la etiqueta de Alimentación junto con sus métricas de rendimiento.

### 5.3.3 Etiqueta de Atención Dispersa

La siguiente etiqueta nos muestra un desempeño bajo con una *precisión* del 63% este porcentaje muestra la calidad con la cual la etiqueta clasifica y 61% de la métrica de *recall*, mostrando la cantidad de textos que es capaz de clasificar, además 17 *verdaderos positivos* textos que fueron clasificados en esta etiqueta de manera correcta, 256 *verdaderos negativos* textos que verdaderamente no pertenecen a la etiqueta, 10 *falsos positivos* textos que fueron clasificados de manera falsa en el escenario que fueron clasificados en esta etiqueta, cuando realmente no hacen parte de esta y 11 *falsos negativos* cantidad de textos que fueron clasificados de manera falsa como si no pertenecieran a esta etiqueta cuando realmente si pertenecen a la *etiqueta de Atención Dispersa*, de manera adicional se observa que las palabra clave en esta etiqueta es *concentrarme*. Estos resultados significan que de 10 textos 6 se clasificaron de manera correcta en esta etiqueta.

El resultado gráfico obtenido para la etiqueta de Atención Dispersa, puede observarse en la Figura 58. Etiqueta de Atención Dispersa.

## Figura 58

### Etiqueta De Atención Dispersa

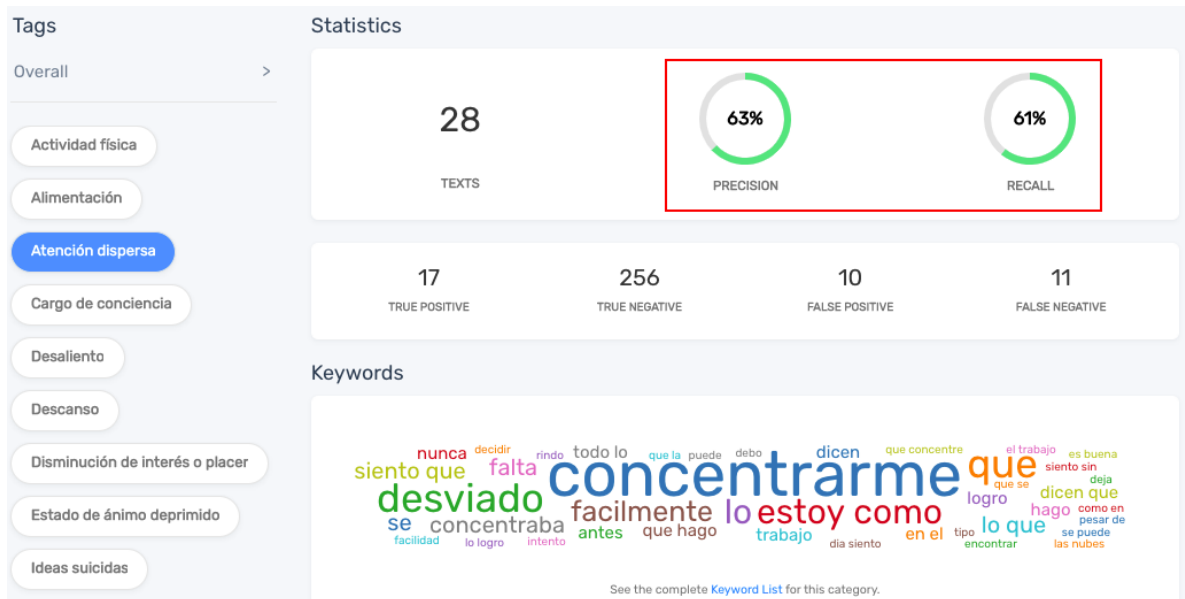


Figura 58. Etiqueta de Atención Dispersa. Fuente:[El Autor]

Nota: Se muestra la etiqueta de Atención Dispersa junto con sus métricas de desempeño.

### 5.3.4 Etiqueta de Cargo de Conciencia

La etiqueta de cargo de conciencia presento los siguientes resultados, una *precisión* del 83% mostrando el porcentaje de la calidad que esta etiqueta presenta al clasificar textos en esta etiqueta, además, con un resultado de la métrica *recall* del 78% mostrando la cantidad de textos que se pueden clasificar en la *etiqueta de Cargo de Conciencia*, con 25 *verdaderos positivos* textos que fueron clasificados de manera correcta en esta etiqueta, 257 *verdaderos negativos* textos que no pertenecían a esta etiqueta y fueron clasificados de manera correcta, 5 *falsos positivos* textos que fueron clasificados de manera incorrecta, en este caso se clasificaron como si pertenecieran a la etiqueta pero no pertenecían, 7 *falsos negativos* textos que hacen parte de la *etiqueta de cargo de conciencia* pero fueron clasificados como si no lo fueran , la palabra clave de esta etiqueta fue *culpa*. Teniendo en cuenta estos resultados, de cada 10 textos 8 se clasificaron de manera correcta en esta etiqueta.

El resultado gráfico obtenido para la etiqueta de Cargo de Conciencia, puede observarse en la Figura 59. Etiqueta de Cargo de Conciencia.

## Figura 59

### *Etiqueta De Cargo De Conciencia*

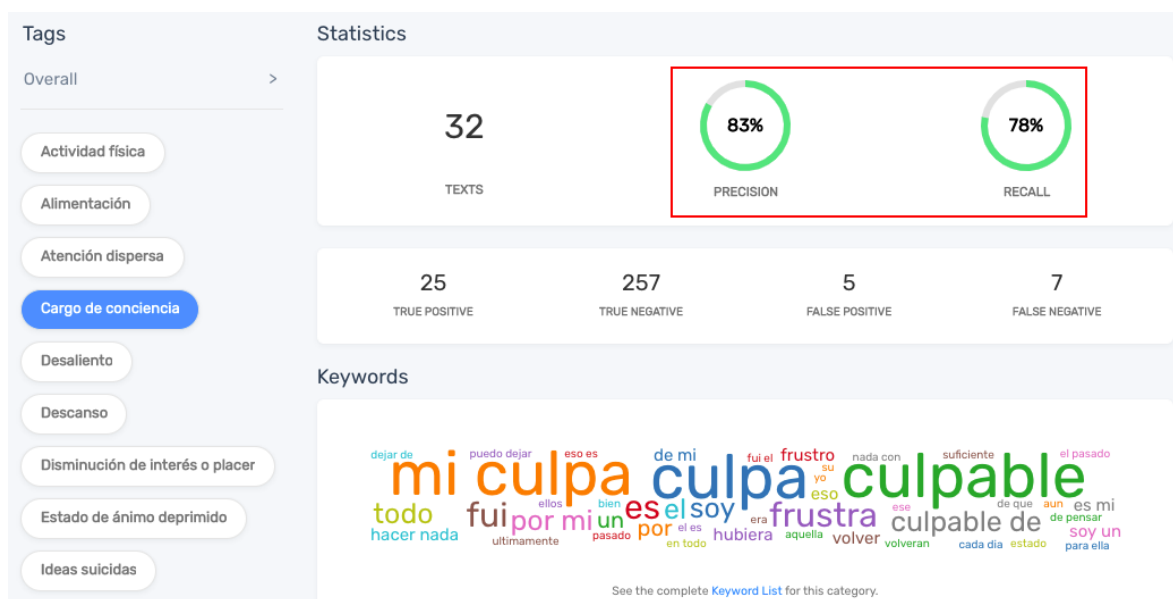


Figura 59. Etiqueta de Cargo de Conciencia. Fuente:[El Autor]

Nota: En esta figura se muestra la etiqueta de cargo de conciencia con sus métricas.

### 5.3.5 Etiqueta de Desaliento

En esta etiqueta se dieron los siguientes resultados en las métricas de desempeño, una *precisión* del 72% este porcentaje quiere decir la calidad del modelo al clasificar en esta etiqueta, un resultado de *recall* del 72%, mostrando el porcentaje de la cantidad que se pueden clasificar en la *etiqueta de Desaliento*, se obtuvieron 18 *verdaderos positivos* textos clasificados de manera correcta en esta etiqueta, 262 *verdaderos negativos* textos que no pertenecen a esta etiqueta y fueron clasificados de manera asertiva, 7 *falsos positivos* textos que fueron clasificados de manera errónea, 7 *falsos negativos* textos clasificados de manera equivocada, ya que pertenecían a la etiqueta pero fueron clasificados como que no pertenecieran a la *etiqueta de Desaliento*, tuvo como palabras clave *decaído* y *debilidad*.

El resultado gráfico obtenido para la etiqueta de Desaliento, puede observarse en la Figura 60. Etiqueta de Desaliento.

## Figura 60

### Etiqueta de Desaliento

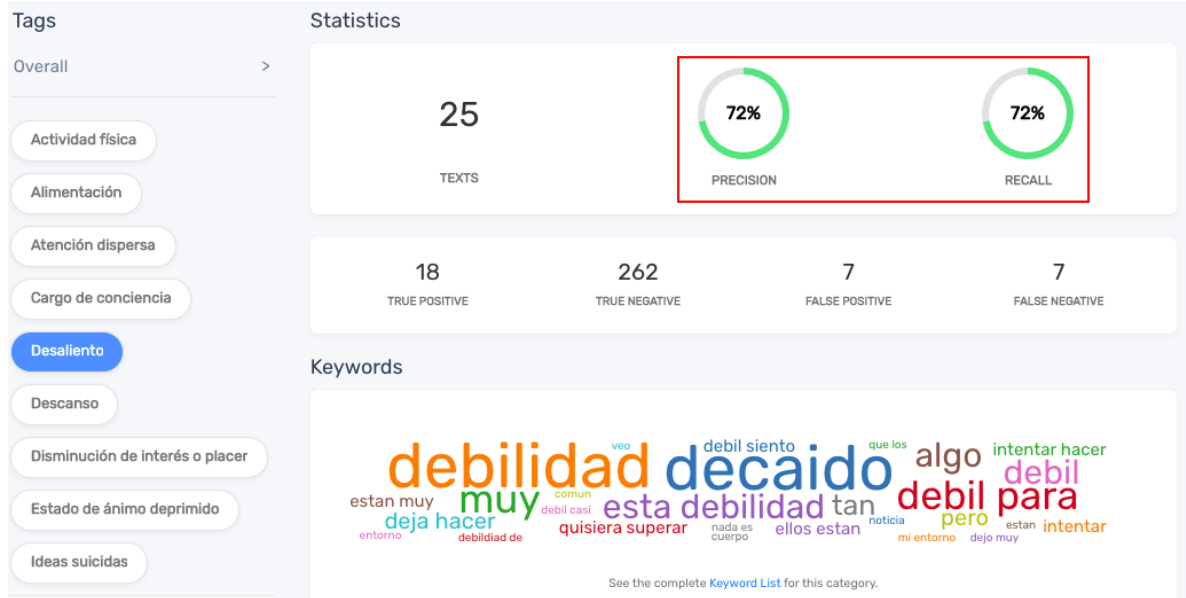


Figura 60. Etiqueta de Desaliento. Fuente:[El Autor]

Nota: En esta figura se muestra la etiqueta de desaliento en conjunto con sus métricas.

### 5.3.6 Etiqueta de Descanso

Los resultados obtenidos en la etiqueta de Descanso, mostraron una *precisión* del 72% este porcentaje muestra la calidad al momento de clasificar en esta etiqueta y un 77% de *recall* este porcentaje determina la cantidad de textos que se pueden clasificar de manera asertiva en esta etiqueta , mostrando 23 predicciones *verdaderas positivas* textos que se clasificaron en esta etiqueta de manera correcta, 255 *verdaderos negativos* textos que no pertenecen a la etiqueta y fueron clasificados de manera correcta, 9 *falsos positivos* textos que fueron clasificados de manera errónea debido a que no pertenecían a la etiqueta y fueron clasificados como si lo fuera, y 7 *falsos negativos* textos que fueron clasificados como si no pertenecieran a la etiqueta, cuando realmente pertenecían, las palabras a resaltar son *duermo* y *noches*, se tuvieron porcentajes promedio a comparación con las otras etiquetas.

El resultado gráfico obtenido para la etiqueta de descanso, puede observarse en la Figura 61. Etiqueta de descanso.

## Figura 61

### Etiqueta de Descanso

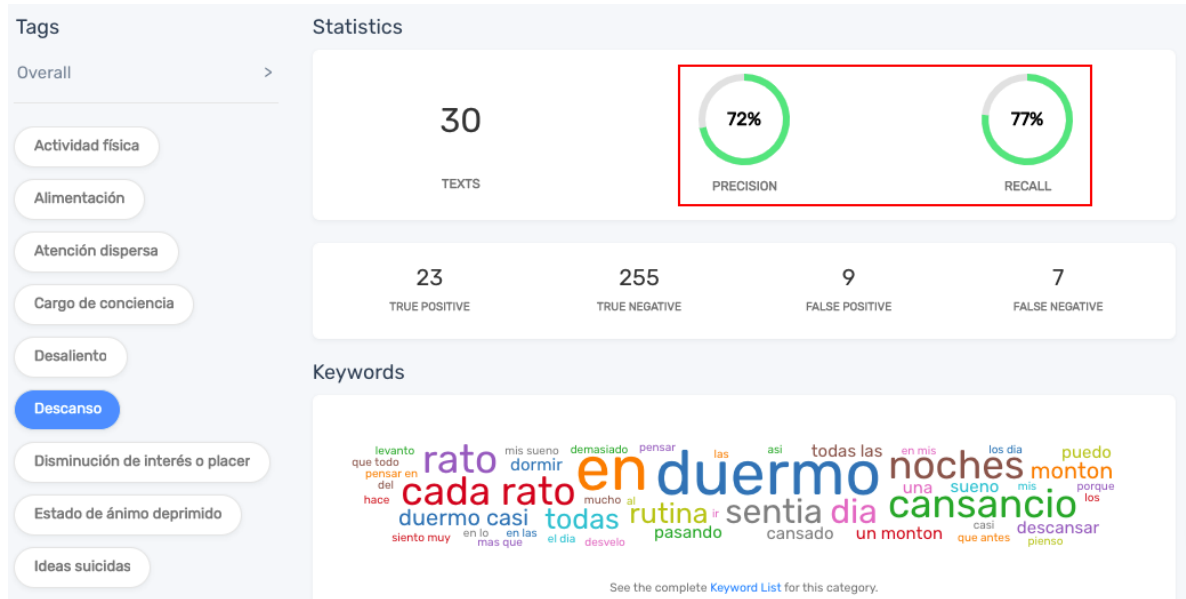


Figura 61. Etiqueta de Descanso. Fuente:[El Autor]

Nota: En la figura se muestra la etiqueta de descanso junto con sus métricas.

### 5.3.7 Etiqueta Disminución de interés o placer

Con la etiqueta de disminución de interés o placer se obtuvieron los siguientes resultados, una *precisión* del 88% lo cual tiene como significado que la calidad con la cual se clasifica con esta etiqueta es muy alta, con un resultado de la métrica *recall* del 83% este porcentaje significa la cantidad que se clasifica con la etiqueta Disminución de interés o placer, estos son resultados de buen desempeño, De manera adicional las palabras claves de esta etiqueta son *interés* y *placer*, se tienen 29 *verdaderos positivos* textos clasificados de manera adecuada en esta etiqueta , 255 *verdaderos negativos* textos que no pertenecen a esta etiqueta y fueron clasificados de manera correcta, 4 *falsos positivos* textos clasificados de manera errónea como si pertenecieran a esta etiqueta pero realmente no lo están, 6 *falsos negativos* textos clasificados de manera que no pertenecieran a la *etiqueta de Disminución de Interés o placer*, pero realmente si pertenecían. Con los resultados anteriormente mencionados se puede decir que de manera aproximada 10 textos que tengan las características para ser clasificados en esta etiqueta 9 serán clasificados de manera correcta.

El resultado gráfico obtenido para la etiqueta de Disminución de interés o placer, puede observarse en la Figura 62. Etiqueta Disminución de interés o placer.

## Figura 62

### Etiqueta Disminución De Interés o Placer.

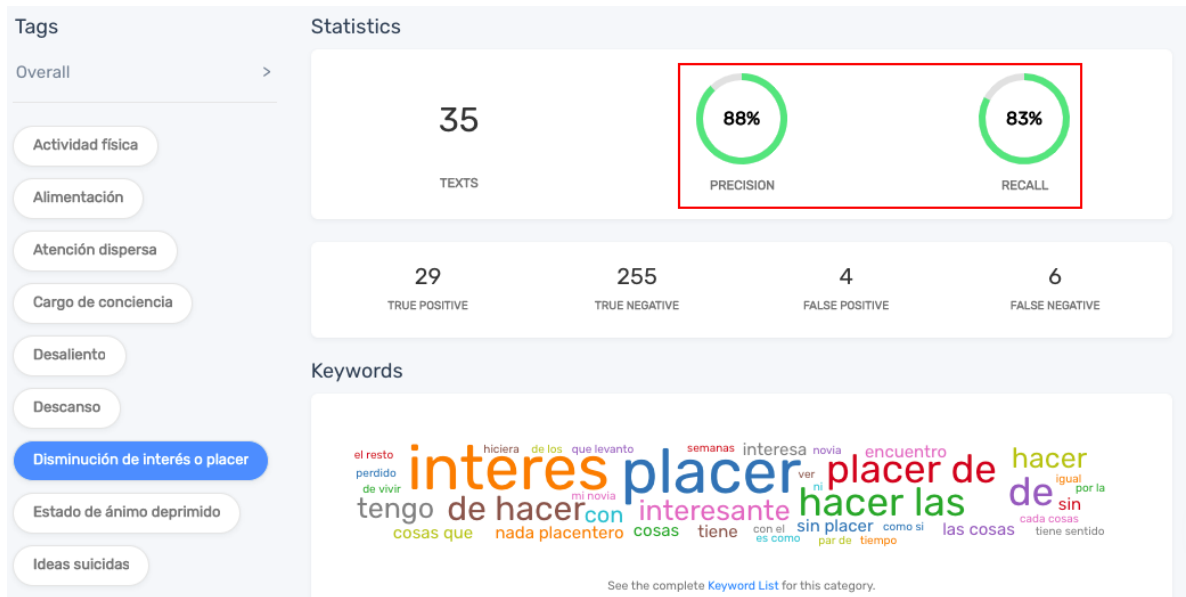


Figura 62. Etiqueta Disminución de interés o placer. Fuente:[El Autor]

Nota: Se aprecia la etiqueta de disminución de interés o placer junto a sus métricas.

### 5.3.8 Etiqueta De Estado De Ánimo Deprimido

En la etiqueta de estado de ánimo deprimido se obtuvieron los siguientes resultados con una *precisión* del 80% lo cual quiere decir que la calidad del modelo es basada en este porcentaje por tanto cuantos casos que tengan características de estado de ánimo serán clasificados de manera eficaz, un porcentaje de *recall* del 79%, significando el porcentaje que representa la cantidad la cual el modelo podrá clasificar en esta etiqueta, con 41 *verdaderos positivos* siendo el número de textos clasificados en esta etiqueta de manera adecuada, 232 *verdaderos negativos* textos clasificados de manera adecuada que no pertenecían a esta etiqueta, 10 *falsos positivos* textos que fueron clasificados como positivos en esta etiqueta, cuando en realidad eran negativos y 11 *falsos negativos* textos que fueron clasificados como negativos en la etiqueta, pero en realidad eran positivos, las palabras claves son *deprimido* y *tristeza*.

El resultado gráfico obtenido para la etiqueta de Estado De Ánimo Deprimido, puede observarse en la Figura 63. Etiqueta De Estado De Ánimo deprimido.

### Figura 63

#### *Etiqueta De Estado De Ánimo Deprimido*

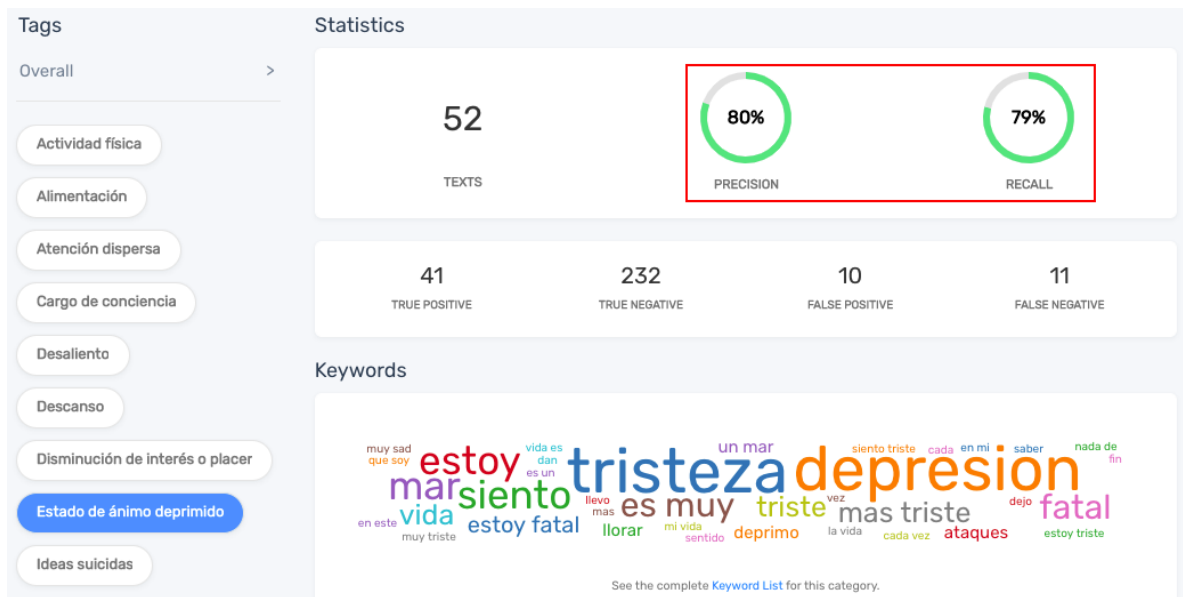


Figura 63. Etiqueta De Estado De Ánimo deprimido. Fuente[El Autor]

Nota: En la figura anterior se aprecia la etiqueta de estado de ánimo deprimido con sus métricas.

### 5.3.9 Etiqueta De Ideas Suicidas

Se obtuvieron los siguientes resultados con la etiqueta de ideas suicidas, una *precisión* del 77% mostrando el porcentaje de calidad del modelo respecto a la *etiqueta de Ideas Suicidas* y un resultado de la métrica de *recall* del 79% indicando el porcentaje de la cantidad de textos que pueden ser clasificados de manera óptima en esta etiqueta, se tienen 30 *verdaderos positivos*, textos clasificados de manera positiva en esta etiqueta, 247 *verdaderos negativos*, textos clasificados de manera negativa y correcta, 9 *falsos positivos*, textos clasificados de manera positiva erróneamente, y 8 *falsos negativos* textos clasificados negativamente de manera errónea, las palabras clave de esta etiqueta fueron *muerte* y *suicidio*. Con los resultados obtenidos en esta etiqueta 8 de cada 10 textos serán clasificados de manera correcta.

El resultado gráfico obtenido para la etiqueta de Ideas Suicidas, puede observarse en la Figura 64. Etiqueta De Ideas Suicidas.



## Figura 64

### Etiqueta De Ideas Suicidas

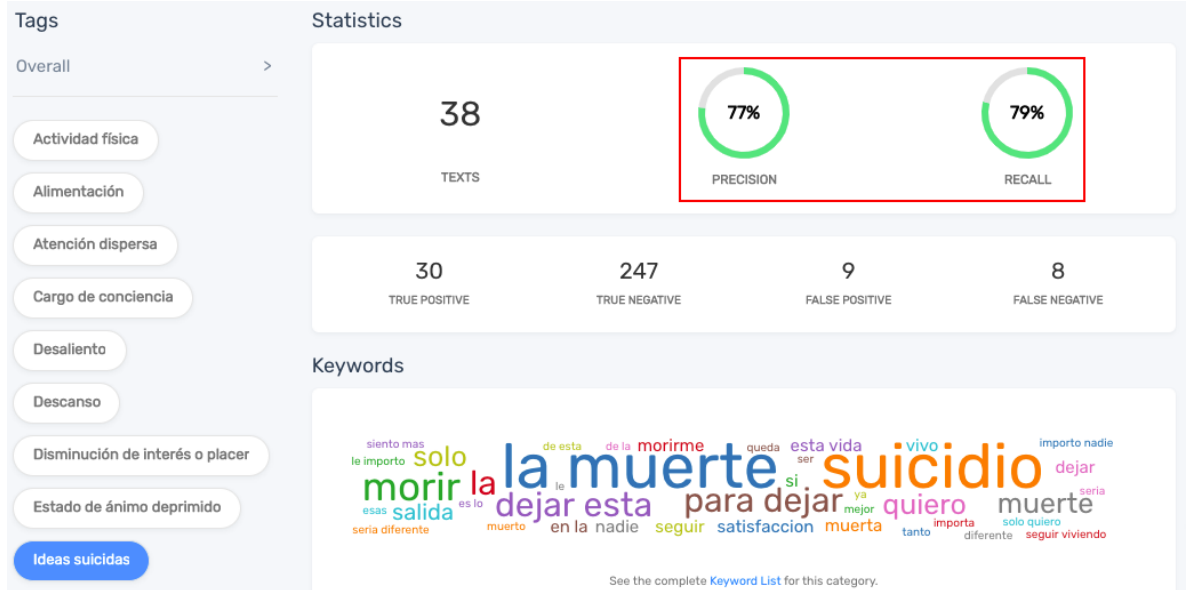


Figura 64. Etiqueta De Ideas Suicidas. Fuente:[El Autor]

Nota: En la figura se aprecia la etiqueta de ideas suicidas junto con sus métricas.

## 5.4 TABLA DE RESULTADOS DE LAS MÉTRICAS

En la *tabla 15 de resultados de etiquetas* se observan distintos resultados los cuales ayudan a validar el modelo entrenado, como se puede ver se tienen 9 etiquetas las cuales tuvieron distintos resultados, a resaltar las etiqueta de *alimentación* y la etiqueta de *disminución de interés o placer*, ya que fueron las más precisas del modelo mostrando resultados de más del 80% en las métricas de rendimiento, las cuales son *Precision*, *Recall*, *Exactitud* y *F1-Score*, esto significa que si alguno de los textos tiene las características para ser etiquetado con uno de estos síntomas de 10 textos 8 serán clasificados de manera efectiva, la etiqueta de *atención dispersa* fue la que mostró el rendimiento más bajo como se puede observar en la *Tabla de Resultados de las Métricas* con unos resultados más bajos del 65% en las métricas de precisión, recall y F1-Score en esta tabla se tienen la cantidad de textos pertenecientes a cada etiqueta, como también las métricas de *precisión*, *recall*, *exactitud* y *F1-Score*, con una exactitud de 92.86%, esto significa que de 10 textos de manera aproximada 6 serán clasificados de manera correcta en la etiqueta de *atención dispersa* del adicionalmente se muestra el promedio general del modelo con todas las etiquetas.

**Tabla 15***Métricas De Rendimiento Por Etiqueta*

MÉTRICAS DE RENDIMIENTO POR ETIQUETA					
ETIQUETA	TEXTO	PRECISION	RECALL	EXACTITUD	F1-SCORE
Actividad Física	25	68%	76%	94,90%	71,77%
Alimentacion	29	83%	83%	96,60%	83,00%
Atención Dispersa	28	63%	61%	92,86%	61,98%
Cargo de Conciencia	32	83%	78%	95,92%	80,42%
Desaliento	25	72%	72%	95,24%	72,00%
Descanso	30	72%	77%	94,56%	74,41%
Disminución de interés o placer	35	88%	83%	96,60%	85,42%
Estado de Ánimo Deprimido	52	80%	79%	92,86%	79,49%
Ideas Suicidas	38	77%	79%	94,22%	77,98%
<b>PROMEDIO</b>	<b>33</b>	<b>76,22%</b>	<b>76,44%</b>	<b>94,86%</b>	<b>76,27%</b>

*Tabla 15. Tabla de Resultados de las Métricas. Fuente: [El autor]*

*Nota:* En la tabla anterior se aprecian los resultados de todas las etiquetas del proyecto con sus respectivas métricas.

## 6. RESULTADOS Y ANÁLISIS DE RESULTADOS

Entre los procesos e implementaciones que se llevaron a cabo para la construcción del método automático que apoye en la identificación de sintomatología depresiva, encontramos una Ontología que cubre un total de 624 conceptos que identifican la posible presencia de depresión. Así mismo, el método automático integra un total de 11 funcionalidades específicas que permiten un manejo de datos para procesos de identificación y clasificación de conceptos con relación a 9 síntomas generales de la depresión. Con lo anterior se genera un informe de resultados, los cuales tienen una aprobación promedio del 77% de efectividad de las calificaciones generadas.

Detallando los procesos e implementaciones anteriormente señaladas encontramos que la construcción de la ontología desarrollada en este proyecto de grado, que es uno de los procesos más significativos del mismo, puesto que dicha ontología está construida por un equipo interdisciplinario conformado por profesionales de la psicología y estudiantes de Ingeniería de sistemas y computación, que mediante su trabajo colaborativo han generado una alternativa para los procesos de detección y tratamiento de la depresión.

Esta ontología permite tener una visualización de los 9 principales síntomas depresivos identificados mediante el análisis de los términos utilizados por los bogotanos.

Destacando los valores cuantitativos derivados de la ontología podemos decir que esta inició con la recolección de más de 1000 palabras relacionadas a los 9 síntomas de depresión. Posteriormente, se realizó un proceso de depuración y clasificación hasta que se obtuvo un total de 9 jerarquías de *clases* e *instancias*, manteniendo internamente alrededor de 624 conceptos interrelacionados, tal como se aprecia en la *Figura 65. Total de conceptos por síntoma estipulados en la Ontología.*

**Figura 65**

*Total Conceptos Por Síntoma*

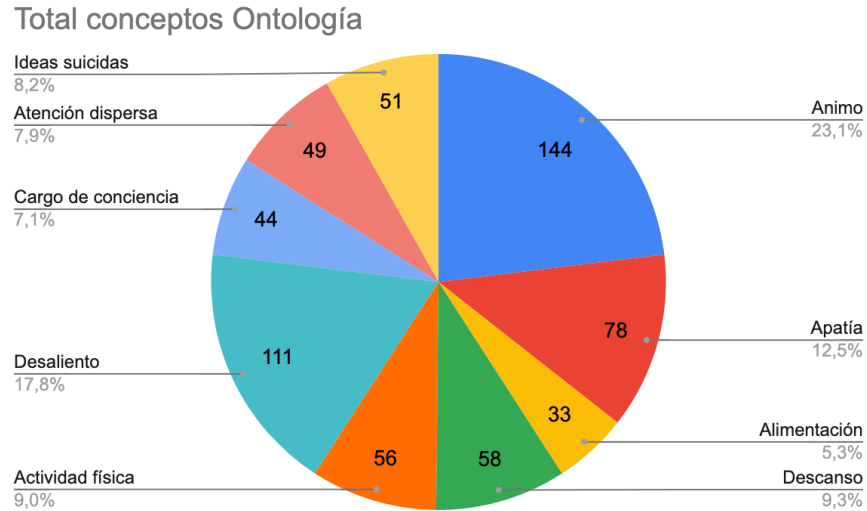


Figura 65. Total de conceptos por síntoma estipulados en la Ontología. Fuente:[El Autor]

Nota: En la anterior figura se tienen en cuenta los conceptos estipulados por la ontología construida y se muestran en forma de gráfica.

En cuanto a la construcción de la Ontología puede decirse que permite compartir el conocimiento de una manera estructurada y clara, reconociendo lo que se pretende desde el área de salud mental y apoyándolo desde el trabajo conjunto y los procedimientos propios de la ingeniería de sistemas y computación. En este sentido, la ontología permite la reutilización del conocimiento sobre este dominio que puede ser implementado en futuras investigaciones, constituyéndose como un referente investigativo en el campo de conocimiento sistémico y, específicamente, en el campo de la sintomatología depresiva.

Con respecto a la novedad, se destaca que el enfoque que presenta esta ontología es único para esta área de estudio de la depresión para adultos jóvenes bogotanos, puesto que en la revisión documental de las investigaciones no se encontró una propuesta con este mismo enfoque, razón por la cual se implementó la metodología “*Ontology Development 101*” que permitió la interdisciplinariedad en los trabajos relacionados con los departamentos de psicología e ingeniería, permitiendo que la propuesta construida fuera innovadora, constituyéndose como un aporte para los programas de diagnóstico y tratamiento de salud mental en nuestro país.

Sin embargo, a pesar de haber generado un considerable número de conceptos se espera que este aumente gracias a labores constantes de selección de otros

conceptos los cuales posiblemente identifiquen algún síntoma depresivo, estos procesos de identificación serán llevados a cabo una vez que se implemente de manera formal el método automático en las consultas o en diferentes proyectos institucionales como lo son desarrollo videojuegos en el área, recolección de información en redes sociales, entre otros.

Los diseños que fueron llevados a cabo para el desarrollo del método automático permiten obtener información general del sistema a trabajarse, el objetivo de estos diseños consiste en representar procesos de obtención, manejo, aseguramiento, almacenamiento y consulta de los datos relacionados con el sistema. Este proceso de diseño es fundamental para el desarrollo de un software que, en este caso, se integra a una ontología de sintomatología depresiva y con los procesos de análisis generando un informe descriptivo de los resultados obtenidos.

El diagrama de diseño general del método automático que se estipuló para la construcción del método automático, podemos identificar que se implementa un total de 7 librerías, 11 funcionalidades específicas y 5 algoritmos de computación. Esto ha sido trabajado con 5 necesidades y requerimientos solicitados por parte de los profesionales del área de salud mental, 6 diseños por sección específica y 7 pseudocódigos para el correcto entendimiento y buena guía de ejecución.

Los diseños aquí presentados son la base fundamental para el desarrollo futuro de un producto software específico, ya que, con los diagramas de flujo estipulados se permite un entendimiento del proceso que puede ser aplicado a diferentes desarrollos de software que permitan todo lo que el método automático trabajó, pero de una manera más directa al usuario ya sea con un enfoque *on premise* o con una aplicación web.

Los diferentes diseños y diagramas que se estipulan para la construcción del método automático, fueron entregados como resultado a estudiantes del semillero SMART que esperan integrar los procesos que se contemplan en un software para la identificación de síntomas depresivos en una segunda etapa de desarrollo.

La implementación y desarrollo de este método automático se trabajaron de la mano con los diseños y necesidades de los profesionales de salud mental establecidos, sin embargo, el objetivo desarrollado en el capítulo cuarto se concentra en la posibilidad de consolidar una descripción del método automático que resulte entendible tanto para los profesionales del área de software como para los profesionales de salud mental y hasta para cualquier persona que se vea relacionada con el método automático, posibilitando el entendimiento y la navegabilidad en los procesos de análisis operacionales que se implementaron en el diseño, planeación y ejecución del mismo. Esto se hace con el fin de facilitar el acceso a la información, su análisis y la generación de informes descriptivos

acerca de la sintomatología depresiva que se encontró en las respuestas de los consultantes.

Para un correcto proceso de transferencia del conocimiento se genera un *Manual de uso para el Método automático en apoyo a la identificación de sintomatología depresiva*. Este manual se encuentra en el anexo 14 y viene acompañado de 22 pasos secuenciales que especifican cuál debe ser el correcto manejo y procedimiento de ejecución para el método automático desarrollado en Google Colaboratory, se representa con la Figura 66. Manual de uso para el Método automático en apoyo a la identificación de sintomatología depresiva.

## Figura 66

### *Manual De Uso Para El Método Automático*

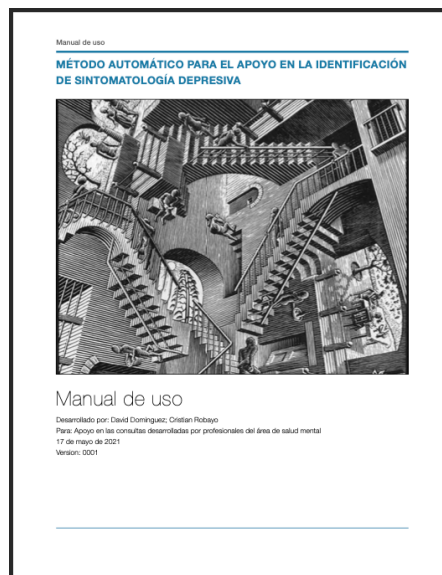


Figura 66. Manual de uso para el Método automático en apoyo a la identificación de sintomatología depresiva.  
Fuente:[El Autor]

*Nota:* En la figura se muestra el manual elaborado para el método automático desarrollado.

Así mismo, con la implementación de la primera versión de este método automático se espera dar inicio a un proceso continuo de generación de Software que permita entender y analizar los sentimientos y emociones de las personas siempre enfocándose en el tratamiento, prevención de enfermedades mentales y en el mejoramiento de la salud mental y de la calidad de vida de las personas.

En la Tabla 15. *Resultados de las Métricas de rendimiento*, se observa un rendimiento promedio del 76.22% en *Precisión*, lo que significa que el modelo

general con todas las métricas de calidad ya inmersas permite clasificar de manera adecuada aproximadamente 8 de 10 consultantes, además, en la métrica *Recall* se tiene un valor de 76.44% significando la cantidad de textos que serán clasificados de manera adecuada sobre un 100% de consultantes. El modelo presenta una *Exactitud* del 94.86% mostrando el porcentaje de textos clasificados de manera correcta y finalmente la métrica *F1-score* con un porcentaje del 76.27% de rendimiento, da a conocer la relación de las métricas de *Precisión* y *Recall*.

Respecto a la relación de la calidad y cantidad del modelo se muestra un valor alto en la identificación de síntomas depresivos en la mayoría de las etiquetas. Es importante destacar que la etiqueta con mejor rendimiento es la de *disminución de interés o placer*, teniendo una precisión del 88% junto con un resultado de la métrica *recall* del 83%, con una *exactitud* del 96.60% y un resultado de *F1 Score* del 85.42%, esto nos da a conocer que aproximadamente de 10 textos que cumplan las características para ser clasificados en esta etiqueta 8 de estos serán clasificados en esta de manera eficaz. Adicionalmente, la etiqueta con el menor rendimiento es la de *atención dispersa* con un 63% de *precisión*, un 61% de la métrica de *recall*, un 92.86% en *exactitud* y, finalmente, un 61.98% en la métrica de desempeño de *F1 Score*. Lo cual significa que de 10 textos 6 serán clasificados de manera correcta en esta etiqueta.

La oportunidad de mejora en todas las etiquetas aquí presentadas puede llevarse a cabo apoyándose de frases prueba y entrenamiento que serán identificadas al momento en el que el método automático se implemente de manera directa en las consultas realizadas, y se tenga un seguimiento a aquellas respuestas más significativas dadas por los consultantes en tiempo real.<sup>170</sup>

**Tabla 15**

*Métricas De Rendimiento Por Etiqueta*

MÉTRICAS DE RENDIMIENTO POR ETIQUETA					
ETIQUETA	TEXTO	PRECISION	RECALL	EXACTITUD	F1-SCORE
Actividad Física	25	68%	76%	94,90%	71,77%
Alimentacion	29	83%	83%	96,60%	83,00%
Atención Dispersa	28	63%	61%	92,86%	61,98%
Cargo de Conciencia	32	83%	78%	95,92%	80,42%
Desaliento	25	72%	72%	95,24%	72,00%

<sup>170</sup> Para más detalles al respecto de cómo se calculan las métricas de desempeño implementadas y su significado teórico en la ciencia de datos, puede dirigirse al [marco referencial](#) de este proyecto de grado.

Descanso	30	72%	77%	94,56%	74,41%
Disminución de interés o placer	35	88%	83%	96,60%	85,42%
Estado de Ánimo Deprimido	52	80%	79%	92,86%	79,49%
Ideas Suicidas	38	77%	79%	94,22%	77,98%
<b>PROMEDIO</b>	<b>33</b>	<b>76,22%</b>	<b>76,44%</b>	<b>94,86%</b>	<b>76,27%</b>

*Tabla 15. Resultados de las Métricas de rendimiento*

*Nota: En la tabla anterior se tienen en cuenta las métricas de rendimiento de todas las etiquetas trabajadas en el proyecto.*



## 7. CONCLUSIONES

En el presente trabajo se logró la construcción de un método automático para la identificación de sintomatología depresiva mediante el uso de narrativas escritas de una persona en respuesta a un formulario de preguntas que se enfoca en identificar la presencia de síntomas depresivos. Esto se realiza implementado un procesamiento de lenguaje natural que se integra con procesos de limpieza y análisis de textos, que son identificados mediante una ontología sistémica que reconoce conceptos de depresión. Todo este proceso se genera y valida por profesionales del área de salud mental e ingenieros de sistemas, para generar informes descriptivos que serán de apoyo en las consultas que se desarrollen.

En la implementación del método automático propuesto se utilizó el entorno de *Google Workspace*, por medio del cual se buscó mayor accesibilidad, seguridad y herramientas que presentan altos estándares de calidad en el análisis de información y datos, tal es el caso de la implementación de los cuadernillos de *Google Colaboratory*. De este modo, los resultados generados en la identificación de sintomatología depresiva van respaldados por métricas de desempeño generadas en *Monkeylearn*, ante las cuales puede concluirse la presencia de etiquetas con un alto valor porcentual de rendimiento (por ejemplo la etiqueta de *disminución de interés o placer* presentó 88% de *precisión*, un resultado de *recall* del 83%, una *exactitud* de 96.60% y un porcentaje de 85.42% de *F1 Score*). Sin embargo, como el modelo fue entrenado en su primera versión se presenta una etiqueta que no llega a cubrir un valor porcentual considerable (por ejemplo la etiqueta de *atención dispersa* con una *precisión* del 63%, con una métrica *recall* de 61%, una *exactitud* del 92.86%, y un porcentaje de *F1 Score* de 61.98%). Con esto, podemos concluir que el modelo puede ser debidamente ajustado con la implementación de nuevas clasificaciones de textos, encaminados a la detección de síntomas depresivos.

De acuerdo con lo señalado en la pregunta de investigación, el objetivo general y los resultados obtenidos de este trabajo de grado, el método automático desarrollado puede ser alineado, no solo con la ontología de sintomatología depresiva, sino que también permite su aplicación a otros tipos de ontologías con dominios y conceptos diferentes, permitiendo así una usabilidad flexible, variable y aplicable a distintas áreas de estudio y análisis de información. De este modo, el método automático al constituirse como la primera versión de una herramienta tecnológica que permita estos análisis en pro de la detección y prevención de la depresión, puede ser la apertura a un camino de desarrollos enfocados en entornos afines.

En este sentido, resulta importante aclarar que, si bien la intención que presentaban las pruebas funcionales que se realizaron para este método

automático era describir un campo de funcionalidad en consultas reales, no se excluyen los procesos de mejoramiento y adecuación de nuestra herramienta a futuros trabajos. Esta posibilidad se ejemplifica en el hecho que muestra que a medida que este método automático se vaya implementando en un mayor número de consultas, los análisis e informes generados serán los parámetros para la reasignación y reformulación de los procesos y objetivos que el método automático puede llegar a alcanzar.

Para garantizar mejores resultados a largo plazo es necesario establecer una retroalimentación continua tanto de la ontología de conceptos de la sintomatología depresiva propuesta como en el trabajo de refinamiento de los procesos de entrenamiento y análisis de información que presenta el método automático. Esto con el fin de garantizar que los resultados obtenidos sean aplicables a todo modelo analítico o descriptivo y no solo en intenciones o entidades de la sintomatología depresiva.

## 8. TRABAJOS FUTUROS

El presente proyecto ha presentado un resultado metodológico que bien puede llamarse *proceso abierto*, con este concepto se explica la posibilidad de ampliar las ideas contenidas dentro del marco de desarrollo y teórico de este proyecto, trabajando colaborativamente con otros estudios e investigaciones que tengan relación con la identificación de depresión, al implementar proyectos de esta área de estudio con nuestro proyecto se permite el trabajo y la mejora continua de la teoría y procedimientos aquí descritos.

En este sentido, el hecho de que este proyecto se constituya como un *proceso abierto* se asocia con algunos de los requerimientos exigidos para los trabajos de grado, no solo por la Universidad Católica de Colombia sino por la academia nacional e internacional. Dentro de estos requerimientos se destaca el lineamiento con respecto a los trabajos futuros, es decir, que cada proyecto debe ser la inspiración de diferentes nuevos trabajos. Además, los trabajos de grado deben inscribirse dentro de una línea de investigación propia de facultad que, en el caso particular de este proyecto, es “Gestión y Tecnología al Servicio de la Sociedad”. Ante esto, se identificará a continuación una lista de posibles trabajos futuros y desarrollos específicos que constituirán una valiosa contribución para el análisis de la emotividad humana y de los trastornos que la aquejan, para así aportar a la mejora de los tratamientos psicológicos de las personas que lo necesiten:

1. Como segunda etapa de este proyecto se plantea desarrollar y consolidar el software, para que siga y reproduzca el método automático descrito en este proyecto de grado.
2. Siguiendo la metodología, ontología y estructuras específicas para el desarrollo de este trabajo, llegar al análisis de textos aplicado a *blogs*, libros *online* e, incluso, a las publicaciones en las distintas redes sociales, para así identificar posibles trastornos o síntomas de depresión que esté padeciendo el usuario.
3. Realizar implementaciones de modelos que se enfoque en la búsqueda de depresión a través de imágenes (o actualmente *memes*), que es uno de los medios de transmisión de información, más populares y aplicados por la sociedad en estos momentos.
4. Trabajar en la implementación de algoritmos que con el apoyo de una ontología o un dataset de información de la persona, se permita hacer entrenamiento por técnicas de deep learning, con el fin de identificar sintomatología depresiva.

## BIBLIOGRAFÍA

- [1] NIMH “Depresión: Información Básica”. {En línea}. {3 de Septiembre de 2020} disponible en: (<https://www.nimh.nih.gov/health/publications/espanol/depresion-sp/index.shtm>)
- [2] MARIA DEL CARMEN BRAVO GARCIA “¿Cuáles son las consecuencias de la depresión?” {En línea}. {3 de Septiembre de 2020} disponible en: (<https://blog.cognifit.com/es/consecuencias-depresion/>)
- [3] PERIODICO EL TIEMPO - SALUD “Depresión en Colombia es más alta que el promedio en el mundo” {En línea}. {5 de Septiembre de 2020} disponible en: (<https://www.eltiempo.com/salud/cifras-sobre-depresion-en-colombia-y-en-el-mundo-segun-la-oms-61454>)
- [4] MINSALUD - COLOMBIA “Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles” {En línea}. {10 de Septiembre de 2020} disponible en: (<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)
- [5] OMS “Campañas mundiales de salud pública de la OMS- Aspectos esenciales de la campaña Día Mundial de la Salud 2017” {En línea}. {10 de Septiembre de 2020} disponible en: (<https://www.who.int/campaigns/world-health-day/2017/campaign-essentials/es>)
- [6] MINSALUD - COLOMBIA - DEPRESION “Depresion #hablemos” {En línea}. {15 de Septiembre de 2020} disponible en: (<https://www.minsalud.gov.co/salud/publica/SMental/Paginas/depresion.aspx>)
- [7] MINSALUD - COLOMBIA “Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles” {En línea}. {15 de Septiembre de 2020} disponible en: (<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)
- [8] Hospital San Juan Capistrano “ Signos y síntomas de la depresión” {En línea}. {20 de Septiembre de 2020} disponible en: (<https://www.sanjuancapistrano.com/depresion/sintomas-efectos/>)
- [9] Pontificia Universidad Javeriana - Cali “ Catálogo de pruebas psicológicas” {En línea}. {20 de Septiembre de 2020} disponible en: ([https://www2.javerianacali.edu.co/sites/ujc/files/catalogo\\_liip\\_2018.pdf](https://www2.javerianacali.edu.co/sites/ujc/files/catalogo_liip_2018.pdf))
- [10] Mariñelarena-Dondena, L., Errecalde, M. L., & Solano, A. C. (2017). Extracción de conocimiento con técnicas de minería de textos aplicadas a la psicología. Revista Argentina de Ciencias del Comportamiento, 9(2), 65-76.
- [11] Palomo, S. R. G., & Gil, E. M. (2020). Aproximación a la ingeniería del software. Editorial Centro de Estudios Ramón Areces SA.
- [12] Margarita M. ÁLVAREZ, Graciela E. BARCHINI, Mercedes DÍAZ, Luis G. CHANFERONII “SISTEMAS DE INFORMACIÓN BASADOS EN ONTOLOGÍAS. UN ÁREA EMERGENTE” {En línea}. {25 de Septiembre de 2020} disponible en: (<https://core.ac.uk/download/pdf/15777408.pdf>)

- [13] González, J. A. M. (2004). El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural. Rev. Esp. Doc. Cient, 27(4), 573.
- [14] Krauskopof, Dina. El desarrollo psicológico en la adolescencia. {En línea}. 1999. {17 septiembre de 2020}. Disponible en ([http://www.scielo.sa.cr/scielo.php?script=sci\\_arttext&pid=S1409-41851999000200004](http://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1409-41851999000200004))
- [15] Ministerio de salud {En línea}. {27 de Octubre de 2020} disponible en: (<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)
- [16] Hector G. Barnes” {En línea}. {27 de Agosto de 2020} disponible en: ([https://www.elconfidencial.com/alma-corazon-vida/2015-08-28/las-razones-por-las-que-cada-vez-cometemos-mas-errores-ortograficos-y-gramaticales\\_988271/](https://www.elconfidencial.com/alma-corazon-vida/2015-08-28/las-razones-por-las-que-cada-vez-cometemos-mas-errores-ortograficos-y-gramaticales_988271/))
- [17] Paul Koeck {En línea} {27 de Agosto de 2020} “¿Test de depresión?¿Que implica un buen test de depresión?” disponible en:(<https://www.mis15minutos.com/depresion/test-de-depresion-que-implica-un-buen-test-de-depresion/>)
- [18] Salud mental y emocional {En línea}. {27 de Agosto de 2020} "Terapia Integral para Depresión y Ansiedad" 2 a.. 2016, disponible en: (<https://www.saludmentalyemocional.com/consulta-particular>)
- [19] Artur samuel (1959)” {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)
- [20] Nick McCrea” {En línea}. {24 de Agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)
- [21] Juan Zambrano - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de Agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)
- [22] Victor Roman - Machine Learning Supervisado: Fundamentos de la Regresión Lineal {En línea}. {28 de Agosto de 2020} disponible en: (<https://medium.com/datos-y-ciencia/machine-learning-supervisado-fundamentos-de-la-regresi%C3%B3n-lineal-bbcb07fe7fd>)
- [23] Joaquin Amat Rodrigo - Regresión logística simple y múltiple {En línea}. {28 de Agosto de 2020} disponible en: ([https://www.cienciadedatos.net/documentos/27\\_regresion\\_logistica\\_simple\\_y\\_multiple](https://www.cienciadedatos.net/documentos/27_regresion_logistica_simple_y_multiple))
- [24] Joaquín Amat Rodrigo - Regresión logística simple y múltiple {En línea}. {28 de Agosto de 2020}: ([https://www.cienciadedatos.net/documentos/27\\_regresion\\_logistica\\_simple\\_y\\_multiple](https://www.cienciadedatos.net/documentos/27_regresion_logistica_simple_y_multiple))
- [25] Sitio Big Data - {Clasificación de Aprendizaje automático supervisado (En línea). {24 de Agosto de 2020} disponible en: (<https://sitiobigdata.com/2019/12/24/clasificacion-de-aprendizaje-automatico-supervisado/>)

- [26] Juan Zambrano - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de Agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)
- [27] Clasificación de Aprendizaje automático supervisado {En línea}. {29 de Agosto de 2020} disponible en: (<https://sitiobigdata.com/2019/12/24/clasificacion-de-aprendizaje-automatico-supervisado/>)
- [28] Shubham Panchal - k Nearest Neighbor Classifier ( kNN ) - Machine Learning Algorithms {En línea}. {29 de Agosto de 2020}: (<https://medium.com/@equipintelligence/k-nearest-neighbor-classifier-knn-machine-learning-algorithms-ed62feb86582>)
- [29] Shubham Panchal - k Nearest Neighbor Classifier ( kNN ) - Machine Learning Algorithms {En línea}. {29 de Agosto de 2020} disponible en: (<https://medium.com/@equipintelligence/k-nearest-neighbor-classifier-knn-machine-learning-algorithms-ed62feb86582>)
- [30] Esacademic - Hiperplano {En línea}. {29 de Agosto de 2020} disponible en: (<https://esacademic.com/dic.nsf/eswiki/574819>)
- [31] Jose Martinez Heras - Máquinas de Vectores de Soporte (SVM) {En línea}. {29 de Agosto de 2020} disponible en: (<https://www.iaartificial.net/maquinas-de-vectores-de-soporte-svm/>)
- [32] Java T Point - Support Vector Machine Algorithm {En línea}. {29 de Agosto de 2020}: (<https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm>)
- [33] Sitio Big Data - Árbol de decisión en Machine Learning (Parte 1) {En línea}. {29 de Agosto de 2020} disponible en: (<https://sitiobigdata.com/2019/12/14/arbore-de-decision-en-machine-learning-parte-1/>)
- [34] Prashant Gupta - Decision Trees in Machine Learning {En línea}. {29 de Agosto de 2020}: (<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)
- [35] Prashant Gupta - Decision Trees in Machine Learning {En línea}. {29 de Agosto de 2020} disponible en: (<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)
- [36] ABHISHEK SHARMA, - Decision Tree vs. Random Forest – Which Algorithm Should you Use (Mayo de 2020) {En línea}. {10 de enero de 2021}: (<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)
- [37] Juan Zambrano - Introducción a la Teoría de Aprendizaje de Máquina y sus Aplicaciones: Un Tutorial Visual con Ejemplos {En línea}. {24 de Agosto de 2020} disponible en: (<https://www.toptal.com/machine-learning/introduccion-a-la-teoria-de-aprendizaje-de-maquina-y-sus-aplicaciones-un-tutorial-visual-con-ejemplos>)
- [38] Universidad de Oviedo - El algoritmo k-means aplicado {En línea}. {29 de Agosto de 2020} disponible en: ([https://www.unioviedo.es/compnum/laboratorios\\_py/kmeans/kmeans.html](https://www.unioviedo.es/compnum/laboratorios_py/kmeans/kmeans.html))
- [39] Prashant Gupta - Decision Trees in Machine Learning {En línea}. {29 de Agosto de

2020):(<https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052>)

[40] Goica21 , Monografias.com. (s. f.). Redes neuronales (página 2) - Monografias.com. monografías. <https://www.monografias.com/trabajos12/redneuro/redneuro2.shtml>

[41] Atria Innovation - (22 de Octubre 2019)" {En línea}. {10 de enero de 2021} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

[42] Atria Innovation - (22 de Octubre 2019)" {En línea}. {10 de enero de 2021} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

[43] Juan ignacio barrios - Everything you Should Know about Confusion Matrix for Machine Learning La matriz de confusión y sus métricas (2019)" {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

[44] Aniruddha Bhandari - Everything you Should Know about Confusion Matrix for Machine Learning {En línea}. {29 de Agosto de 2020}:(<https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/>)

[45] Juan ignacio barrios - La matriz de confusión y sus métricas (2019)" {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

[46] Juan ignacio barrios - La matriz de confusión y sus métricas (2019)" {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>)

[47] Jose martinez Heras - precisión, Recall, F1, Accuracy en clasificación {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.iartificial.net/precisión-recall-f1-accuracy-en-clasificacion/>)

[48] Xue Y., Li Q., Jin L., Feng L., Clifton D.A., Clifford G.D. (2014) Detecting Adolescent Psychological Pressures from Micro-Blog. En: Zhang Y., Yao G., He J., Wang L., Smalheiser N.R., Yin X. (eds) Health Information Science. HIS 2014. Lecture Notes in Computer Science, vol 8423. Springer, Cham. {18 de Agosto de 2020} disponible en: [https://link.springer.com/chapter/10.1007/978-3-319-06269-3\\_10](https://link.springer.com/chapter/10.1007/978-3-319-06269-3_10)

[49] Farruque N., Zaiane O., Goebel R. (2019) Augmenting Semantic Representation of Depressive Language: From Forums to Microblogs. En: Brefeld U., Fromont E., Hotho A., Knobbe A., Maathuis M., Robardet C. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2019. {18 de Agosto de 2020} disponible en: <https://sci-hub.tw/10.1007/978-3-030-46133-1>

[50] Developers Google - Clasificación: Exactitud {En línea}. {29 de Agosto de 2020} disponible en:(<https://developers.google.com/machine-learning/crash-course/classification/accuracy?hl=es-419>)

[51] Maria Manuela Valenzuela Sosa - Herramienta Protégé para la Web Semántica {En línea}. {19 de Agosto de 2020} disponible en: (<https://es.slideshare.net/Mvalenzuelasosa/herramienta-protg-para-la-web-semntica>)

[52] Universidad de Stanford - Protégé {En línea}. {25 de Agosto de 2020} disponible en: (<https://Protégé.stanford.edu/about.php>)

- [53] Brandwatch - Entendiendo el análisis de sentimiento: qué es y para qué se usa {En línea}. {24 de Agosto de 2020} disponible en: (<https://www.brandwatch.com/es/blog/analisis-de-sentimiento/>)
- [54] Manuel Delgado - ¿Qué es el análisis del sentimiento? {En línea}. {24 de Agosto de 2020} disponible en: (<https://manueldelgado.com/que-es-el-analisis-del-sentimiento/>)
- [55] Antonio Feregrino Bolaños - Introducción al análisis de textos{En línea}. {10 de enero de 2021} disponible en: (<https://tacosdedatos.com/analisis-texto>)
- [56] Brandwatch - Entendiendo el análisis de sentimiento: qué es y para qué se usa {En línea}. {24 de Agosto de 2020} disponible en: (<https://www.brandwatch.com/es/blog/analisis-de-sentimiento/>)
- [57] Elena Álvarez Mellado - La lematización de textos{En línea}. {10 de Enero de 2021} disponible en: (<https://es.slideshare.net/lirondos/la-lematizacion-de-textos>)
- [58] Elena Álvarez Mellado - La lematización de textos{En línea}. {10 de Enero de 2021} disponible en: (<https://es.slideshare.net/lirondos/la-lematizacion-de-textos>)
- [59] Elena Álvarez Mellado - La lematización de textos{En línea}. {10 de Enero de 2021} disponible en: (<https://es.slideshare.net/lirondos/la-lematizacion-de-textos>)
- [60] Natalya F. Noy and Deborah L. McGuinness “Ontology Development 101: A Guide to Creating Your First Ontology”. {En línea}. {01 Octubre de 2020} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).
- [61] Jose Martinez Heras - precisión, Recall, F1, Accuracy en clasificación {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.iartificial.net/precisión-recall-f1-accuracy-en-clasificacion/>)
- [62] Datahack - INTRODUCCIÓN A GOOGLE COLAB PARA DATA SCIENCE {En línea}. {25 de Agosto de 2020} disponible en: (<https://www.datahack.es/google-colab-para-data-science/#:~:text=Colab%20es%20un%20servicio%20cloud.disponible%20para%20R%20y%20Scala.>)
- [63] SAS - What is natural language processing nlp {En línea}. {18 de Agosto de 2020} disponible en:([https://www.sas.com/es\\_co/insights/analytics/what-is-natural-language-processing-nlp.htmlhttps://www.sas.com/es\\_co/insights/analytics/what-is-natural-language-processing-nlp.html](https://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.htmlhttps://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.html))
- [64] Organización Mundial De la Salud - Depresión {En línea}. {18 de Agosto de 2020} disponible en: (<https://www.who.int/es/news-room/fact-sheets/detail/depression>)
- [65] Pablo Turmero - Métodos automáticos de análisis {En línea}. {19 de Agosto de 2020} disponible en: (<https://www.monografias.com/trabajos109/metodos-automaticos-analisis/metodos-automaticos-analisis.shtml>)
- [66] Manuel Delgado Tenorio - ¿Qué es el análisis del sentimiento? {En línea}. {19 de Agosto de 2020} disponible en: (<https://manueldelgado.com/que-es-el-analisis-del-sentimiento/>)
- [67] García Higuera - Trastornos psicológicos {En línea}. {25 de Agosto de 2020} disponible en: (<http://www.psicoterapeutas.com/Trastornos.html>)



- [68] Natalya F. Noy and Deborah L. McGuinness “Ontology Development 101: A Guide to Creating Your First Ontology”. [En línea]. {01 Octubre de 2020} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).
- [69] Alba Sanz - Procesamiento del Lenguaje Natural (PLN o NLP): qué es y para qué se utiliza [En línea]. {25 de Agosto de 2020} disponible en: (<https://decidesoluciones.es/procesamiento-del-lenguaje-natural-pln-o-nlp-que-es-y-para-que-se-utiliza/>)
- [70] Miguel Angel Alvarez - Qué es Python [En línea]. {25 de Agosto de 2020} disponible en: (<https://desarrolloweb.com/articulos/1325.php>)
- [71] Mintic - Colombia aprende: ¿QUÉ ES UN ALGORITMO? [En línea]. {25 de Agosto de 2020} disponible en: ([http://aprende.colombiaprende.edu.co/sites/default/files/naspublic/curriculos\\_ex/n1g10\\_fproy/nive11/programacion/unidad1/leccion1.html](http://aprende.colombiaprende.edu.co/sites/default/files/naspublic/curriculos_ex/n1g10_fproy/nive11/programacion/unidad1/leccion1.html))
- [72] Rafael Gonzalez - R para visualización [En línea]. {25 de Agosto de 2020} disponible en: (<https://gonzalezgouveia.com/que-es-r-y-rstudio/>)
- [73] UniPython [En línea]. {10 de Enero de 2021} disponible en: (<https://uniPython.com/scipy-funciones-principales/>)
- [74] Requests. [En línea]. {16 de Enero de 2021} PyPI. [En línea] disponible en: <https://pypi.org/project/requests/>
- [75] Beautiful Soup Documentation — Beautiful Soup 4.9.0 documentation. (s. f.). Beautiful Soup Documentation. [En línea] {18 de Enero de 2021} disponible en: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- [76] deplacy. [En línea] {20 de Enero de 2021}. PyPI. [En línea] disponible en: <https://pypi.org/project/deplacy/>
- [77] Spanish · spaCy Models Documentation. (s. f.). Spanish. [En línea] {20 de Enero de 2021} disponible en: <https://spacy.io/models/es/>
- [78] Package Discovery and Resource Access using pkg\_resources — setuptools 51.1.2 documentation. (s. f.). Package Discovery and Resource Access using pkg\_resources. [En línea] {22 de Enero de 2021} disponible en: [https://setuptools.readthedocs.io/en/latest/pkg\\_resources.html](https://setuptools.readthedocs.io/en/latest/pkg_resources.html)
- [79] imp — Access the import internals — Python 3.9.1 documentation. (s. f.). imp. [En línea] {22 de Enero de 2021} disponible en: <https://docs.Python.org/3/library/imp.html>
- [80] Python Examples of Impreload. (s. f.). Python Examples of imp.reload. [En línea] {22 de Enero de 2021} disponible en: <https://www.programcreek.com/Python/example/6950/imp.reload>
- [81] Hellmann, D. (s. f.). Counter - Python Module of the Week. Counter - Python Module of the Week. [En línea] {24 de Enero de 2021} disponible en: <https://pymotw.com/2/collections/counter.html>

- [82] Masum, M.(2020) How to import and Export Data sets Colab - Towards Data Science. Medium [En línea] {24 de Enero de 2021} disponible en: <https://towardsdatascience.com/google-colab-import-and-export-datasets-eccf801e2971>
- [83] Openpyxl - A Python library to read/write Excel 2010 xlsx/xlsm files — openpyxl 3.0.5 documentation. (s. f.). openpyxl A Python library to read/write Excel 2010 xlsx/xlsm files. [En línea] {26 de Enero de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/>
- [84] GeeksforGeeks. (2019, 18 junio). Python | os.read() method. [En línea] {26 de Enero de 2021} disponible en: [https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read\(\)%20method%20in%20Python.bytes%20left%20to%20be%20read.](https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read()%20method%20in%20Python.bytes%20left%20to%20be%20read.)
- [85] Python String split() Method. (s. f.). Python String split() Method. [En línea] {26 de Enero de 2021} disponible en: [https://www.w3schools.com/Python/ref\\_string\\_split.asp](https://www.w3schools.com/Python/ref_string_split.asp)
- [86] Simple usage — openpyxl 3.0.5 documentation. (s. f.). Simple usage — openpyxl 3.0.5 documentation. [En línea] {01 de Febrero de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/usage.html>
- [87] Openpyxl 3.0.5 documentation. (s. f.). iter\_rows openpyxl.worksheet.worksheet module. [En línea] {01 de Febrero de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/api/openpyxl.worksheet.worksheet.html>
- [88] Extracting Information from Text. (s. f.). Extracting Information from Text. [En línea] {01 de Febrero de 2021} disponible en: <https://www.nltk.org/book/ch07.html>
- [89] Code Faster with Line-of-Code Completions, Cloudless Processing. (s. f.). [En línea] {01 de Febrero de 2021} Code Faster with Line-of-Code Completions. disponible en: [https://www.kite.com/Python/docs/collections.Counter.most\\_common](https://www.kite.com/Python/docs/collections.Counter.most_common)
- [90] Python Language - Iterando sobre los partidos usando `re.finditer` | Python Tutorial. (s. f.). Python Language - Iterando sobre los partidos usando `re.finditer`. [En línea] {01 de Febrero de 2021} disponible en: <https://riptutorial.com/es/Python/example/12097/iterando-sobre-los-partidos-usando--re-finditer>
- [91] Ministerio de salud, del gobierno de Colombia. "Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles" [En línea] {03 de Febrero de 2021} disponible en: (<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-depresion-marzo-2017.pdf>)
- [92] Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. "Self-compassion and Symptoms of Stress, Anxiety, and Depression" [En línea] {03 de Febrero de 2021} disponible en: ([\)](https://www.scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85082776505&origin=resultslist&sort=plf-f&src=s&st1=sintomas+depresi%c3%b3n&st2=&sid=60b9ad8431f907b88a1e1093beb09ff8&sot=b&sdt=b&sl=33&s=TITLE-ABS-KEY%28sintomas+depresi%c3%b3n%29&relpos=2&citeCnt=1&searchTerm=)
- [93] Chaiwuttisak, P. Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand. "Text Mining Analysis of Comments in Thai Language for Depression from Online Social Networks " [En línea] {05 de Febrero de 2021}

disponible en:  
(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85088469231&origin=resultslst&sort=plf-f&src=s&st1=Depression+and+Software&st2=&sid=10a2d25e36c518b4f31ef2fda413dbc6&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Depression+and+Software%29&relpos=1&citeCnt=0&searchTerm=>)

[94] Trifan, A. , Antunes, R. Oliveira, J.L. DETI/IEETA, University of Aveiro, Aveiro, Portugal {En línea}. "Machine Learning for Depression Screening in Online Communities " [En línea] {05 de Febrero de 2021}  
(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85089220717&origin=resultslst&sort=plf-f&src=s&st1=Language+for+Depression&st2=&sid=4507d7f667902348aa8d12e0318915fd&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Language+for+Depressio n%29&relpos=1&citeCnt=0&searchTerm=>)

[95] Ding, J.a,b, Chen, K.c, Liu, H.d, Huang, L.e, Chen, Y.a,f, Lv, Y.g, Yang, Q.h, Guo, Q.e , Han, Z.a , Lambon Ralph, M.A.i University, Beijing, China, Department of Neurosurgery, Baylor College of Medicine, Houston, TX, United States, Fudan University, Shanghai, China. "A unified neurocognitive model of semantics language social behaviour and face recognition in semantic dementia" [En línea] {05 de Febrero de 2021}  
(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85085265414&origin=resultslst&sort=plf-f&src=s&st1=Language+for+Depression&st2=&sid=4507d7f667902348aa8d12e0318915fd&sot=b&sdt=b&sl=38&s=TITLE-ABS-KEY%28Language+for+Depressio n%29&relpos=4&citeCnt=0&searchTerm=>)

[96] Maria Li Zou, Mandy Xiaoyang Li, Vincent Cho. "Depression and disclosure behavior via social media: A study of university students in China" [En línea] {05 de Febrero de 2021}  
(<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S2405844020302139>)

[97] Ingrid Nunes, Simon Miles, Michael Luck, Simone Barbosa, Carlos Lucena. "Decision making with natural language based preferences and psychology-inspired heuristics" [En línea] {05 de Febrero de 2021}  
(<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/abs/pii/S0952197615000603>)

[98] S.R.Goldman, M.B.W.Wolfe - International Encyclopedia of the Social & Behavioral Sciences {En línea}. "Text Comprehension: Models in Psychology" {10 de Febrero de 2021} disponible en:  
(<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/B0080430767015424>)

[99] JJ Bradac, S.A. Reid - Encyclopedia of Language & Linguistics (Second Edition) {En línea}. "Social Psychology and Language" {10 de Febrero de 2021} disponible en:  
(<https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/B0080448542013079>)

[100] Stankevich, M. , Isakov, V., Devyatkin, D. , Smirnov, - Institute for Systems Analysis, Federal Research Center "Computer Science and Control" of RAS, Moscow, Russian Federation {En línea}. " Feature engineering for depression detection in social media " {10 de Febrero de 2021} disponible en:  
(<https://www-scopus-com.ucatolica.basesdedatosezproxy.com/record/display.uri?eid=2-s2.0-85052014689&origin=resultslst&sort=plf-f&src=s&st1=depression+AND+machine+learning&st2=&sid=9fab>)

[b56fd23cdd630985120d318a1db8&sot=b&sdt=b&sl=45&s=TITLE-ABS-KEY%28depression+AND+machine+learning%29&relpos=0&citeCnt=8&searchTerm=\)](https://www.researchgate.net/publication/350812031/b56fd23cdd630985120d318a1db8&sot=b&sdt=b&sl=45&s=TITLE-ABS-KEY%28depression+AND+machine+learning%29&relpos=0&citeCnt=8&searchTerm=)

[101] Vanessa Rosselló Villán “Las metodologías ágiles más utilizadas y sus ventajas dentro de la empresa”. {En línea}. {10 de Febrero de 2021} disponible en: (<https://www.iebschool.com/blog/que-son-metodologias-agiles-agile-scrum/>).

[102] Gráfica obtenida de: 2020 © AIN Asociación de la Industria Navarra “Transformación Cultural Ágil y Responsable (TCAR)” {En línea}. {13 de Febrero de 2021} disponible en: <https://www.ain.es/archivo-proyectos/metodologia-agile-tcar/>.

[103] Natalya F. Noy and Deborah L. McGuinness “Ontology Development 101: A Guide to Creating Your First Ontology”. {En línea}. {15 de Febrero de 2021} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).

[104] Natalya F. Noy and Deborah L. McGuinness “Ontology Development 101: A Guide to Creating Your First Ontology”. {En línea}. {05 de Marzo de 2021} disponible en: ([https://Protégé.stanford.edu/publications/ontology\\_development/ontology101.pdf](https://Protégé.stanford.edu/publications/ontology_development/ontology101.pdf)).

[105] Ontology 101 - Protégé Wiki. (s. f.). Protégé Wiki.Stanford.Edu. {En línea}. {05 de Marzo de 2021} disponible en: <https://Protégéwiki.stanford.edu/wiki/Ontology101>

[106] Colombia, S. P. S. D. M. Y. (s. f.). Páginas - Ciclo de Vida. [minsalud.gov.co/](https://www.minsalud.gov.co/). {En línea}. {05 de Marzo de 2021} disponible en: [https://www.minsalud.gov.co/proteccionsocial/Paginas/cicloVida.aspx#:~:text=6%20%2D%2011%20a%C3%B1os\)-,Adolescencia%20\(12%20%2D%2018%20a%C3%B1os\),o%20mas\)%20envejecimiento%20y%20vejez](https://www.minsalud.gov.co/proteccionsocial/Paginas/cicloVida.aspx#:~:text=6%20%2D%2011%20a%C3%B1os)-,Adolescencia%20(12%20%2D%2018%20a%C3%B1os),o%20mas)%20envejecimiento%20y%20vejez)

[107] Psi.uba.ar. 2021. {En línea}. {15 de Marzo de 2021} disponible en: ([https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios\\_catedras/obligatorias/070\\_psicoterapias1/material/inventario\\_beck.pdf](https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios_catedras/obligatorias/070_psicoterapias1/material/inventario_beck.pdf))

[108] Escala de Autoevaluación para la Depresión de Zung (SDS). [En línea] Disponible en: [Mentalhealthministries.net website:](http://www.mentalhealthministries.net/website/) {En línea}. {15 de Marzo de 2021} disponible en: [http://www.mentalhealthministries.net/resources/flyers/zung\\_scale/zung\\_scale\\_sp.pdf](http://www.mentalhealthministries.net/resources/flyers/zung_scale/zung_scale_sp.pdf)

[109] di AGnósticos, C. (s/f). Guía de consulta.Disponible en:Edu.co website: {En línea}. {15 de Marzo de 2021} disponible en: <https://www.eafit.edu.co/ninos/reddelaspreguntas/Documents/dsm-v-guia-consulta-manual-diagnostico-estadistico-trastornos-mentales.pdf>

[110] Diagnósticos, C.-10-E. (s/f). MANUAL DE CODIFICACIÓN. [En línea] Disponible en Gob.es website: {En línea}. {15 de Marzo de 2021} disponible en: [https://www.mscbs.gob.es/estadEstudios/estadisticas/normalizacion/CIE10/UT\\_MANUAL\\_DIAG\\_2016\\_prov1.pdf](https://www.mscbs.gob.es/estadEstudios/estadisticas/normalizacion/CIE10/UT_MANUAL_DIAG_2016_prov1.pdf)

[111] Bases de Datos. (s. f.). U Católica Bases de Datos. {En línea}. {20 de Marzo de 2021} disponible en: [https://login.ucatolica.basesdedatosezproxy.com/public/menu\\_ext.htm#cat%C3%A1logo](https://login.ucatolica.basesdedatosezproxy.com/public/menu_ext.htm#cat%C3%A1logo)

[112] Jung, H., Park, H. A., & Song, T. M. (2017). Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals. *Journal of medical internet research*, 19(7), e259. [Consultado el 20 de Marzo de 2021].

[113] Jung, H., Park, H. A., & Song, T. M. (2016). Development and Evaluation of an Adolescents' Depression Ontology for Analyzing Social Data. *Studies in health technology and informatics*, 225, 442-446. [Consultado el 20 de Marzo de 2021].

[114] Larsen, R. R., & Hastings, J. (2018). From affective science to psychiatric disorder: Ontology as a semantic bridge. *Frontiers in psychiatry*, 9, 487. [Consultado el 20 de Marzo de 2021].

[115] Chang, Y. S., Fan, C. T., Lo, W. T., Hung, W. C., & Yuan, S. M. (2015). Mobile cloud-based depression diagnosis using an ontology and a Bayesian network. *Future Generation Computer Systems*, 43, 87-98. [Consultado el 20 de Marzo de 2021].

[116] Hu, B., Hu, B., Wan, J., Dennis, M., Chen, H. H., Li, L., & Zhou, Q. (2010). Ontology-based ubiquitous monitoring and treatment against depression. *Wireless communications and mobile computing*, 10(10), 1303-1319. [Consultado el 20 de Marzo de 2021].

[117] Yamada, D. B., Bernardi, F. A., Miyoshi, N. S. B., de Lima, I. B., Vinci, A. L. T., Yoshiura, V. T., & Alves, D. (2020, June). Ontology-Based Inference for Supporting Clinical Decisions in Mental Health. In *International Conference on Computational Science* (pp. 363-375). Springer, Cham. [Consultado el 20 de Marzo de 2021].

[118] "(En informática, un complemento o plug-in es una aplicación que se relaciona con otra para agregarle una función nueva y generalmente muy específica.)"Definición de plugin — Definición.de. (s. f.). Definición.de. {En línea}. {15 de Marzo de 2021} disponible en: <https://definicion.de/plugin/>

[119] (S/f). Recuperado el 20 de Marzo de 2021, de Stanford.edu website: {En línea}. {20 de Marzo de 2021} disponible en: <https://Protégéwiki.stanford.edu/wiki/ProtégéReasonerPlugin>

[120] Vista General del Lenguaje de Ontologías Web (OWL). (s. f.). w3.org {En línea}. {20 de Marzo de 2021} disponible en: <https://www.w3.org/2007/09/OWL-Overview-es.html>

[121] ¿Qué son los Logs y por qué deben interesarte? Recuperado el 12 de abril de 2021, de Dbibyhavas.io website: {En línea}. {20 de Marzo de 2021} disponible en: <https://dbibyhavas.io/es/blog/que-son-los-logs/>

[122] OntoDebug - Protégé Wiki. (s. f.). Protégéwiki.Stanford.Edu. {En línea}. {20 de Marzo de 2021} disponible en: <https://Protégéwiki.stanford.edu/wiki/OntoDebug>

[123] Marciszack, M. M. (s. f.). Construcción de una ontología para gramáticas formales y máquinas abstractas utilizando Protégé para la elicitación de requerimientos. sedici.unlp.edu.ar. {En línea}. {20 de Marzo de 2021} disponible en: <http://sedici.unlp.edu.ar/handle/10915/19743>

**Documento:**

Marciszack, M., Pérez Cota, M., Antonelli, R., Cardenas, M. and Giandini, R., 2009. *Construcción De Una ontología Para Gramáticas Formales Y Máquinas Abstractas Utilizando Protégé Para La Elicitación De Requerimientos*. [ebook] Argentina: SEDICI, pp.296-300. Disponible en: <[https://www.profesores.frc.utn.edu.ar/sistemas/ssl/marciszack/Documentos/ontologías\\_CNIT2009.pdf](https://www.profesores.frc.utn.edu.ar/sistemas/ssl/marciszack/Documentos/ontologías_CNIT2009.pdf)> [Consultado en 20 Marzo 2021].

- [124] Ramos, E. (s. f.). Esquema para evaluar ontologías únicas para un dominio de conocimiento. scielo.org. {En línea}. {20 de Marzo de 2021} disponible en: [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1690-75152009000100005](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1690-75152009000100005)
- [125] spaCy · Industrial-strength Natural Language Processing in Python. (2021). spacy. [En línea] disponible en: {En línea}. {20 de Marzo de 2021} disponible en: <https://spacy.io/>
- [126] Requests. {En línea}. {25 de Marzo de 2021} disponible en: disponible en: <https://pypi.org/project/requests/>
- [127] Beautiful Soup Documentation — Beautiful Soup 4.9.0 documentation. (s. f.). Beautiful Soup Documentation. {En línea}. {25 de Marzo de 2021} disponible en: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- [128] deplacy. PyPI. {En línea}. {25 de Marzo de 2021} disponible en: <https://pypi.org/project/deplacy/>
- [129] Spanish · spaCy Models Documentation. (s. f.). Spanish. {En línea}. {25 de Marzo de 2021} disponible en: <https://spacy.io/models/es/>
- [130] Package Discovery and Resource Access using pkg\_resources — setuptools 51.1.2 documentation. (s. f.). Package Discovery and Resource Access using pkg\_resources. {En línea}. {25 de Marzo de 2021} disponible en: [https://setuptools.readthedocs.io/en/latest/pkg\\_resources.html](https://setuptools.readthedocs.io/en/latest/pkg_resources.html)
- [131] imp — Access the import internals — Python 3.9.1 documentation. (s. f.). imp. {En línea}. {25 de Marzo de 2021} disponible en: <https://docs.python.org/3/library/imp.html>
- [132] Python Examples of imp.reload. (s. f.). Python Examples of imp.reload. {En línea}. {25 de Marzo de 2021} disponible en: <https://www.programcreek.com/Python/example/6950/imp.reload>
- [133] Hellmann, D. (s. f.). Counter - Python Module of the Week. Counter - Python Module of the Week. {En línea} {01 de Abril de 2021} disponible en: <https://pymotw.com/2/collections/counter.html>
- [134] Masum, M. (2020, 27 octubre). How to Import and Export Datasets in Google Colab - Towards Data Science. Medium. {En línea} {01 de Abril de 2021} disponible en: <https://towardsdatascience.com/google-colab-import-and-export-datasets-eccf801e2971>
- [135] openpyxl - A Python library to read/write Excel 2010 xlsx/xlsm files — openpyxl 3.0.5 documentation. (s. f.). openpyxl A Python library to read/write Excel 2010 xlsx/xlsm files. {En línea} {01 de Abril de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/>
- [136] GeeksforGeeks. (2019, 18 junio). Python | os.read() method. {En línea} {01 de Abril de 2021} disponible en: [https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read\(\)%20method%20in%20Python.bytes%20left%20to%20be%20read.](https://www.geeksforgeeks.org/Python-os-read-method/#:%7E:text=read()%20method%20in%20Python.bytes%20left%20to%20be%20read.)
- [137] Python String split() Method. (s. f.). Python String split() Method. {En línea} {05 de Abril de 2021} disponible en: [https://www.w3schools.com/Python/ref\\_string\\_split.asp](https://www.w3schools.com/Python/ref_string_split.asp)



- [138] Simple usage — openpyxl 3.0.5 documentation. (s. f.). Simple usage — openpyxl 3.0.5 documentation. {En línea} {05 de Abril de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/usage.html>
- [139] openpyxl.worksheet.worksheet module — openpyxl 3.0.5 documentation. (s. f.). iter\_rows openpyxl.worksheet.worksheet module. {En línea} {05 de Abril de 2021} disponible en: <https://openpyxl.readthedocs.io/en/stable/api/openpyxl.worksheet.worksheet.html>
- [140] 7. Extracting Information from Text. (s. f.). Extracting Information from Text. {En línea} {05 de Abril de 2021} disponible en: <https://www.nltk.org/book/ch07.html>
- [141] Code Faster with Line-of-Code Completions, Cloudless Processing. (s. f.). Code Faster with Line-of-Code Completions. {En línea} {15 de Abril de 2021} disponible en: [https://www.kite.com/Python/docs/collections.Counter.most\\_common](https://www.kite.com/Python/docs/collections.Counter.most_common)
- [142] Python Language - Iterando sobre los partidos usando `re.finditer` | Python Tutorial. (s. f.). Python Language - Iterando sobre los partidos usando `re.finditer`. {En línea} {15 de Abril de 2021} disponible en: <https://riptutorial.com/es/Python/example/12097/iterando-sobre-los-partidos-usando--re-finditer->
- [143] N. (2019, 6 abril). Procesamiento del Lenguaje Natural (NLP). Aprende Machine Learning. {En línea} {15 de Abril de 2021} disponible en: <https://www.aprendemachinelearning.com/procesamiento-del-lenguaje-natural-nlp/>
- [144] J., JayJay, V. A. P. B. (2018, 16 abril). Limpieza de texto utilizando Python. Data Science. [En línea] disponible en: {En línea} {21 de Abril de 2021} disponible en: <http://datascience.esy.es/limpieza-de-texto-utilizando-Python/>
- [145] Captura de Información – Tandem. (2018, 10 julio). tandem. {En línea} {21 de Abril de 2021} disponible en: <https://www.tandemweb.com/tandem/tag/captura-de-informacion/#:%7E:text=Eficacia%20y%20eficiencia%20en%20la.datos%20y%20aplicaciones%20de%20terceros.>
- [146] X.D. (2019 - 2020). Desarrollo de ontologías. CS-FIB-UPC cbea. {En línea} {21 de Abril de 2021} disponible en: <https://www.cs.upc.edu/~bejar/ecsdi/Teoria/ECSDI05b-Desarrollo.pdf>
- [147] O.M. (2020, 27 agosto). Paso a paso para definir métricas de desempeño - Obed Medina. Medium {En línea} {21 de Abril de 2021} disponible en: <https://obedm.medium.com/paso-a-paso-para-definir-m%C3%A9tricas-de-desempe%C3%B1o-88a4207b5042#:~:text=Tener%20objetivos%20te%20permite%20enfocar.en%20el%20desempe%C3%B1o%20del%20equipo.>
- [148] Ofima s.a.s, “Ejemplo\_Formato\_Requerimiento\_SoftwareVersion2.0.doc” {En línea} {21 de Abril de 2021} disponible en: [https://www.ofima.com/wp-content/uploads/2018/05/Ejemplo\\_Formato\\_Requerimiento\\_SoftwareVersion2.0.doc](https://www.ofima.com/wp-content/uploads/2018/05/Ejemplo_Formato_Requerimiento_SoftwareVersion2.0.doc)
- [149] (S/f-b). Recuperado el 12 de abril de 2021, de Gov.co website: {En línea} {25 de Abril de 2021} disponible en: <https://www.unidadvictimas.gov.co/sites/default/files/documentosbiblioteca/formatodeespecificacionderequerimientosdesoftwarev1.doc>

[150] Pierre Bourque, Richard E. (Dick) Fairley, "SWEBOK Guide to the Software Engineering Body of Knowledge v3.0" {En línea} {25 de Abril de 2021} disponible en: <https://ieeecs-media.computer.org/media/education/swebok/swebok-v3.pdf>

[151] Qué es el lenguaje unificado de modelado (UML). (s. f.). Lucidchart. {En línea} {25 de Abril de 2021} disponible en: <https://www.lucidchart.com/pages/es/que-es-el-lenguaje-unificado-de-modelado-uml>

[152] (S/f-c). Recuperado el 25 de abril de 2021, de Edu.mx website:{En línea} {25 de Abril de 2021} disponible en: [https://www.uaeh.edu.mx/docencia/P\\_Presentaciones/tizayuca/ing\\_computacion/Fundamentos%20de%20Programacion.pdf](https://www.uaeh.edu.mx/docencia/P_Presentaciones/tizayuca/ing_computacion/Fundamentos%20de%20Programacion.pdf)

[153] Universal POS tags. (s. f.). Universaldependencies.Org. {En línea} {25 de Abril de 2021} disponible en: <https://universaldependencies.org/docs/u/pos/>

[154] MonkeyLearn API Reference. (s. f.). monkeylearn.{En línea} {27 de Abril de 2021} disponible en: <https://monkeylearn.com/api/v3/>





**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 1**  
**Consentimiento Informado**

**ANDRÉS DAVID DOMÍNGUEZ ROZO**  
**CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA**  
**FACULTAD DE INGENIERÍA**  
**PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**  
**BOGOTÁ**  
**2021**

## Anexo 1. Consentimiento Informado

### CONSENTIMIENTO INFORMADO

Yo \_\_\_\_\_ mayor de edad, identificado con C.C. \_\_\_\_\_ de \_\_\_\_\_, expreso de manera libre y voluntaria que **ACEPTO** participar en la investigación que tiene como diseñar y validar un software de apoyo para la detección de sintomatología emocional en jóvenes entre los 18 y 24 años, a través de una ontología semántica y algoritmos de aprendizaje de máquina en narrativas textuales y publicación en redes sociales por medio de medidas subjetivas y objetivas de malestar psicológico. Esta investigación es realizada por \_\_\_\_\_ docentes del programa de la Facultad de Ingeniería y Psicología de Universidad Católica de Colombia.

Declaro que la explicación que recibí, acerca de la investigación y la naturaleza de mi participación en ella, ha sido clara y suficiente, y comprendo que:

1. Si accede a participar en el estudio y cumple con los criterios de inclusión, usted debe contestar varios cuestionarios antes de iniciar la investigación, durante la misma y un mes después de finalizar.
2. La información obtenida es de carácter confidencial, lo que significa que por ninguna razón será divulgada mi identidad o se entregará información a instituciones ajenas a las suscritas en esta investigación, y que si fuere preciso se tomarán las medidas necesarias para salvaguardar mis datos personales y mi identidad, utilizando códigos para identificar a los participantes garantizando su anonimato y encriptando los datos con la protección adecuada a nivel informático. Lo anterior en conformidad con las Leyes 1266 de 2008, 1273 de 2009, 1581 de 2012 y la Resolución 2654 de 2019.
3. La plataforma web en la cual se alojará el software se acoge a todas normas y estatutos aplicables en territorio nacional (Ley 527 de 1999, Ley 1266 de 2008, Ley 1273 de 2009, Ley 1581 de 2012, Resolución 2654 de 2019), garantizando la seguridad privacidad y protección de los datos personales, clínicos y de contacto de los participantes.
4. La confidencialidad que obliga a los investigadores a guardar el secreto profesional solo puede romperse en situaciones en las que de no hacerlo llevaría a un daño inminente de la persona o a terceros, en concordancia con la Ley 1090 de 2006.
5. Se trabajará con la información publicada por los participantes en sus redes sociales ya sea de manera pública o privada, para este último caso nos ajustaremos al consentimiento asistido y al contrato de protección de información que tenga el usuario con la red social.
6. Los resultados de la investigación serán utilizados con fines académicos y pueden ser divulgados a través de publicaciones académicas, como establece la ley 1090 de 2006. 7. Al tratarse de un programa de apoyo a la evaluación psicológica, la participación en este estudio podría implicar un riesgo mínimo para mi salud emocional o psicológica,

sin embargo no existe ningún riesgo conocido para mi integridad física o la de las personas que me rodean.

1

8. En caso de detectarse una situación de alto riesgo para los participantes, se ofrecerán rutas de acceso a los servicios de salud pertinentes de acuerdo con el caso y se interrumpirá la participación en el estudio indefinidamente.
9. Toda la información que voy a proporcionar será suficiente y sincera, de acuerdo con los requerimientos de la investigación.
10. En cualquier momento puedo expresar mis inquietudes o solicitar la ampliación de la información acerca del estudio. Asimismo, me ha sido informado que mi participación en la investigación es completamente voluntaria, por lo tanto, puedo retirarme en cualquier momento si lo deseo, incluso posteriormente a la firma de este documento.
11. Mi participación en la investigación no implica el pago de remuneración alguna, ni el acceso a otros beneficios económicos o en especie.
12. Si desea mayor información o conocer los resultados del estudio, puede comunicarse al correo electrónico [metodoautomaticodepresion@gmail.com](mailto:metodoautomaticodepresion@gmail.com).



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 2**  
**Estado del Arte Del Proyecto**

**ANDRÉS DAVID DOMÍNGUEZ ROZO**  
**CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA**  
**FACULTAD DE INGENIERÍA**  
**PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**  
**BOGOTÁ**  
**2021**

Etiquetas de búsqueda	Ubicación en la red	Revista	Título	Autor	Año	Palabras clave	Resumen
Depression and Colombia and Datos	<a href="https://www.minsalud.gov.co/files/7/6/1/76102/Bolet%C3%ADN%20N.%20197/Bolet%C3%ADN%20N.%20197.pdf">https://www.minsalud.gov.co/files/7/6/1/76102/Bolet%C3%ADN%20N.%20197/Bolet%C3%ADN%20N.%20197.pdf</a>	Ministerio de salud de Colombia	Boletín de salud mental Depresión Subdirección de Enfermedades No Transmisibles	Ministerio de salud de Colombia	Marzo de 2017	Salud, Depresión, Enfermedades, Enfermedad, Depresión, Moderada, Perseón, Casos, Edad.	La Subdirección de Enfermedades No Transmisibles, a través del grupo de Gestión Integrada para la Salud Mental, publica un boletín trimestral, proporcionando información de interés en salud mental al personal que trabaja en el sector salud y a la comunidad en general; este boletín será el primero de 2017 y se centrará en la depresión. América Latina y el Caribe experimentan una elevada carga de la enfermedad en términos de morbilidad, mortalidad y discapacidad; dentro de este grupo de enfermedades, la depresión se ha convertido en un diagnóstico frecuente en los servicios de atención primaria en salud a nivel de atención de primer nivel. Se encuentra entre los primeros cinco trastornos que generan mayor discapacidad de largo plazo y dependencia. El Sistema Integral de Información de la Protección Social reporta que las atenciones por depresión se han incrementado desde el año 2009 y el diagnóstico de depresión moderada se realizó en 36.304 atenciones en 2013.
Depression and Symptoms	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Universidade Federal de Minas Gerais, Belo Horizonte, Brazil	Self-compassion and Symptoms of Stress, Anxiety, and Depression (Autoempatía y síntomas de estrés, ansiedad y depresión)	de Souza, L.R. (Autor de correo electrónico, Polícaro, D. A., Hutz, C.S e	Volume 26, Número 1, 1 de marzo de 2020	Autoempatía, Ansiedad, Estrés, Ansiedad, Estudios, Estrés	Este estudio ha examinado las relaciones entre la autoempatía y variables como la ansiedad, depresión y estrés. La mayoría de los estudios encontraron una correlación negativa entre la autoempatía y la ansiedad, el estrés y la depresión. Se encontró una correlación negativa significativa entre la autoempatía y todos los puntajes DAS-21. También se confirmaron diferencias significativas de género con respecto a la autoempatía, con los participantes masculinos con una puntuación más alta que las mujeres, a pesar de que no hubo diferencias entre hombres y mujeres en el DAS-21. Los practicantes de religión cristiana presentaron puntuaciones más bajas en el DAS-21 que los participantes sin religión.
depression and symptoms and identification methods	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand	Text Mining Analysis of Comments in Thai Language for Depression (Análisis de minería de textos de los comentarios en tailandés para la depresión de las redes sociales en línea)	Chaiyaprasit, PEmail Author,	Volume 899, 2021, Pages 302-313	Autoempatía, Comentarios, Palabras, Clasificación de sentimientos, Opinión Minera, Revisión de producto.	Los objetivos de esta investigación fueron analizar la relación de las frases o palabras comúnmente encontradas en los comentarios de depresión en Twitter usando las reglas de asociación. Según los comentarios en tailandés en las redes sociales recopiladas mediante el uso de software RapporMiner Studio 9 para analizar las respuestas de las palabras de un comentario de texto para obtener el formato Asociación de datos. De acuerdo con el rendimiento de modelos en cada uno de los métodos anteriores, se encontró que la Distancia euclidiana proporcionó el mejor resultado debido a la distancia promedio más pequeña en todos los puntos de cada grupo, que fue igual a 153.504.
depression and software	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	DETI/FEPA, University of Aveiro, Aveiro, Portugal	Machine Learning for Depression Screening in Online Communities (Aprendizaje automático para la detección de depresión en comunidades en línea)	Trifun, A. Email Author, Antunes, F. Email Author, Oliveira, J.L.	Volume 1240 AISC, 2021, Pages 102-111	Claves, Redes sociales, búsqueda, línea	En este artículo se presenta un enfoque computacional para la clasificación de usuarios deprimidos de redes sociales. Realizamos un estudio de evaluación cruzada basado en dos conjuntos de datos públicos, recopilados de la red social para comprender el impacto de la transferencia de aprendizaje cuando la fuente de datos es prácticamente la misma.
depression and technology and identification	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Nature Communications	A unified neurocognitive model of semantic language social behaviour and face recognition in semantic dementia (Un modelo neurocognitivo unificado de semántica, lenguaje, comportamiento social y reconocimiento facial en la demencia semántica)	Ding, J. a,b, Chen, K. a, Liu, H. d, Huang, L. e, Chen, Y. a, Lu, Y. a, Tang, C. h, Guo, Q. e Email Author, Han, Z. a Email Author, Lamborn, R. a, n, U.S.A.	Volume 11, Issue 1, 1 December 2020, Article number 2995	Memoria Semántica, Lóbulo temporal, Relaciones temáticas	Los lóbulos temporales anteriores de han convertido en una región central clave de interés en la neurociencia cognitiva basada en investigaciones neuropsicológicas de la demencia semántica. Aquí mostramos que el grado de deterioro semántico generado está relacionado con la atrofia AT, bilateral total de los pacientes. Se encuentra que la atrofia se relaciona positivamente con el grado de atrofia orbitofrontal.
depression and technology and identification	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Heliyon	Depression and disclosure behavior via social media: A study of university students in China (Comportamiento de depresión y divulgación a través de las redes sociales: un estudio de estudiantes universitarios en China)	Marie U. Zou, Mandy Xiaoyang Lin, Vincent Chu	Volume 6, Issue 2, February 2020, e03368	Psicología, Depresión, Intención de divulgación, Comportamiento de divulgación, Redes sociales, Estudiantes universitarios, China	El estrés severo y la depresión constituyen serios desafíos para la salud tanto personal como pública. Siguiendo las tendencias de la época, las redes sociales han proporcionado una plataforma que mejorará y se han convertido en una nueva plataforma para divulgar reflexiones a la depresión. Además de la intención de divulgación, este estudio también exploró los encuentros han presentado su depresión a través de partes de redes sociales.
psychology and text	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Engineering Applications of Artificial Intelligence	Decision making with natural language based preferences and preferences (Un modelo neurocognitivo unificado de semántica, lenguaje, comportamiento social y reconocimiento facial en la demencia semántica)	Ingrid Nunes, Simon Miles, Michael Luck, Jimmie Barbosa, Carlos Lucena	Junio 2013	Decisiones, Técnica, Preferencias, Sistemas	La toma de decisiones es un requisito para muchas tareas, como la compra, hoy en día asistida por sistemas de software, y dar soporte al proceso de toma de decisiones es una característica que mejorará significativamente dichos sistemas. En este artículo, presentamos una técnica de toma de decisiones automatizada, que elige una opción de conjunto de las disponibles en función de las preferencias y prioridades expresadas en un lenguaje de preferencia de alto nivel, explorando términos del lenguaje natural, como los actos de habla expresivos.
psychology and language	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	International Encyclopedia of the Social & Behavioral Sciences	Text Comprehension: Models in Psychology (Comprensión de textos: modelos en psicología)	Autor links open overlay panel R. G. Dixon & M. Wolf	2001	Comprensión, Texto, Textos, Modos	Los modelos de comprensión de textos describen y explican los procesos involucrados en la comprensión y el recuerdo de la información verbal. La comprensión de textos se evalúa tanto en tareas de lectura de alta a bajo, o examinando los índices de la representación mental de texto que resulta de la comprensión. Actualmente, los modelos computacionales de comprensión de textos más influyentes son el modelo de sistema de producción basado en activación contextualizado de Kintsch y Carpenter, y el modelo Construcción-Integración de Kintsch.
depression and language	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Expert Systems with Applications	A text classification framework for simple and effective early depression detection (Un marco de clasificación de textos para la detección temprana simple y efectiva de la depresión a través de las redes sociales)	Jorge G. Burdick, Marcelo Escrivada, Manuel Montoya-Gómez	1 de noviembre de 2020	Modelo, Aprendizaje, Detección, Temprana, Modos, Clasificación, Depresión	Con el auge de Internet, existe una necesidad creciente de construir sistemas inteligentes que sean capaces de tratar de manera eficiente los problemas de detección temprana de riesgo en las redes sociales, como la detección temprana de depresión, la detección temprana de tumores o la identificación de depresores reactivos. Los resultados experimentales muestran que nuestro clasificador fue capaz de sustraer a estos modelos y clasificadores estándar, a pesar de ser mucho menos computacionalmente intensos la capacidad de explicar su razón de ser. Evaluamos nuestro modelo en la tarea piloto Affect 2017 de CLEF sobre detección temprana de depresión. La mayoría de las 30 contribuciones enviadas a este concurso utilizaron métodos de última generación. Los resultados experimentales muestran que nuestro clasificador fue capaz de sustraer a estos modelos y clasificadores estándar, a pesar de ser mucho menos computacionalmente intensos la capacidad de explicar su razón de ser.
depression and identification by text	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	Encyclopedia of Language & Linguistics (Second Edition)	Social Psychology and Language (Psicología social y lenguaje)	J.J. Braxos, S.A. Reid	2006	Verbalidad etnolingüística, Lengua y psicología social, actitudes lingüísticas, identidad social.	Los humanos son criaturas que usan símbolos. El argumento de que los seres humanos son usuarios de símbolos esencial y exclusivamente ha sido formulado con fuerza por varios estudiosos. Sin embargo, los humanos también son animales psicológicos, de ahí el surgimiento de la psicolingüística. Además, los humanos son animales sociales, de ahí la sociolingüística.
depression AND machine AND learning	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	SciTePress	Feature engineering for depression detection in social media	Stankovic, M., Isakov, V., Devayn, D., Siminovic, I.	2016	clasificación, detección de depresión, machine learning, social media	Si investigación se basa en la tarea piloto Affect 2017 que se centra en la detección temprana del riesgo de depresión. El conjunto de datos de CLEF Affect 2017 consta de ejemplos de texto recopilados de mensajes de 87 usuarios de Reddit. La idea principal de la tarea es diseñar a los usuarios en dos grupos: caso de riesgo de depresión y caso no de riesgo.
predicting AND depression	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	AAAI press	Predicting depression via social media	De Choudhury, M., Gamon, M., Counts, S., Horvitz, E.	2013	clasificación, depresión, twitter, medio social	La depresión mayor constituye un serio desafío para la salud pública y personal. Cada año, decenas de millones de personas sufren depresión y solo una fracción recibe el tratamiento adecuado. Exploramos el potencial del uso de las redes sociales para detectar y diagnosticar el trastorno depresivo mayor en las personas. Primero empleamos el crowdsourcing para compilar un conjunto de usuarios de Twitter que informan haber sido diagnosticados con depresión clínica, según un instrumento psicométrico estándar.
twitter AND analysis	<a href="https://www.sciopus.com/ucsf/doi/10.1002/ps.2242">https://www.sciopus.com/ucsf/doi/10.1002/ps.2242</a>	European Language Resource Association (ELRA)	Twitter as a corpus for sentiment analysis and opinion mining	Pak, A., Paroubek, P.	2010	lingüística, herramientas de comunicación, análisis de sentimientos, minería de datos	El microblogging hoy en día se ha convertido en una herramienta de comunicación muy popular entre los usuarios de Internet. Millones de usuarios comparten opiniones sobre diferentes aspectos de la vida cotidiana. Por lo tanto, los sitios web de microblogging son fuentes ricas de datos para la extracción de opiniones y el análisis de sentimientos. Dado que el microblogging ha aparecido hace relativamente poco tiempo, existen algunos tabajos de investigación que se dedicaron a este tema. En nuestro artículo, nos centramos en el uso de Twitter, la plataforma de microblogging más popular, para la tarea de análisis de sentimientos.

Machine learning AND python	<a href="https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning">https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning</a>	JMLA	Scikit-learn: Machine learning in Python	Hedgrera F, Vanouque G, Gramfort, A, Michel, V, Thirion, B, Grisel, O, Blondel, M, Prettenhofer, R, Weiss, R, Dubourg, V, Vanderplas, J, Passos, A, Cournapeau, D, Brucher, M, Perrot, M, Duchesnay, E.	2011	selección de modelo, python, aprendizaje supervisado, aprendizaje no supervisado	Scikit-learn es un módulo de Python que integra una amplia gama de algoritmos de aprendizaje automático de última generación para problemas supervisados y no supervisados de manera sencilla. Esta paquete se centra en llevar el aprendizaje automático a los no especialistas mediante un lenguaje de alto nivel de uso general, de fácil integración en la documentación y la concurrencia de la API.
Depression AND model	<a href="https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning">https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning</a>	springer	A depression detection model based on sentiment analysis in micro-blog social network	Wang, X., Zhang, C., Li, Y., Sun, L., Wu, L., Bao, Z	2013	análisis de sentimientos, minería de datos, depresión, detección de modelos	Los conjuntos de datos que se originan en las redes sociales son valiosos para muchos campos como la sociología y la psicología. Pero los apoyos desde la perspectiva técnica están lejos de ser suficientes y se necesitan con urgencia enfoques específicos. Este trabajo aplica la minería de datos al área de psicología para la detección de usuarios deprimidos en servicios de redes sociales. En primer lugar, se propone un método de análisis de sentimientos que utiliza vocabulario y reglas creadas por el hombre para calcular la inclinación a la depresión de cada microblog. En segundo lugar, se construye un modelo de detección de depresión basado en el método propuesto y 10 características de los usuarios deprimidos derivadas de la Investigación psicológica.
Depression AND machine learning	<a href="https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning">https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning</a>	springer	Depression Detection Among Social Media Users Using Machine Learning	Verma, P., Sharma, K., Wala, G.S	2020	matriz de confusión, knn, naive bayes, Twitter API, um	En el mundo actual, la depresión entre las personas se está volviendo muy común y, según la Organización Mundial de la Salud (OMS) para 2020, la depresión se convertirá en la segunda causa principal de muerte y discapacidad. La depresión se puede medir en humanos midiendo sus actividades diarias y condiciones físicas, y no es tan fácil y ha sido un desafío en los últimos días. Es poco probable que las personas compartan fácilmente sus verdaderos sentimientos / emociones internas con nadie, por lo que es importante detectar la depresión humana mediante cualquier máquina en lugar de ser humano y proporcionar un sistema de apoyo inmediatamente basado en el nivel de depresión.
Predicting AND depression	<a href="https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning">https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning</a>	IEEE	Predicting Depression Levels Using Social Media Posts	Aldawish, M.M., Ahmad, H.F	2017	social network social, maquina de soporte vectorial, diagnóstico, salud	El uso de los sitios de redes sociales (SNS) está aumentando hoy en día especialmente entre las generaciones más jóvenes. La disponibilidad de SNS permite a los usuarios expresar sus intereses, sentimientos y compartir la rutina diaria. Muchos investigadores demuestran que usar contenido generado por el usuario (CGU) de manera correcta puede ayudar a determinar los niveles de salud mental de las personas. Extraer el UGC podría ayudar a predecir los niveles de salud mental y la depresión.
detection AND social media	<a href="https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning">https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning</a>	IEEE	MIDAS: Mental illness detection and analysis via social media	Saravia, E., Chang, C.-H., De Lorenzo, R. J., Chen, Y.-S.	2016	adquisición de información, social networking, servicios web, patrones de comportamiento, enfermedades mentales	Las enfermedades mentales se encuentran entre las condiciones más discapacitantes y afectan a millones de personas en todo el mundo. En general, el principal desafío de los trastornos mentales es que siguen siendo difíciles de detectar en los pacientes que los sufren. En un entorno en línea, el desafío se extiende a la recopilación de datos de los pacientes y la implementación de algoritmos adecuados para ayudar en la detección de tales enfermedades.
depression AND behavior AND twitter	<a href="https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning">https://www.scribd.com/document/400758464/Depression-Detection-Model-Using-Machine-Learning</a>	IEECC	Social media, big data, and public health informatics: Rumormongering behavior of depression revealed through twitter	Nambisan, P., Luo, Z., Kapoor, A., Patrick, T.B., Coker, B.A.	2015	big data, salud, información comunicativa, búsqueda de comportamiento	Los trastornos depresivos no diagnosticados y no tratados se han convertido en un problema grave de salud pública y prevalecen entre personas de todas las edades, sexos y razas. Los sitios de redes sociales, como Twitter, se han convertido en un lugar importante para que las personas expresen / muestren sus sentimientos y sentimientos. Los tweets de estos sitios de microblogs podrían usarse para detectar y potencialmente detectar depresión. Hasta la fecha, los estudios en esta área se han centrado en desarrollar y validar los términos y vocabulario utilizados por los usuarios con depresión, o evaluar los tweets relacionados con la depresión utilizando términos que son sinónimos de depresión.



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 3  
PREGUNTAS CUESTIONARIO ESCRITO AJUSTADAS  
DESPUÉS DE LA VALIDACIÓN**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

**ANEXO 3. MÉTODO AUTOMÁTICO PARA EL APOYO EN LA IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A TRAVÉS DE NARRATIVA ESCRITA**

**PREGUNTAS CUESTIONARIO ESCRITO AJUSTADAS DESPUES DE LA VALIDACIÓN**

**INSTRUCCIÓN**

Te animamos a responder las siguientes preguntas, evita las respuestas cortas, siéntete en la libertad de expresar lo que sientes.

<b>CRITERIOS</b>	<b>PREGUNTA</b>	<b>FORMATO</b>
<b>ESTADO DE ÁNIMO ANHEDONIA</b>	<b>1.</b> Piensa en algunas situaciones, en las dos últimas semanas, en las que te hayas sentido triste o desanimado(a) y descríbenos qué pensabas y qué sentías	ESCRITO
<b>AGITACIÓN/ RETRASO PSICOMOTOR FATIGA DETERIORO EN LOS ÁMBITOS DISMINUCIÓN DE LA CONCENTRACIÓN</b>	<b>2.</b> ¿En las dos últimas semanas te has sentido con menos energía (de lo normal) o con dificultades para concentrarte en las actividades diarias? Descríbenos más detalladamente... ¿En las últimas semanas te has sentido más acelerado o lento (de lo normal) al enfrentar las actividades diarias?	ESCRITO



	<p>Desríbenos más detalladamente...</p>	
--	---	--

<p><b>SUEÑO ALIMENTACIÓN</b></p>	<p><b>3.</b> ¿Has tenido cambios en tus hábitos o rutina de alimentación? (has mantenido tus horarios habituales de alimentación, te saltas alguna comida o ha aumentado o disminuido tu apetito) Desríbenos cómo fueron esos cambios...</p>	<p>ESCRITO</p>
<p><b>INUTILIDAD DESESPERANZA BAJA EFICACIA</b></p>	<p><b>4.</b> ¿Qué tan satisfecho(a) y cómodo(a) te sentiste con la manera en la que enfrentaste las situaciones que fueron difíciles en las últimas dos semanas? ¿Las cosas resultaron como querías? Desríbenos la situación...</p>	<p>ESCRITO</p>
<p><b>PENSAMIENTOS DE MUERTE ESPERANZA EN EL FUTURO FRUSTRACIÓN REDES DE APOYO</b></p>	<p><b>5.</b> ¿Has tenido deseos o ideas de morir en las últimas dos semanas? ¿O has pensado que no vale la pena vivir? Desríbenos al respecto...</p>	<p>ESCRITO</p>



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 4**  
**Enumeración y Clasificación De Conceptos**

**ANDRÉS DAVID DOMÍNGUEZ ROZO**  
**CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA**  
**FACULTAD DE INGENIERÍA**  
**PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**  
**BOGOTÁ**  
**2021**

Síntomas depresivos	General	Agrupación	Descripción
1	Ánimo	Estado de ánimo deprimido	Estado de ánimo deprimido la mayor parte del día, casi todos los días, según se desprende de la información subjetiva (p. ej., se siente triste, vacío, sin esperanza) o de la observación por parte de otras personas (p. ej., se le ve lloroso). (Nota: En niños y adolescentes, el estado de ánimo puede ser irritable.)
2	Apatía	Disminución de interés o placer	Disminución importante del interés o el placer por todas o casi todas las actividades la mayor parte del día, casi todos los días (como se desprende de la información subjetiva o de la observación).
3	Alimentación	Pérdida de peso	Pérdida importante de peso sin hacer dieta o aumento de peso (p. ej., modificación de más de un 5% del peso corporal en un mes) o disminución o aumento del apetito casi todos los días. (Nota: En los niños, considerar el fracaso para el aumento de peso esperado.)
		Aumento de peso	Incremento del peso corporal que ocurre cuando una persona consume más calorías de las que el cuerpo necesita o utiliza. El comer en exceso o la inactividad como resultado de la depresión pueden provocar aumento de peso. Algunas personas adelgazan como consecuencia de la depresión. En cambio, es posible que una mejora en el apetito asociada con una mejora en el estado de ánimo tenga como consecuencia un aumento de peso
4	Descanso	Hipersomnia	La hipersomnia es el síntoma (queja subjetiva del paciente) consistente en tendencia al sueño en situaciones anormales. Es decir, es una facilidad excesiva para conciliar el sueño durante el día, en situaciones que a la mayoría de las personas no le sucede. Es un síntoma que puede aparecer en múltiples enfermedades
		Insomnio	El insomnio es un trastorno del sueño común y afecta con mayor frecuencia a los hombres. El insomnio puede causar somnolencia diurna y falta de energía. También puede hacerle sentir

			ansioso, deprimido o irritable.
5	Actividad física	Agitación	Agitación o retraso psicomotor casi todos los días (observable por parte de otros, no simplemente la sensación subjetiva de inquietud o de enlentecimiento).
		Retraso psicomotor	Observable por parte de otros, no simplemente la sensación subjetiva de inquietud o de enlentecimiento.
6	Desaliento	Fatiga o pérdida de energía	Es una sensación de falta de energía, de agotamiento o de cansancio.
7	Carga de conciencia	Sentimiento de inutilidad	Sentimiento de inutilidad o culpabilidad excesiva o inapropiada (que puede ser delirante) casi todos los días (no simplemente el autorreproche o culpa por estar enfermo).
		Sentimiento de culpabilidad	la culpa alude a la mala conciencia y al sentimiento negativo que la persona experimenta. Para que el sentimiento de culpa aparezca tiene que haber tres elementos principales: la causa o la razón que la origina, ya sea real o imaginaria; la percepción y autovaloración negativa de la situación, es decir la mala conciencia; y por último, la emoción negativa derivada de la culpa propiamente dicha: el remordimiento.
8	Atención dispersa	Disminución de la capacidad de pensar	Disminución de la capacidad para pensar o concentrarse, o para tomar decisiones, casi todos los días (a partir de la información subjetiva o de la observación por parte de otras personas).
		Disminución de la capacidad de concentrarse	es un trastorno en la capacidad de atención que impide la realización de las tareas más cotidianas, tales como escuchar en una conversación, leer, ver una película, estudiar o llevar a cabo otro tipo de actividad.
		Disminución de la capacidad de tomar decisiones	Una persona siente que en una toma de decisiones va acompañado de sensación excesiva de miedo a equivocarse y miedo a la crítica por parte de los demás, se puede tener un problema de toma de decisiones
		Pensamientos de muertes recurrentes	Pensamientos de muerte recurrentes (no sólo miedo a morir), ideas suicidas recurrentes sin un plan determinado, intento de suicidio o un plan específico para llevarlo a cabo
9	Ideas suicidas		

		Miedo a morir	Los individuos que presentan este problema suelen tener personalidades obsesivas y concentran su tiempo en pensamientos referentes a la muerte, desde cuidarse en exceso para no enfermar, hasta evitar acudir al entierro de un familiar para no tener que ver el cuerpo.
--	--	---------------	--



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 5**  
**Enumeración y Clasificación De Conceptos**

**ANDRÉS DAVID DOMÍNGUEZ ROZO**  
**CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA**  
**FACULTAD DE INGENIERÍA**  
**PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**  
**BOGOTÁ**  
**2021**

Ontology(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>>

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Abandono>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Abatimiento>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Abrumada>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Abrumado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Aburrido>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Aburrimiento>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Actividad\_física>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Acusación>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Acusado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Afectado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Aflicción>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Afligida>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Afligido>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Agitacion>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Agitado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Agobiado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Aislado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Aislamiento>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresio>  
n#Alejado>))

Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Alimentación>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Amargado>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Amargura>>))  
Declaration(Class(<[http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Amor\\_propio](http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Amor_propio)>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Angustia>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Ansioso>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Apagado>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Apatico>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Apenada>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Apenado>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Apetito>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Aplazar>>))  
Declaration(Class(<<http://www.semanticweb.org/arusu/ontologies/2021/3/depresion#Apático>>))





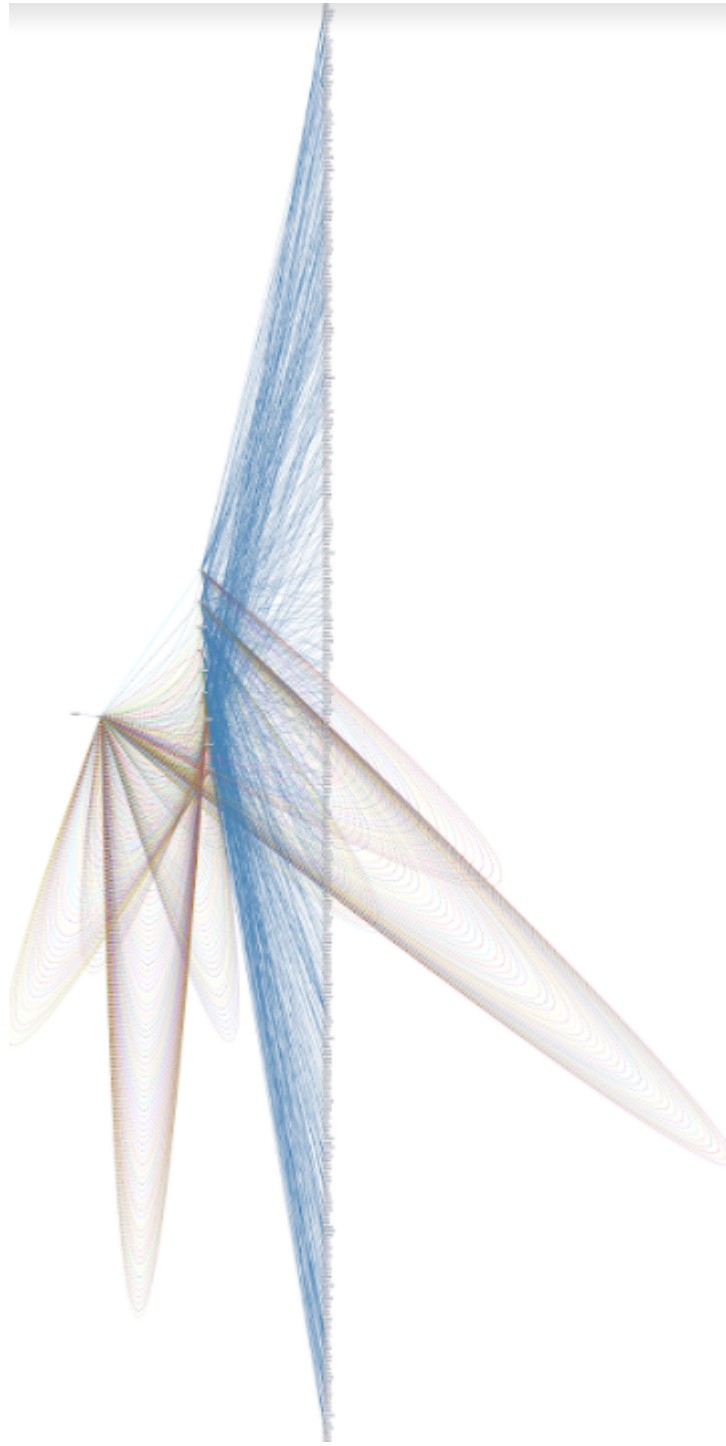
**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 6**  
**Mapa Ontologia**

**ANDRÉS DAVID DOMÍNGUEZ ROZO**  
**CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA**  
**FACULTAD DE INGENIERÍA**  
**PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**  
**BOGOTÁ**  
**2021**





**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 7**

**Log generado por Reasoner-Protégé**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

```

        {\rtf1\ansi\ansicpg1252\cocoartf2578
\cocoatextscaling0\cocoaplatform0{\fonttbl\f0\fswiss\fcharset0
        Helvetica;}
        {\colortbl;\red255\green255\blue255;}
        {\*\expandedcolortbl;;}
        \margl1440\margr1440\vieww11520\viewh8400\viewkind0
\pard\tx566\tx1133\tx1700\tx2267\tx2834\tx3401\tx3968\tx4535\tx5102\tx
5669\tx6236\tx6803\pardirnatural\partightenfactor0

\f0\fs24 \cf0 2021-04-11 20:18:44.182 [Close OSGi Session] INFO
  Launcher ----- Shutting down Protege
        -----\
2021-04-11 20:18:44.183 [Close OSGi Session] INFO Launcher
  Cleaning up temporary directories\
2021-04-11 20:27:38.990 [main] INFO Launcher
*****\
        2021-04-11 20:27:38.999 [main] INFO Launcher **
  Protege **\
2021-04-11 20:27:39.000 [main] INFO Launcher
*****\
        2021-04-11 20:27:39.000 [main] INFO Launcher \
2021-04-11 20:27:39.000 [main] INFO Launcher -----
  Initialising and Starting the OSGi Framework -----\
2021-04-11 20:27:39.000 [main] INFO Launcher FrameworkFactory
  Class: org.apache.felix.framework.FrameworkFactory\
2021-04-11 20:27:39.003 [main] INFO Launcher \
2021-04-11 20:27:39.577 [main] INFO Launcher The OSGi framework
  has been initialised\
2021-04-11 20:27:40.809 [main] INFO Launcher
  ----- Starting Bundles
        -----\
2021-04-11 20:27:40.812 [main] INFO Launcher Starting bundle
  org.protege.common\
2021-04-11 20:27:40.814 [main] INFO Launcher Starting bundle
  org.eclipse.equinox.common\
2021-04-11 20:27:40.815 [main] INFO Launcher Starting bundle
  org.eclipse.equinox.supplement\
2021-04-11 20:27:40.816 [main] INFO Launcher Starting bundle
  org.protege.editor.core.application\
2021-04-11 20:27:40.817 [main] INFO Launcher Starting bundle
  org.eclipse.equinox.registry\
2021-04-11 20:27:40.817 [main] INFO Launcher Starting bundle
  log4j.over.slf4j\
2021-04-11 20:27:40.818 [main] INFO Launcher Starting bundle
  jul.to.slf4j\

```

2021-04-11	20:27:40.818	[main]	INFO	Launcher	Starting bundle	
				ch.qos.logback.core\ slf4j.api\ org.apache.commons.io\ org.semanticweb.owl.owlapi\ ch.qos.logback.classic\ org.eclipse.equinox.common\ org.protege.common\ org.protege.editor.owl\ org.jsr-305\ org.apache.servicemix.bundles.javax-inject\ org.apache.servicemix.bundles.aopalliance\ org.protege.editor.core.application\ com.googlecode.javaewah.JavaEWAH\ org.eclipse.jgit\ com.google.guava\ org.eclipse.equinox.registry\ org.eclipse.equinox.supplement\ org.protege.owl.rdf\ uk.ac.manchester.cs.owl.factplusplus\ org.protege.editor.owl.rdf\ existentialquery\ csv-export-plugin\ uk.ac.manchester.cs.owl.explanation\ org.protege.ontograf\ 		
2021-04-11	20:27:40.819	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.819	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.820	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.820	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.820	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.821	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.821	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.821	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.822	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.823	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.823	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.824	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.824	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.825	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.826	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.826	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.827	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.828	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.829	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.832	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.833	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.833	[main]	INFO	Launcher	Starting bundle	
2021-04-11	20:27:40.834	[main]	INFO	Launcher	Starting bundle	

```

2021-04-11 20:27:40.835 [main] INFO Launcher Starting bundle
                        org.coode.browser\
2021-04-11 20:27:40.835 [main] INFO Launcher Starting bundle
                        com.clarkparsia.protege.plugin.pellet\
2021-04-11 20:27:40.836 [main] INFO Launcher Starting bundle
                        org.coode.owlviz\
2021-04-11 20:27:40.836 [main] INFO Launcher Starting bundle
                        org.semanticweb.hermit\
2021-04-11 20:27:40.837 [main] INFO Launcher Starting bundle
                        cellfie\
2021-04-11 20:27:40.837 [main] INFO Launcher Starting bundle
                        snap-sparql-query-plugin\
2021-04-11 20:27:40.838 [main] INFO Launcher Starting bundle
                        org.coode.dlquery\
2021-04-11 20:27:40.838 [main] INFO Launcher Starting bundle
                        swrltab-plugin\
2021-04-11 20:27:40.838 [main] INFO Launcher Starting bundle
                        code-generation\
2021-04-11 20:27:45.111 [main] INFO Launcher The OSGi framework
                        has been started\
                2021-04-11 20:27:45.112 [main] INFO Launcher \
2021-04-11 20:27:45.116 [FelixDispatchQueue] INFO ProtegeApplication
                ----- Protege
                -----\
2021-04-11 20:27:45.123 [FelixDispatchQueue] INFO ProtegeApplication
                Protege Desktop\
2021-04-11 20:27:45.123 [FelixDispatchQueue] INFO ProtegeApplication
                Version 5.5.0, Build \
2021-04-11 20:27:45.127 [FelixDispatchQueue] INFO ProtegeApplication
                \
2021-04-11 20:27:45.127 [FelixDispatchQueue] INFO ProtegeApplication
                \
2021-04-11 20:27:45.127 [FelixDispatchQueue] INFO ProtegeApplication
                ----- Platform
                -----\
2021-04-11 20:27:45.127 [FelixDispatchQueue] INFO ProtegeApplication
                Java: JVM 1.8.0_121-b13 Memory: 1908M\
2021-04-11 20:27:45.127 [FelixDispatchQueue] INFO ProtegeApplication
                Language: es, Country: ES\
2021-04-11 20:27:45.128 [FelixDispatchQueue] INFO ProtegeApplication
                Framework: Apache Software Foundation (1.8) \
2021-04-11 20:27:45.128 [FelixDispatchQueue] INFO ProtegeApplication
                OS: macosx (10.16)\
2021-04-11 20:27:45.128 [FelixDispatchQueue] INFO ProtegeApplication
                Processor: x86-64\
                \
2021-04-11 20:27:45.128 [FelixDispatchQueue] INFO ProtegeApplication
                \

```

```

2021-04-11 20:27:45.128 [FelixDispatchQueue] INFO ProtegeApplication
----- Plugins
-----\
2021-04-11 20:27:45.129 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: OWLAPI RDF Library (3.0.0)\
2021-04-11 20:27:45.129 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Factplusplus Plug-in (1.6.5)\
2021-04-11 20:27:45.129 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: SPARQL Query Plugin (3.0.0)\
2021-04-11 20:27:45.129 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Existential Query (2.0.0)\
2021-04-11 20:27:45.129 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: CSV Export Plugin (1.0.0)\
2021-04-11 20:27:45.129 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Explanation Workbench (3.0.0)\
2021-04-11 20:27:45.130 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: OntoGraf (2.0.3)\
2021-04-11 20:27:45.130 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Browser View (OWLDoc) (3.0.3)\
2021-04-11 20:27:45.130 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Pellet Reasoner Plug-in (2.2.0)\
2021-04-11 20:27:45.130 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: OWLViz (5.0.3)\
2021-04-11 20:27:45.130 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Hermit Reasoner (1.4.3.456)\
2021-04-11 20:27:45.130 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: Cellfie Protege 5.0+ Plugin (2.1.0)\
2021-04-11 20:27:45.131 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: snap-sparql-query-plugin (6.0.0)\
2021-04-11 20:27:45.131 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: DL Query (4.0.1)\
2021-04-11 20:27:45.131 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: SWRLTab Protege 5.0+ Plugin (2.0.6)\
2021-04-11 20:27:45.131 [FelixDispatchQueue] INFO ProtegeApplication
Plugin: OWL Code Generation Plug-in (2.0.0)\
2021-04-11 20:27:45.131 [FelixDispatchQueue] INFO ProtegeApplication
\
2021-04-11 20:27:46.021 [FelixDispatchQueue] INFO ProtegeApplication
Creating and setting up empty (default) editor kit\
2021-04-11 20:27:46.024 [AWT-EventQueue-0] INFO ProtegeApplication
Received request to edit document at
file:/Users/daviddominguez/Library/Mobile%20Documents/com~apple~CloudD
ocs/Tesis/FINAL%20Ontologia%20Sintomatologia%20Depresiva%20(Escrita).o
wl\
2021-04-11 20:27:46.025 [AWT-EventQueue-0] INFO ProtegeApplication
Application is initialized. Opening URI.\
2021-04-11 20:27:46.026 [AWT-EventQueue-0] INFO ProtegeApplication
Creating and setting up (default) editor kit for

```

```

file:/Users/daviddominguez/Library/Mobile%20Documents/com~apple~CloudD
ocs/Tesis/FINAL%20Ontologia%20Sintomatologia%20Depresiva%20(Escrita).o
wl\
2021-04-11 20:27:46.113 [FelixDispatchQueue] INFO OWLEditorKit    OWL
      API Version: 4.5.9.2019-02-01T07:24:44Z\
2021-04-11 20:27:47.851 [FelixDispatchQueue] INFO OWLModelManagerImpl
      Cannot generate ontology catalog for ontology at
http://www.semanticweb.org/daviddominguez/ontologies/2021/3/untitled-o
ntology-36. URI scheme is not "file"\
2021-04-11 20:27:49.251 [FelixDispatchQueue] INFO ProtegeApplication
      ----- Auto-update Check
      -----\
2021-04-11 20:27:49.251 [AWT-EventQueue-0] INFO OWLEditorKit    OWL
      API Version: 4.5.9.2019-02-01T07:24:44Z\
2021-04-11 20:27:49.261 [FelixDispatchQueue] INFO ProtegeApplication
      Auto-update last performed: Sun Apr 11 12:14:10 COT 2021\
2021-04-11 20:27:49.281 [FelixDispatchQueue] INFO ProtegeApplication
      Auto-update has been performed today. Not performing it again.\
2021-04-11 20:27:49.281 [FelixDispatchQueue] INFO ProtegeApplication
      \
2021-04-11 20:27:49.298 [AWT-EventQueue-0] INFO OWLModelManagerImpl
      ----- Loading Ontology
      -----\
2021-04-11 20:27:49.298 [AWT-EventQueue-0] INFO OWLModelManagerImpl
      Loading ontology from
file:/Users/daviddominguez/Library/Mobile%20Documents/com~apple~CloudD
ocs/Tesis/FINAL%20Ontologia%20Sintomatologia%20Depresiva%20(Escrita).o
wl\
2021-04-11 20:27:50.034 [pool-27-thread-1] INFO OntologyLoader
      Finished loading
file:/Users/daviddominguez/Library/Mobile%20Documents/com~apple~CloudD
ocs/Tesis/FINAL%20Ontologia%20Sintomatologia%20Depresiva%20(Escrita).o
wl\
2021-04-11 20:27:50.279 [AWT-EventQueue-0] INFO OWLModelManagerImpl
      Loading for ontology and imports closure successfully completed in 980
ms\
2021-04-11 20:27:50.282 [AWT-EventQueue-0] INFO OWLModelManagerImpl
      \
2021-04-11 20:27:51.019 [AWT-EventQueue-0] INFO OWLEditorKit
      ----- Disposing of Workspace
      -----\
2021-04-11 20:27:51.046 [AWT-EventQueue-0] INFO TabbedWorkspace
      Saved tab state for 'OntoGraf' tab\
2021-04-11 20:27:51.048 [AWT-EventQueue-0] INFO TabbedWorkspace
      Saved tab state for 'Individuals by class' tab\
2021-04-11 20:27:51.050 [AWT-EventQueue-0] INFO TabbedWorkspace
      Saved tab state for 'DL Query' tab\

```



```

2021-04-11 20:27:51.052 [AWT-EventQueue-0] INFO TabbedWorkspace
    Saved tab state for 'Entities' tab\
2021-04-11 20:27:51.052 [AWT-EventQueue-0] INFO TabbedWorkspace
    Saved workspace\
2021-04-11 20:27:51.054 [AWT-EventQueue-0] INFO TabbedWorkspace
    Disposed of 'OntoGraf' tab\
2021-04-11 20:27:51.056 [AWT-EventQueue-0] INFO TabbedWorkspace
    Disposed of 'Individuals by class' tab\
2021-04-11 20:27:51.058 [AWT-EventQueue-0] INFO TabbedWorkspace
    Disposed of 'DL Query' tab\
2021-04-11 20:27:51.074 [AWT-EventQueue-0] INFO TabbedWorkspace
    Disposed of 'Entities' tab\
2021-04-11 20:27:51.075 [AWT-EventQueue-0] INFO TabbedWorkspace
    Disposed of workspace\
2021-04-11 20:27:51.082 [AWT-EventQueue-0] INFO OWLEditorKit \
    2021-04-11 20:35:50.206 [Classification Thread] INFO
OWLReasonerManager ----- Running Reasoner
    -----\
    2021-04-11 20:35:50.274 [Classification Thread] INFO
    OWLReasonerManager Pre-computing inferences:\
    2021-04-11 20:35:50.274 [Classification Thread] INFO
    OWLReasonerManager - class hierarchy\
    2021-04-11 20:35:50.275 [Classification Thread] INFO
    OWLReasonerManager - object property hierarchy\
    2021-04-11 20:35:50.275 [Classification Thread] INFO
    OWLReasonerManager - data property hierarchy\
    2021-04-11 20:35:50.275 [Classification Thread] INFO
    OWLReasonerManager - class assertions\
    2021-04-11 20:35:50.277 [Classification Thread] INFO
    OWLReasonerManager - object property assertions\
    2021-04-11 20:35:50.277 [Classification Thread] INFO
    OWLReasonerManager - same individuals\
    2021-04-11 20:35:50.397 [Classification Thread] INFO
    OWLReasonerManager Ontologies processed in 189 ms by FaCT++\
    2021-04-11 20:35:50.399 [Classification Thread] INFO
    OWLReasonerManager \
    }

```



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 8  
MÉTRICA DE VALIDACIÓN PARA ONTOLOGIA**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

		Métrica de validación de Ontología		Valoración	
		Ontology Development 101		1 = Regular	
		Protégé (versión Escritorio) v5.5.0		2 = Aceptable	
		Metodología		3 = Sobresaliente	
		Herramienta		4 = Excelente	
Métrica de evaluación para ontologías					
Sintoma	Categoría	Criterio evaluado	Especificación	Valoración	Observaciones
Generales	Suficiencia	Potencialidad para abstraer esencia	La información referida a las etapas de diseño e implementación		
	Metodología	Replicación de metodología	La metodología utilizada para el modelado es fácil de replicar en		
Estado de ánimo deprimido	Simplicidad	Entendimiento de la Ontología	La ontología de síntomas depresivos es entendible para el		
	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
Disminución importante del interés o el placer	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
Pensamientos de peso sin hacer dieta o aumento de peso, o disminución o ausencia del apetito	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
Insomnio o hipersomnia	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
Aglación o retraso psicomotor	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
Fatiga o pérdida de energía	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
Sentimiento de inutilidad o culpabilidad excesiva o inapropiada	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
Disminución de la capacidad para pensar o concentrarse	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
	Relevancia	Claridad de vinculación de las palabras	Se puede determinar la vinculación de las palabras		
Pensamientos de muerte recurrentes, ideas suicidas	Claridad	Definición de los síntomas	La definición de los síntomas es adecuada y correspondiente a		
	Coherencia	Numero de conceptos	El número de palabras identificadas por cada categoría es		
<b>Proyecto</b>					
Preguntas formuladas		Anotar las preguntas o inquietudes que surgen, al respecto de la			
Sugerencias		Sugerencias al desarrollo y aplicabilidad al respecto de la			
Oportunidades		Oportunidades encontradas para un mejor desarrollo y			
Firma		Nombre el profesional que participa en la validación de la ontología desarrollada.			



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 9  
REQUERIMIENTOS Y NECESIDADES DEL MÉTODO  
AUTOMÁTICO**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

## Anexo 9. Requerimientos y necesidades del método automático

### 1) Requerimiento número 1: Ejecución práctica del método automático

<b>PROYECTO</b>	Método automático para la identificación de sintomatología depresiva por medio de la narrativa escrita	
<b>Código:</b>	01	
Nombre Requerimiento:	Ejecución práctica del método automático.	
Responsable(s) Solicitud:	Profesionales del área de salud mental	
Responsable Funcional designado por el equipo de desarrollo:	Andres David Dominguez Rozo Cristian Andres Robayo Merchan	
<b>Descripción de la solicitud</b>		
<b>Usuario Solicitante</b>		
Es necesario que el método automático que será construido no tenga una accesibilidad compleja, por ejemplo, en los procesos de carga y en el análisis de las respuestas dadas por los consultantes; lo que fue expresado por los psicólogos de la siguiente manera: “El programa sea lo más claro y sencillo posible de ejecutar”.		
<b>Descripción del desarrollo estipulado</b>		
<p><b>Descripción:</b> Con el uso de figuras, flechas y diferentes imágenes descriptivas, se propone una clara descripción del paso a paso a seguir para la correcta ejecución del método automático desarrollado en el entorno de Google Colab.</p> <p>En el caso de que se presentara alguna confusión, se puede consultar la documentación de guía del usuario y videos de uso, que serán desarrollados y entregados conjuntamente con el método automático.</p>		
<b>Flujo para el desarrollo estipulado</b>		
<b>Flujo lógico</b>	<b>Paso 1:</b>	El usuario se encuentra en la cuenta de Google que tiene el método automático en los cuadernillos de Google Colab almacenado.
	<b>Paso 2:</b>	El usuario abre el entorno de Google Colab, seguido del <i>notebook</i> de desarrollo del método automático
	<b>Paso 3:</b>	El usuario ejecuta el boton de inicio de cada seccion de codigo que se encuentra al costado izquierdo señalado por la flecha de color rojo.
	<b>Paso 4:</b>	Internamente el <i>notebook</i> de desarrollo ejecuta

		operaciones como: carga de librerías, análisis del lenguaje, predicciones y generación de informe, según sea la sección correspondiente.
	<b>Paso 5:</b>	El usuario espera a que la carga se complete para cada celda de ejecución y que esta le permita cargar archivos o que esta celda descargue automáticamente del informe de análisis generados.
	<b>Paso 6:</b>	Repetiría el paso 3 una vez por cada sección de código, hasta que se finalice el notebook y se generación el informe de relación de conceptos.
<b>Flujo alternativo para el desarrollo estipulado</b>		
	<b>Paso 1:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar los videos guías de ejecución
	<b>Paso 2:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar la documentación “Manual de uso de Método automático para la identificación de síntomas depresivos”.

## 2) Requerimiento número 2: Integrar librerías específicas de análisis de sentimientos

<b>PROYECTO</b>	Método automático para la identificación de sintomatología depresiva por medio de la narrativa escrita
<b>Código:</b>	02
<b>Nombre Requerimiento:</b>	Integrar librerías específicas de análisis de sentimientos
<b>Responsable(s) Solicitud:</b>	Profesionales del área de salud mental
<b>Responsable Funcional designado por el equipo de desarrollo:</b>	Andres David Dominguez Rozo Cristian Andres Robayo Merchan
<b>Descripción de la solicitud</b>	
<b>Usuario Solicitante</b>	
Para los diferentes procesos que llevará a cabo el método automático es necesario que se enfoque en el lenguaje natural, para que así se identifique, mediante el análisis de las respuestas dadas por el consultante, una relación con el análisis de sentimientos.	
<b>Descripción del desarrollo estipulado</b>	
<b>Descripción:</b> Seleccionar e implementar librerías y funciones de análisis de texto y de lenguaje	

natural más adecuadas para cumplir con los requerimientos a el análisis solicitado por los expertos de la salud mental.

Flujo para el desarrollo estipulado		
<b>Flujo lógico:</b>	<b>Paso 1:</b>	El usuario se encuentra en la cuenta de Google que tiene el método automático en los cuadernillos de Google Colab almacenado.
	<b>Paso 2:</b>	El usuario abre el entorno de Google Colab, seguido del <i>notebook</i> de desarrollo del método automático.
	<b>Paso 3:</b>	El usuario ejecuta el botón de inicio de la sección de carga de librerías, esta sección es específica y será la primera en realizarse.
	<b>Paso 4:</b>	Internamente el <i>notebook</i> de desarrollo ejecuta las operaciones de carga de librerías internas del entorno, o descarga de internet las funciones específicas de análisis del lenguaje y sentimientos.
	<b>Paso 5:</b>	El usuario luego de que este proceso finalice puede seguir ejecutando las demás celdas de código que cumplen diversas funciones.
	<b>Paso 6:</b>	Internamente, en el método automático cada vez que una función requiera el uso de una librería hará el llamado correspondiente o la descarga momentáneamente para su uso.
Flujo alternativo para el desarrollo estipulado		
	<b>Paso 1:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar los videos guías de ejecución.
	<b>Paso 2:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar la documentación "Manual de uso de Método automático para la identificación de síntomas depresivos".

### 3) Requerimiento número 3: Identificar información del consultante mediante un código único (ID)

<b>PROYECTO</b>	Método automático para la identificación de sintomatología depresiva por medio de la narrativa escrita
<b>Código:</b>	03
<b>Nombre Requerimiento:</b>	Identificar información del consultante mediante un código único (ID)

Responsable(s) Solicitud:	Profesionales del área de salud mental	
Responsable Funcional designado por el equipo de desarrollo:	Andres David Dominguez Rozo Cristian Andres Robayo Merchan	
<b>Descripción de la solicitud</b>		
<b>Usuario Solicitante</b>		
Dado que este formulario de preguntas va a ser entregado y resuelto por más de un paciente, es necesario que esta información cuente con un código único (ID), que permita identificar tanto la información original del consultante como los resultados del método automático.		
<b>Descripción del desarrollo estipulado</b>		
<b>Descripción:</b> Generar un código único (ID) que permita la identificación de todas las respuestas dadas por un consultante específico, para que el experto de salud mental pueda individualizar el análisis generado por el método automático.		
<b>Flujo para el desarrollo estipulado</b>		
<b>Flujo lógico</b>	<b>Paso 1:</b>	El usuario se encuentra en la cuenta de Google que tiene el método automático en los cuadernillos de Google Colab almacenado.
	<b>Paso 2:</b>	El usuario abre el entorno de Google Colab, seguido del <i>notebook</i> de desarrollo del método automático.
	<b>Paso 3:</b>	El usuario ejecuta el botón correspondiente a la sección de carga del archivo de respuestas dadas por el consultante a las preguntas.
	<b>Paso 4:</b>	Internamente el archivo que almacena todas las respuestas dadas por los consultantes, genera automáticamente un código único (ID) específico para cada paciente y sus respuestas
	<b>Paso 5:</b>	Internamente el <i>notebook</i> de desarrollo correlaciona la hoja de cálculo con las respuestas al entorno de análisis del método automático, señalando como "llave" la columna de los códigos únicos.
	<b>Paso 6:</b>	El usuario digita el código único (ID) del consultante que desee tratar, este lo digita en el recuadro específico para esta información.
	<b>Paso 7:</b>	En adelante el <i>notebook</i> de desarrollo realiza toda la serie de análisis y demás clasificaciones de las respuestas generadas por el consultante pertenecientes específicamente a el ID señalado.



Flujo alternativo para el desarrollo estipulado	
	<p><b>Paso 1:</b> El usuario si llega a presentar inconvenientes o errores puede consultar los videos guías de ejecución.</p>
	<p><b>Paso 2:</b> El usuario si llega a presentar inconvenientes o errores puede consultar la documentación “Manual de uso de Método automático para la identificación de síntomas depresivos”.</p>

#### 4) Requerimiento número 4: Identificar fuentes de clasificación para la sintomatología depresiva

<b>PROYECTO</b>	Método automático para la identificación de sintomatología depresiva por medio de la narrativa escrita
<b>Código:</b>	04
<b>Nombre Requerimiento:</b>	Identificar fuentes de clasificación para la sintomatología depresiva
<b>Responsable(s) Solicitud:</b>	Profesionales del área de salud mental
<b>Responsable Funcional designado por el equipo de desarrollo:</b>	Andres David Dominguez Rozo Cristian Andres Robayo Merchan
Descripción de la solicitud	
<b>Usuario Solicitante</b>	
Se solicita que el informe generado por el método automático, en donde se especifican los análisis y clasificaciones respecto a las respuestas generadas por el consultante, estén fundamentadas en los nueve principales síntomas de la depresión. Para esto se propone una serie de reuniones entre psicólogos e ingenieros para que desarrollen e implementen una ontología de conceptos que será el referente teórico fundamental para las operaciones del método automático y de este trabajo.	
Descripción del desarrollo estipulado	
<b>Descripción:</b> La ontología que fue previamente desarrollada y validada por los expertos, debe ser la misma ontología con la cual el método automático operará y generará las clasificaciones pertinentes a las respuestas generadas por los consultantes en el formulario de preguntas.	
Flujo para el desarrollo estipulado	
<b>Flujo lógico</b>	<p><b>Paso 1:</b> El usuario se encuentra en la cuenta de Google que tiene el método automático en los cuadernillos de Google Colab almacenado.</p>
	<p><b>Paso 2:</b> El usuario abre el entorno de Google Colab,</p>

		seguido del <i>notebook</i> de desarrollo del método automático.
	<b>Paso 3:</b>	El usuario ejecuta con el botón de inicio todas las secciones correspondientes a la carga del archivo de respuestas, el análisis de texto y el análisis de lenguaje.
	<b>Paso 4:</b>	El usuario se encuentra en la sección de relacionar las respuestas obtenidas con la ontología de sintomatología depresiva construida.
	<b>Paso 5:</b>	El usuario ejecuta la sección de "Carga de ontología"
	<b>Paso 6:</b>	Internamente el <i>notebook</i> de desarrollo carga la ontología de sintomatología depresiva, está construida por profesionales del área de la salud mental y ingenieros de sistemas
	<b>Paso 7:</b>	El usuario ejecuta la sección de "Relación con la ontología"
	<b>Paso 8:</b>	Internamente el <i>notebook</i> de desarrollo ejecuta procesos de identificación de conceptos de las respuestas obtenidas con relación a los conceptos de la ontología pertenecientes a cada síntoma específico
	<b>Paso 9:</b>	El usuario ejecuta la sección de "Generar análisis de la relación con la ontología"
	<b>Paso 10:</b>	Internamente el <i>notebook</i> de desarrollo genera un archivo el cual describe las relaciones identificadas entre las respuestas generadas con el síntoma depresivo respectivo, además de incluir datos generales del paciente.
	<b>Paso 11:</b>	El informe generado, se descarga automáticamente para uso y entendimiento del usuario (profesional del área de la salud mental)
<b>Flujo alternativo para el desarrollo estipulado</b>		
	<b>Paso 1:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar los videos guías de ejecución.
	<b>Paso 2:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar la documentación "Manual de uso de Método automático para la identificación de síntomas depresivos".

## 5) Requerimiento número 5: Asegurar la privacidad de la información.

<b>PROYECTO</b>	Método automático para la identificación de sintomatología depresiva por medio de la narrativa escrita	
<b>Código:</b>	05	
<b>Nombre Requerimiento:</b>	Asegurar la privacidad de la información	
<b>Responsable(s) Solicitud:</b>	Profesionales del área de salud mental	
<b>Responsable Funcional designado por el equipo de desarrollo:</b>	Andres David Dominguez Rozo Cristian Andres Robayo Merchan	
<b>Descripción de la solicitud</b>		
<b>Usuario Solicitante</b>		
La información brindada por el usuario respecto a sus datos personales, sociodemográficos y las respuestas al formulario que busca identificar sintomatología depresiva, tiene que ser protegida por un entorno de seguridad que a su vez informe con anticipación que se realizará un análisis académico con la información que se brinde bajo leyes de protección y manejo de datos.		
<b>Descripción del desarrollo estipulado</b>		
<b>Descripción:</b> La información brindada por el usuario y el análisis de la misma por parte del método automático y los diferentes procesos de aprendizaje de máquina generarán un informe de resultados que se mantendrá bajo los esquemas de seguridad y de accesibilidad que brinda Google Drive, además de contar con el consentimiento asistido del mismo con la ley de Habeas Data.		
<b>Flujo para el desarrollo estipulado</b>		
<b>Flujo lógico</b>	<b>Paso 1:</b>	El consultante responde al formulario de preguntas las cuales buscan identificar sintomatología depresiva, esto se da solo si se acepta previamente la ley de <i>habeas data</i> para el manejo de sus datos.
	<b>Paso 2:</b>	Internamente <i>Google Forms</i> se encarga de almacenar la información dada por el usuario en una hoja de cálculo, a la cual únicamente se tendrá acceso desde la cuenta de Google donde se encuentre el método automático establecido.
	<b>Paso 3:</b>	El usuario ingresa en la cuenta de Google que tiene el método automático previamente establecido.
	<b>Paso 4:</b>	El usuario abre el entorno de Google Colab, seguido del <i>notebook</i> de desarrollo del método

		automático.
	<b>Paso 5:</b>	El usuario ejecuta todos los procedimientos correspondiente a la diferentes asignaciones del método automático
	<b>Paso 6:</b>	Internamente el <i>notebook</i> de desarrollo almacena información respectiva de los procesos en su carpeta de Google Drive
	<b>Paso 7:</b>	El usuario ejecuta la sección de generar el informe correspondientes a los análisis de respuestas dadas por el consultante
	<b>Paso 8:</b>	Internamente todas las herramientas involucradas en el proceso, almacenan su información en Google Drive la cual brinda la seguridad de información generada por Google y solamente se tendrá acceso para los profesionales del área de la salud involucrados en estas consultas.
<b>Flujo alternativo para el desarrollo estipulado</b>		
	<b>Paso 1:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar los videos guías de ejecución.
	<b>Paso 2:</b>	El usuario si llega a presentar inconvenientes o errores puede consultar la documentación “Manual de uso de Método automático para la identificación de síntomas depresivos”.



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 10  
PSEUDOCODIGOS**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

## PSEUDOCÓDIGOS PARA EL DISEÑO DEL MÉTODO AUTOMÁTICO

### 1. Proceso de limpieza de textos

1. PROCESO Limpieza\_informacion\_respuestas;
2. IMPORTAR (librerías de lectura de archivos Excel);
3. GENERAR (variables que señalan hoja de cálculo);
4. GENERAR (arreglos vacíos donde se almacena información extraída);
5. **INICIO CICLO REPETITIVO**
6.     ESCRIBIR (ciclo repetitivo que recorra la     hoja de cálculo);
7.     ASIGNAR(variable que almacena la información en cada iteración con función append);
8. **SALIR CICLO REPETITIVO**
9. REEMPLAZAR (Caracteres \_\_\_\_\_ al igual que espacios entre líneas);
10. COMPLETAR (Proceso de completación de palabras automáticamente);
11. ALMACENAR(Información resultante en nueva variable);
12. FIN

### 2. Proceso de implementación de algoritmo de lenguaje natural

1. PROCESO Implementacion\_lenguaje\_natural;
2. IMPORTAR (librería de lenguaje natural);
3. GENERAR (variables que llama funciones de la librería de lenguaje natural);
4. OPERAR (funciones de lenguaje natural con los arreglos de información limpia);
5. ALMACENAR (listas generadas con palabras claves seleccionadas por el algoritmo de lenguaje natural);
6. FIN

### 3. Correlación de palabras destacadas con ontología desarrollada

1. PROCESO Correlacion\_Palabras\_ontología
2. IMPORTAR (librería de lectura de ontologías);
3. CARGAR (documento de ontología);

4. GENERAR (variable que almacena información de la ontología)
5. IMPORTACIÓN (Del documento de ontología a la variable que almacena esta información, en tipo de arreglo);
6. **INICIO CICLO REPETITIVO**
7. DECLARACIÓN (Recorrer por medio de una variable arreglos con información de ontología y de las respuestas)
8. **INICIO CONDICIONAL**
9. VALIDACIÓN (COINCIDEN TÉRMINOS)
10. SI CUMPLE (Adicionar elemento coincidente en otro arreglo)
11. NO CUMPLE (Continuar con validación en las demás palabras)
12. **FIN CONDICIONAL**
13. **FIN CICLO REPETITIVO**
14. IMPRIMIR (Arreglo con palabras coincidentes entre respuestas y ontología)
15. FIN

#### 4. Proceso de exportación de información desde Google Drive

PROCESO Importar\_informacion\_HojasDeCalculo;

1. IMPORTAR (modulo files desde google.colab);
2. ESCRIBIR (funcion para importación);
3. Ejecutar
4. **Una vez cargue entorno**
5. OPRIMIR("Elegir archivo");
6. SELECCIONAR(Hoja de cálculo con respuestas a formulario);
7. SELECCIONAR ("cargar");
8. FIN

#### 5. Lectura de la información extraída en el entorno de Google Colab

1. PROCESO Lectura\_informacion\_GoogleColab;
2. IMPORTAR (librería pandas);
3. GENERAR (variable que almacene información);

4. ESCRIBIR (llamado a librería y función de leer archivos con formato "xlsx");
5. IMPRIMIR (variable que almacena información)
6. Ejecutar
7. FIN

## **6. Proceso de selección de palabras destacadas**

1. PROCESO Seleccion\_palabras\_destacadas;
2. GENERAR (variables para almacenar las palabras más destacadas de las respuestas);
3. OPERAR (funciones conteo de palabras y de selección de repeticiones por palabra);
4. ALMACENAR (listas generadas con palabras que presentan más repetición);
5. FIN

## **7. Generación de informe con el cual se relacionan información de las respuestas con la ontología de sintomatología depresiva**

1. PROCESO\_Generacion\_informe\_resultados
2. IMPORTAR (Librería para exportar información en documentos de Google);
3. RESULTADOS (Reunir arreglos en donde especifican las palabras y su referentes en la ontología);
4. ESCRIBIR (función para generar archivo .doc o .txt);
5. ASIGNAR (nombre para el documento de resultados);
6. EJECUTAR
7. FIN



## 8. Desarrollo del formulario



Para llevar a cabo el proceso de desarrollo del formulario de Google, en donde la respuesta a sus preguntas, puedan identificar si la persona posiblemente presente sintomatología depresiva, fue necesaria la utilización de diferentes documentaciones en línea, como también el uso de cuestionarios y manuales reglamentados por profesionales de la salud mental.

Los estudiantes que desarrollan este proyecto de grado, al igual que el formulario, fueron documentados por psicólogos de la Universidad Católica de Colombia, sobre qué cuestionarios y manuales tomar, como referente teórico al momento de desarrollar el formulario de Google. Además de compartir esta información, el formulario final fue validado por los mismos psicólogos, en donde se hicieron las debidas correcciones y diferentes afinidades a las preguntas.



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 11  
DISEÑO GENERAL MÉTODO AUTOMÁTICO**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**





**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 12  
PREGUNTAS PARA IDENTIFICAR TU ESTADO DE ÁNIMO  
(RESPUESTAS)**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

ID	Marca temporal	Acepta el tratamiento de sus datos bajo la ley de Habeas Data:	¿Cómo te llamas?	¿Donde vives?	¿Que edad tienes?
1	4/24/2021 12:31:10	Sí	prueba	prueba	prueba
2	4/24/2021 12:33:58	Sí	Prueba	Prueba	Prueba

**1. Piensa en algunas situaciones, en las dos últimas semanas, en las que te hayas sentido triste o desanimado(a) y descríbenos qué pensabas y qué sentías**

prueba

Prueba

**2. ¿En las dos últimas semanas te has sentido con menos energía (de lo normal) o con dificultades para concentrarte en las actividades diarias? Descríbenos más detalladamente... ¿En las últimas semanas te has sentido más acelerado o lento (de lo normal) al enfrentar las actividades diarias? Descríbenos más detalladamente...**

prueba

Prueba

**3. ¿Has tenido cambios en tus hábitos o rutina de alimentación? (has mantenido tus horarios habituales de alimentación, te saltas alguna comida o ha aumentado o disminuido tu apetito) Descríbenos cómo fueron esos cambios...**

prueba

Prueba

**4. ¿Qué tan satisfecho(a) y cómodo(a) te sentiste con la manera en la que enfrentaste las situaciones que fueron difíciles en las últimas dos semanas? ¿Las cosas resultaron como querías? Descríbenos la situación...**

prueba

Prueba

**5. ¿Has tenido deseos o ideas de morir en las últimas dos semanas? ¿O has pensado que no vale la pena vivir? Descríbenos al respecto...**

prueba

Prueba



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 13  
INFORME GENERADO**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

Este es el informe generado para el paciente con el ID : 8

El escribió estas palabras en sus respuestas: ['Si', 'Mal', 'La', 'Falta', 'De', 'Trabajo', 'Ha', 'Llevado', 'A', 'Que', 'No', 'Solo', 'La', 'Parte', 'Económica', 'Si', 'No', 'Emocional', 'Se', 'Vea', 'Afectada', 'Si', 'El', 'Desanimo', 'Hace', 'Que', 'No', 'Me', 'Den', 'Ganas', 'De', 'Hacer', 'Nada', 'No', 'Si', 'Ser', 'Invisible', 'Si', 'Su', 'Fortaleza', 'Todo', 'Me', 'Gusta', 'Si', 'Estabilidad', 'Económica', 'Felicidad', 'Si', 'Es', 'La', 'Sensación', 'De', 'No', 'Querer', 'Absolutamente', 'Y', 'Que', 'Nada', 'Te', 'Motive', 'Si', 'A', 'Pesar', 'De', 'Las', 'Dificultades', 'Mi', 'Fe', 'En', 'Dios', 'Es', 'Mas', 'Grande', 'Que', 'Ni', 'Debilidad', 'Al', 'Igual', 'Que', 'El', 'Amor', 'Por', 'Mi', 'Familia', 'No', 'Me', 'Gustan', 'Mis', 'Malas', 'Decisiones', 'Que', 'Ne', 'Llevan', 'A', 'Cometer', 'Muchos', 'Errores', 'Pensar', 'Mas', 'En', 'Los', 'Demas', 'Que', 'En', 'Mi', 'Misma', 'Mi', 'Cuerpo', 'Mi', 'Malgenio', 'Y', 'Ansiedad', 'Si', 'No', 'Me', 'Dan', 'Ganas', 'De', 'Cocinar', 'Ni', 'De', 'Comer', 'No', 'Me', 'Da', 'Hambre']

Al relacionar estas palabras con la Ontología encontramos que presenta correlación con:

La sintomatología de : ['Animo', 'Apatía', 'Alimentación', 'Descanso', 'Actividad\_fisica', 'Desaliento', 'Cargo\_de\_conciencia', 'Atencion\_dispersa', 'Ideas\_suicidas']

Estos son los conceptos identificados por rama especifica:

Animo: {'Solo', 'Errores'}

Apatia: {'Nada', 'Solo'}

Alimentación: {'Nada', 'Hambre'}

Descanso: {'Igual', 'Debilidad'}

Actividad\_física: {'Nada', 'Debilidad'}

Desaliento: {'Igual', 'Solo', 'Errores', 'Debilidad'}

Cargo\_de\_conciencia: {'Solo', 'Errores'}

Atencion\_dispersa: {'Errores'}

Ideas\_suicidas: {'Solo'}

Si anteriormente no se ve ninguna relación es porque no se encontraron correlaciones.

Las palabras mas significativas dadas por el consultante son:

['Mal La Falta De Trabajo', 'Solo', 'La Parte Económica', 'Emocional', 'Se', 'Vea Afectada', 'El Desanimo', 'Me', 'Den Ganas', 'Hacer', 'Nada', 'Invisible', 'Su Fortaleza', 'Todo', 'Me', 'Gusta Si Estabilidad Económica Felicidad', 'La Sensación De No', 'Querer Absolutamente', 'Nada Te Motive Si A Pesar De Las Dificultades Mi Fe En Dios Es Mas', 'Debilidad Al Igual Que El Amor Por Mi Familia', 'Me', 'Gustan Mis Malas Decisiones Que Ne Llevan A Cometer Muchos Errores Pensar Mas En Los Demas Que En Mi Misma Mi Cuerpo Mi Malgenio', 'Ansiedad', 'Me', 'Dan Ganas De Cocinar', 'Comer', 'Me', 'Hambre']

Los adjetivos que este utilizo fueron:

['Grande']

Los verbos que este utilizo fueron:

['Llevado', 'Hace', 'Da']

Las palabras mas usadas por el consultante son:

[('Si', 9), ('No', 8), ('Que', 7), ('De', 6), ('Me', 5), ('Mi', 5), ('La', 3), ('A', 3), ('En', 3), ('Economica', 2), ('El', 2), ('Ganas', 2), ('Nada', 2), ('Es', 2), ('Y', 2), ('Mas', 2), ('Ni', 2), ('Mal', 1), ('Falta', 1), ('Trabajo', 1)]

Si en alguno de los anteriores, te genero vacío, es porque el consultante no utilizo ningún concepto relacionado a esa área especifica.





**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

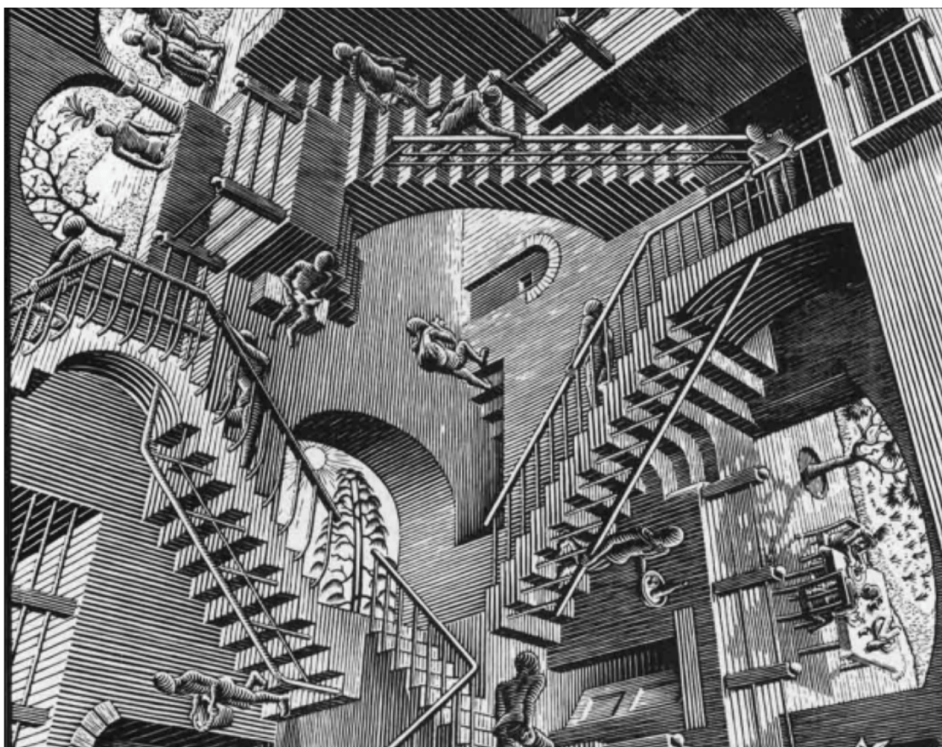
**ANEXO 14  
MANUAL DE USO PARA EL MÉTODO AUTOMÁTICO EN  
APOYO A LA IDENTIFICACIÓN DE SINTOMATOLOGÍA  
DEPRESIVA**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

---

## MÉTODO AUTOMÁTICO PARA EL APOYO EN LA IDENTIFICACIÓN DE SINTOMATOLOGÍA DEPRESIVA



## Manual de uso

Desarrollado por: David Dominguez; Cristian Robayo

Para: Apoyo en las consultas desarrolladas por profesionales del área de salud mental

17 de mayo de 2021

Version: 0001

## RESUMEN GENERAL

### Objetivo

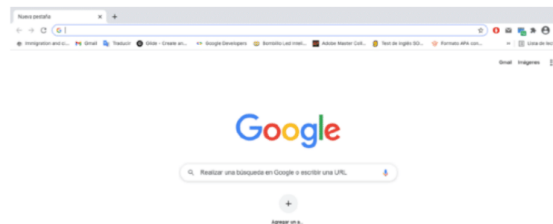
Este manual de uso está enfocado a describir el buen uso del método automático que apoya en la identificación de sintomatología depresiva, que puede sufrir una persona, por medio de las narrativas escritas que el mismo género.

### Metas

Para cumplir con esta meta se llevo a cabo un desarrollo en el entorno de Google Colab, el cual permite de una manera automática facilitar para los profesionales del área de Salud Mental procesos como lo son: carga y recopilación de información, procesos de transformación y entendimiento de información y la generación de un informe final, el cual consolida los aprendizajes obtenidos la correlación entre la información obtenida del consultante con los conceptos estipulados en una ontología de palabras, que apuntan a los posibles términos que escriba una persona que padece de algún síntoma de depresión.

### Desarrollo

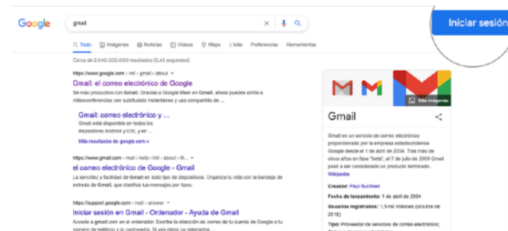
1. Abrimos nuestro navegador de preferencia, ya sea Google Chrome, Opera, Safari, Microsoft Edge, entre otros.



2. Iniciamos sesión de gmail, con las siguientes credenciales:

**Correo:**  
[metodoautomaticodepresion@gmail.com](mailto:metodoautomaticodepresion@gmail.com)

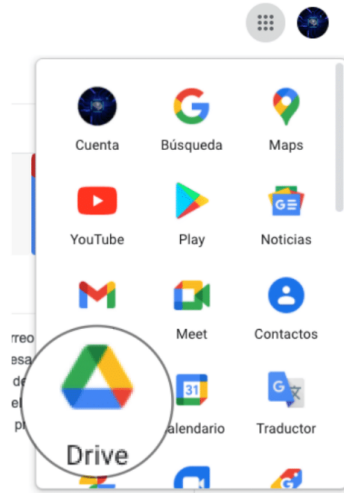
**Contraseña:** Metodoautomaticodepresion1



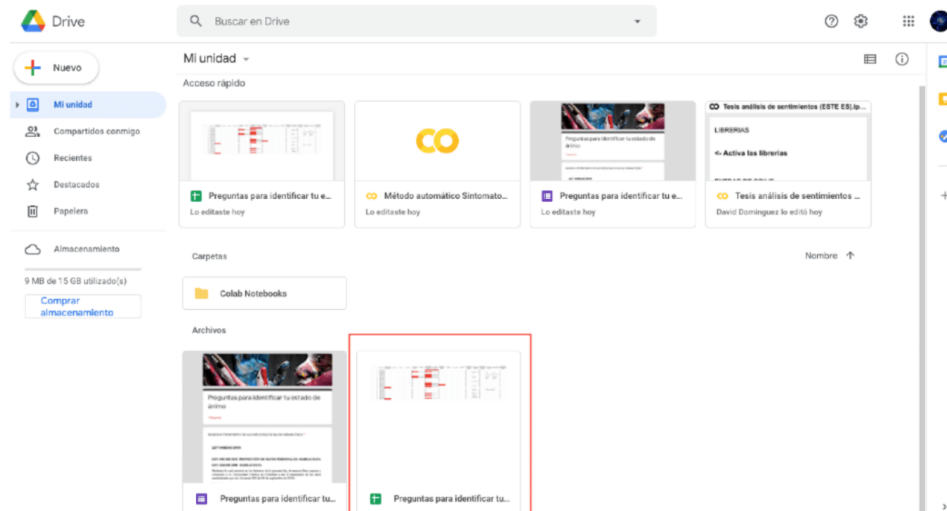
3. En este correo, que será único para este proyecto y herramienta. Dado esta cuenta almacena únicamente:

- El metodo automatico en Google Colab
- El formulario de preguntas
- La hoja de calculo que almacena las respuestas
- Video de uso y ejemplos reales
- Manual de uso del metodo automatico

Ahora, nos iremos en la sección de aplicaciones y herramientas a Google Drive.



4. Una vez estemos en el drive, identificaremos la hoja de calculo de Google, el cual almacena las respuestas de los consultantes:



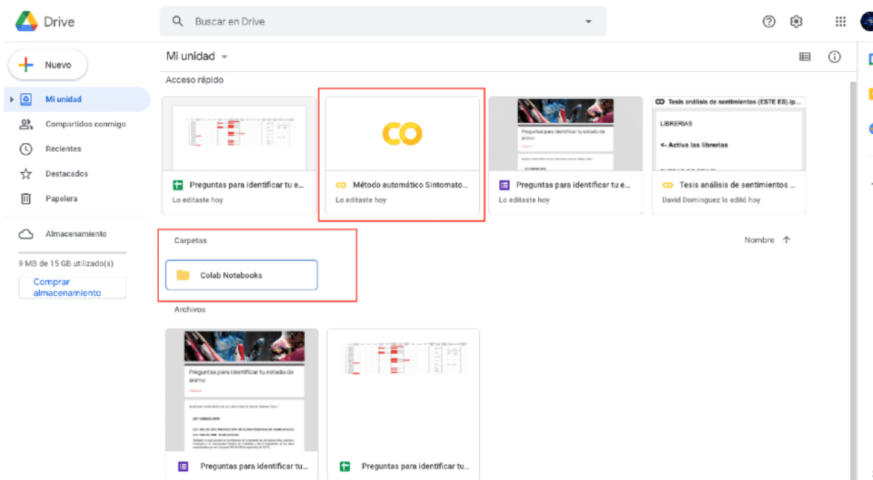
5. Daremos clic en el archivo, para abrirlo y así validar que se encuentre almacenada la ultima respuesta dada por él consultare tratado, con su código único ID y demás información del mismo.

ID	Fecha temporal	¿Acepta el tratamiento de sus datos bajo la ley de Habeas Data?	¿Cómo te llamas?	¿Dónde vives?	¿Que edad tienes?	¿Cómo has estado sentir sus últimos semanas?	¿Cuántas veces lo que ha pasado?	¿Sientes que algo nuevo ha afectado cómo se sienten (últimos dos meses)?
1	13/05/21 20:30:51 S					Bien		No
2	13/05/21 20:42:28 S					Mal	Mucho tiempo en mal y	Si
3	13/05/21 20:44:08 S					Bien		No
4	13/05/21 20:40:34 S					Bien		No
5	13/05/21 20:50:26 S					Bien		No
6	13/05/21 20:41:31 S					Bien		No
7	13/05/21 20:54:42 S					Bien		No
8	13/05/21 20:56:24 S					Mal	La falta de trabajo ha b	Si
9	13/05/21 20:56:36 S					Mal	La falta de trabajo ha b	Si
10	13/05/21 20:57:01 S					Bien		No
11	13/05/21 21:16:58 S					Bien		No
12	13/05/21 21:18:28 S					Bien		No
13	13/05/21 21:19:27 S					Bien		No
14	13/05/21 21:24:57 S					Bien		No
15	13/05/21 21:27:25 S					Bien		No
16	13/05/21 21:32:22 S					Bien		Si
17	13/05/21 21:32:36 S					Bien		No
18	13/05/21 21:36:46 S					Bien		No
19	13/05/21 21:37:50 S					Bien		Si
20	13/05/21 21:34:14 S					Bien		No
21	14/02/21 9:34:23 S					Bien		No
22	14/02/21 9:38:27 S					Bien		No
23	14/02/21 9:38:64 S					Bien		No
24	14/02/21 9:41:41 S					Bien		No

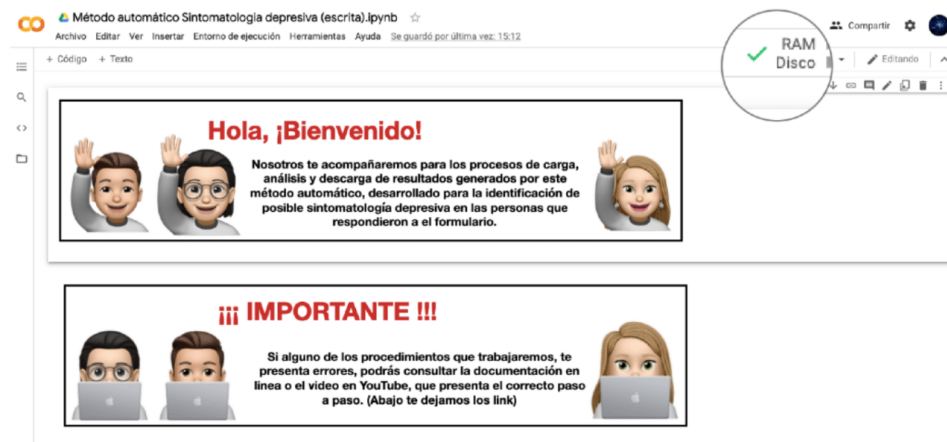
**NOTA:** Veras que información relevante como que no acepte el manejo de datos o que de primer momento diga que se siente mal, esta señalada con el color rojo. Este proceso es importante para verificar que la información del consultante se envió y almaceno correctamente al igual que para identificar cuál es el ID que se le asignaron a estas respuestas generadas.

6. Descargaremos esta hoja de calculo cómo archivo Excel (.xlsx), para cargarlo luego a nuestro método automático.

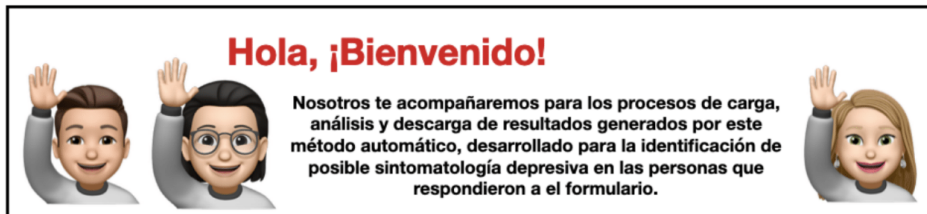
7. Volvemos a los archivos de Google Drive, en donde podemos buscar en la carpeta "Colab Notebooks" o el archivo directamente llamado "Método automático Sintomatología depresiva (escrita).ipynb", y lo abrimos.



8. Esperamos que carga el método automático, hasta ver un " ✓ " de color verde, esto indicara que ya puede empezar a ejecutar este cuadernillo.



9. Encontraras a nuestros colaboradores, que te irán indicando el paso a paso a seguir para llevar un buen proceso y ejecución del método automático.



10. Para cada sección que se trabaje en este método automático estará señalado con una flecha amarilla y un botón de play:



Todas las acciones de ejecución van reglamentadas de la siguiente manera:



•Das click, en ejecutar.



•Esperas a que el procedimiento cargue y si es necesario ingresas la información que se solicitara en los cuadros de texto.



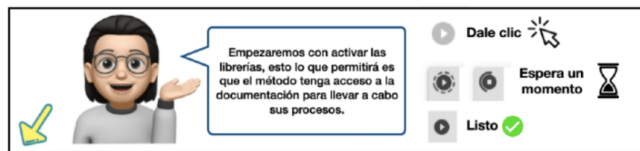
•Una vez finalizado este procedimiento el botón de play, se tornada color negro y eso dará a saber qué se ha ejecutado correctamente.

11. Todos los procesos trabajados en este método automático serán explicados desde el entorno de Google Colab, con nuestros colaboradores y globos de texto. Aquí se exponen brevemente:

### 11.1 Carga de librerías

#### ▼ COMENCEMOS CON LAS LIBRERIAS

---



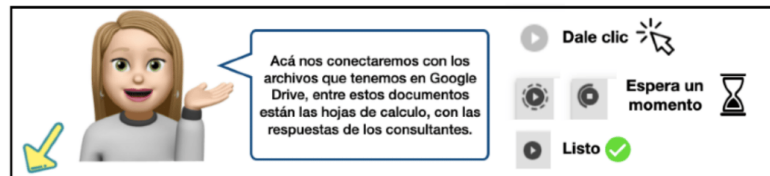
[46] <- Activa las librerias

```
Requirement already satisfied: deplacy in /usr/local/lib/python3.7/dist-packages (1.9.3)
Requirement already satisfied: es_core_news_sm==2.2.5 from https://github.com/explosion/spacy-models/releases/download/es\_core\_news\_sm-2.2.5/es\_core\_news\_sm-2.2.5
Requirement already satisfied: spacy==2.2.2 in /usr/local/lib/python3.7/dist-packages (from es_core_news_sm==2.2.5) (2.2.4)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-packages (from spacy==2.2.2->es_core_news_sm==2.2.5) (1.19.5)
```

11.2 Conexión de información con Google Drive, esta cuenta tiene los permisos validados de conexión previamente establecidos.

#### ▼ EXTRAER INFORMACION DESDE DRIVE

---



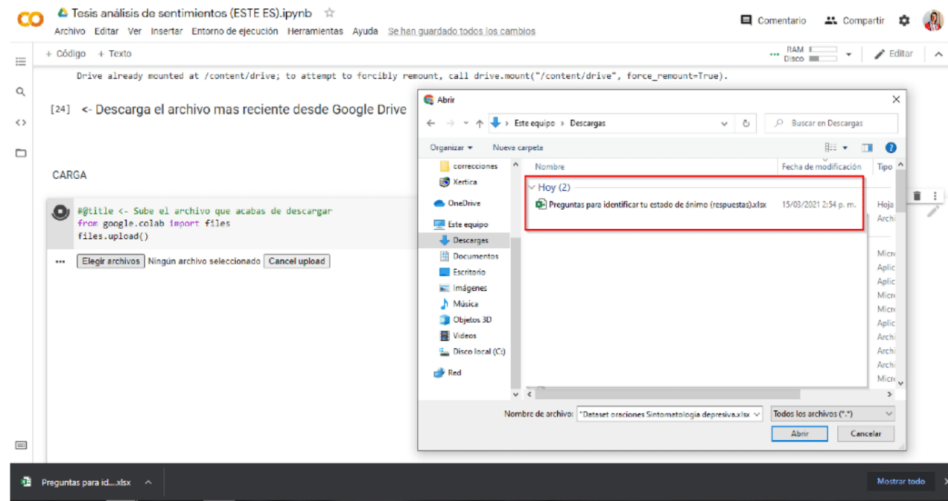
[47] <- Conecta con Drive

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

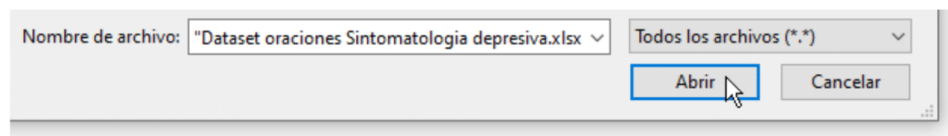




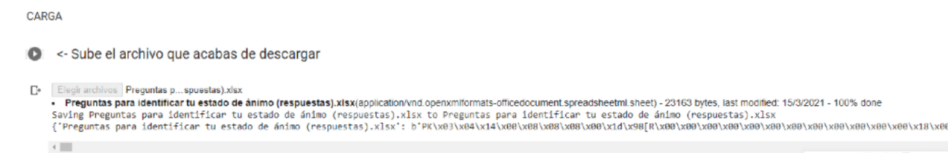
Se abre la ventana de visualizador de archivos y se selecciona de la carpeta descargas el archivo más reciente, que es el descargado en el paso 6, este debe llamarse: 'Preguntas para identificar tu estado de ánimo (respuestas).xlsx'



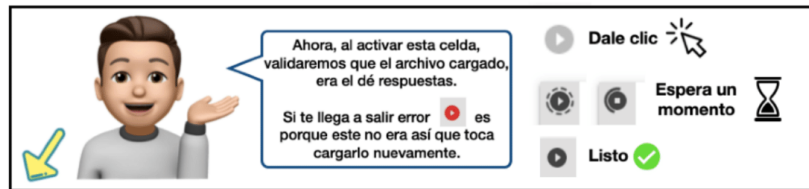
Se selecciona este archivo y se da en "Abrir".



Resultado obtenido:



Para validar que el archivo que anteriormente cargado sea el correcto y procesar los procesos de análisis de respuestas, es obligatorio ejecutar la siguiente sección, la cual valida que el archivo cargado sea el necesario para que el método automático pueda empezar a operar el análisis a las respuestas obtenidas.



▶ <- Valida que el archivo que cargaste se llama "Preguntas para identificar tu estado de ánimo (respuestas).xlsx"

Si la carga del archivo correspondiente fue correcta se genera la siguiente respuesta:

```
▶ <- Valida que el archivo que cargaste se llama "Preguntas para identificar tu estado de ánimo (respuestas).xlsx"
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:4: DeprecationWarning: call to deprecated function get_sheet_by_name (Use wb[sheetname]).
after removing the cud from sys.path.
```

Si la carga del archivo correspondiente fue incorrecta se genera la siguiente respuesta:

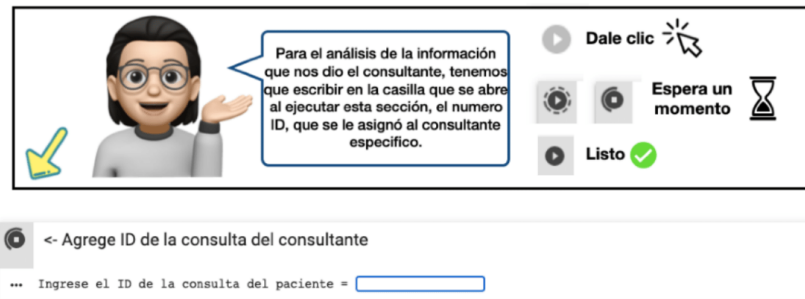
```
▶ <- Valida que el archivo que cargaste se llama "Preguntas para identificar tu estado de ánimo (respuestas).xlsx"
-----
FileNotFoundError                               Traceback (most recent call last)
ipython_input_31_39f125d8a22f in <module>()
----> 1 #!title <- valida que el archivo que cargaste se llama "Preguntas para identificar tu estado de ánimo (respuestas).xlsx"
      2 import openpyxl
----> 3 doc = openpyxl.load_workbook ("Preguntas para identificar tu estado de ánimo (respuestas).xlsx")
      4 hoja = doc.get_sheet_by_name("Respuestas de formulario 1")
      5 preguntas = hoja.delete_rows(1)

~/usr/lib/python3.7/site-packages/openpyxl/workbook.py in _init_(self, file, mode, compression, allowZip64, compresslevel)
    1238 while True:
    1239     try:
-> 1240         self.fp = io.open(file, filemode)
    1241     except OSError:
    1242         if filemode in modeDict:
FileNotFoundError: [Errno 2] No such file or directory: 'Preguntas para identificar tu estado de ánimo (respuestas).xlsx'
SEARCH STACK OVERFLOW
```

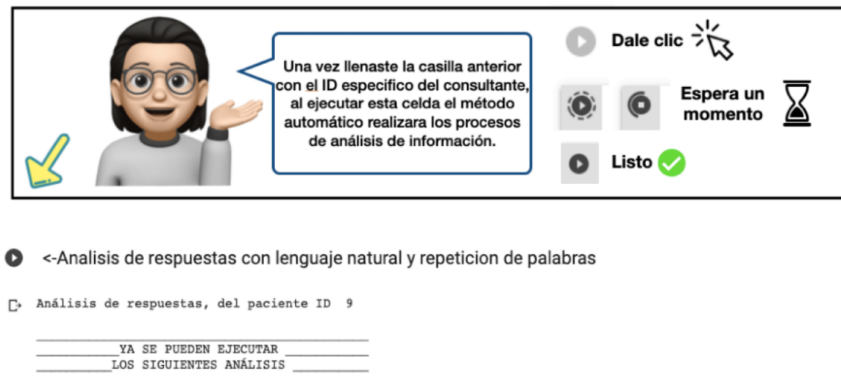
Esta celda de validación nos permite verificar y enviar el archivo específico de respuestas al proceso de análisis del método automático, es importante que esta validación sea correcta para proseguir.

11.4 Analizaremos la respuestas del consultante específico determinando su ID, en el cuadro de texto que se abre al ejecutar.

▼ ANALISIS A EL CONSULTANTE



11.5 Esta sección ejecutara todos los análisis de lenguaje natural y entendimiento computacional de las respuestas, esta sección como resultado te muestra el ID del consultante analizado y a su vez muestra también un mensaje que notifica que ya se pueden ejecutar los siguientes análisis y que la información ya fue analizada y sin problemas.

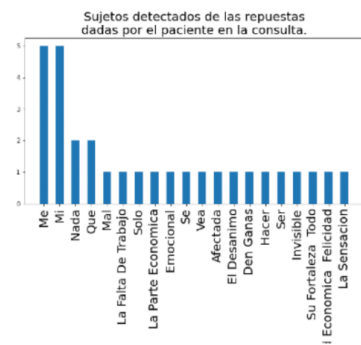


11.6 Las secciones de análisis generan gráficas en las cuales puedes evidenciar:

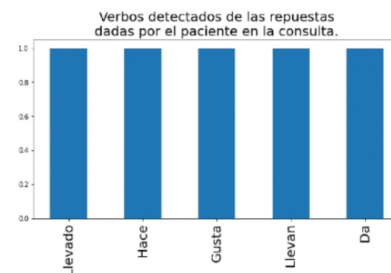
- Cuales fueron los sujetos de las respuestas mas destacados
- Cuales fueron los adjetivos de las respuestas obtenidas
- Cuales fueron los verbos de las respuestas obtenidas
- Cuales fueron las palabras mas utilizadas por el consultante en las respuestas
- Cuantas veces se menciona el concepto de depresión

**Nota:** Si alguna de estas secciones te genera un error, no es relevante, lo que ocurre es que entre las respuestas dadas por el consultante no se encontró una relación con este tipo de conceptos.

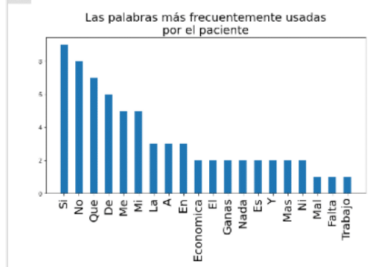
[55] <- Sujetos detectados de las repuestas dadas por el consultante.



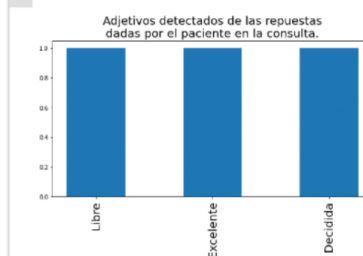
[57] <- Verbos detectados de las repuestas dadas por el consultante.

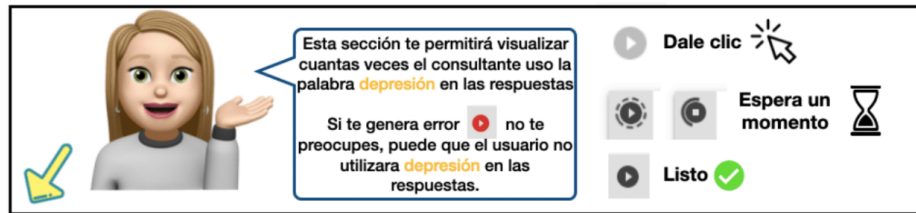


6 <- Palabras mas usadas por el consultante.



6 <- Adjetivos detectados de las repuestas dadas por el consultante.

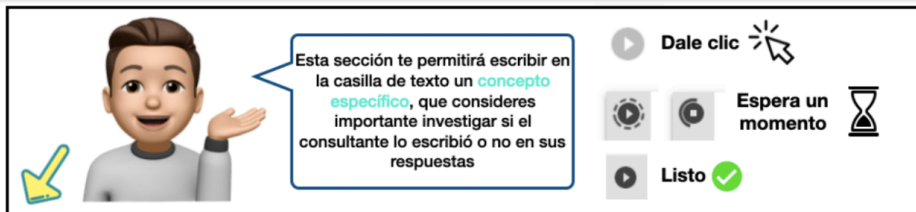




◀- Cuantas veces se encontro el concepto de Depresion:

☐ El concepto de 'Depresion' se encontro 1 vez en las respuestas

11.7 Para finalizar esta sección de análisis, se genera una celda en la cual al ejecutar tienes un cuadro de texto que te permite con el concepto que este escribas, darte a saber cuantas veces este concepto se escribió en las respuestas dadas por el consultante.



⊞ -> Veces se encontro el concepto de:

... Que concepto desea buscar su frecuencia de uso =

▶ -> Veces se encontro el concepto de:

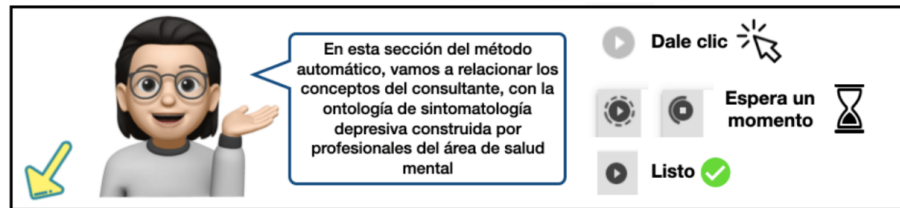
Que concepto desea buscar su frecuencia de uso = bien  
El concepto de bien se encuentra en las respuestas dada por el usuario, 2 veces.

12. Para la relación de conceptos con la Ontología, llevamos a cabo dos procesos:

12.1 Cargamos la Ontología desarrollada con esta celda

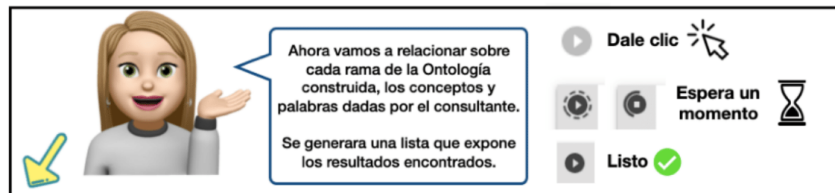
▼ RELACION CON LA ONTOLOGÍA

---



▶ <- Carga la Ontología

12.2 Relacionamos los conceptos dados por el consultante con los términos establecidos en la ontología de sintomatología depresiva.



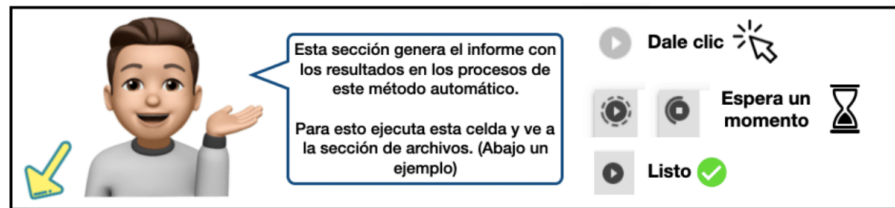
▶ <- Relaciona la Ontología de sintomatología depresiva, con las diferentes respuestas generadas por el paciente.

---

13. Generamos el informe de resultados, para este realizamos el siguiente proceso:

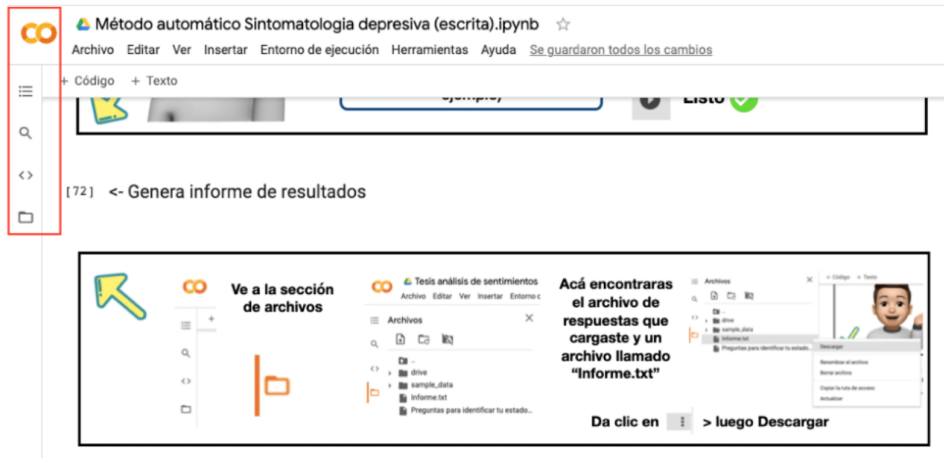
13.1 Ejecutamos la celda que realiza el proceso de crear el archivo y cargarlo con los datos relevantes.

▼ INFORME DE RESULTADOS



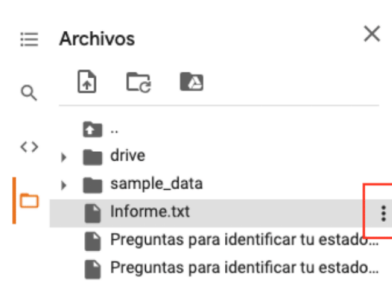
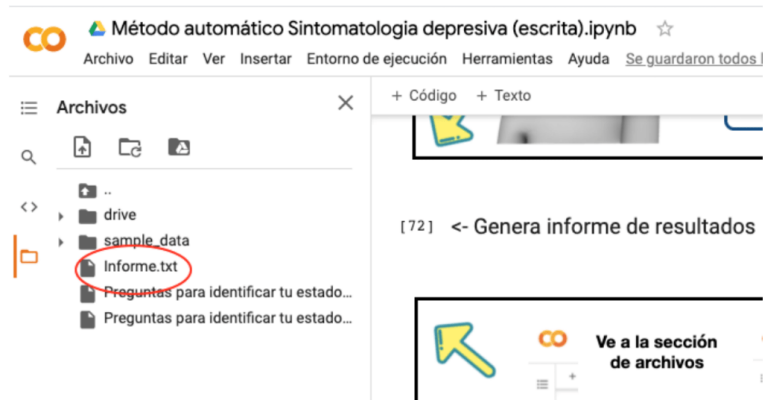
▶ <- Genera informe de resultados

13.2 Vamos a la sección de archivos de Google Colab



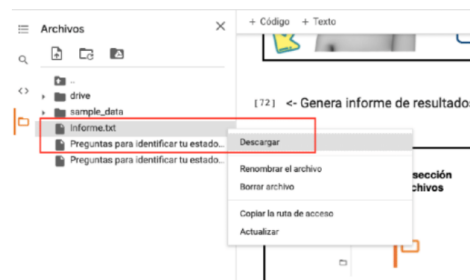


13.3 En el icono de una carpeta abrimos y ahí encontramos los archivos que utilizamos y generamos, como lo es el de informe.txt que es el que necesitamos descargar.



13.3 Damos clic en los tres puntos, que aparecen al lado izquierdo del archivo.

13.4 Damos en la opción de descargar.



13.5 El archivo se descarga y podemos trabajarlo en nuestra computadora.

```
Informe.txt
Este es el informe generado para el paciente con el ID : 9

El escribí estas palabras en sus respuestas: ['Si', 'Mal', 'La', 'Falta', 'De', 'Trabajo', 'Ha',
'Llevado', 'A', 'Que', 'No', 'Solo', 'La', 'Parte', 'Económica', 'Si', 'No', 'Emocional', 'Se',
'Vea', 'Afectada', 'Si', 'El', 'Desanimo', 'Hace', 'Que', 'No', 'Me', 'Den', 'Ganas', 'De', 'Hacer',
'Nada', 'No', 'Si', 'Ser', 'Invisible', 'Si', 'Su', 'Fortaleza', 'Todo', 'Me', 'Gusta', 'Si',
'Estabilidad', 'Económica', 'Felicidad', 'Si', 'Es', 'La', 'Sensacion', 'De', 'No', 'Querer',
'Absolutamente', 'Y', 'Que', 'Nada', 'Te', 'Motive', 'Si', 'A', 'Pesar', 'De', 'Las', 'Dificultades',
'Mi', 'Fe', 'En', 'Dios', 'Es', 'Mas', 'Grande', 'Que', 'Ni', 'Debilidad', 'Al', 'Igual', 'Que',
'El', 'Amor', 'Por', 'Mi', 'Familia', 'No', 'Me', 'Gustan', 'Mis', 'Malas', 'Decisiones', 'Que',
'Ne', 'Llevan', 'A', 'Cometer', 'Muchos', 'Errores', 'Pensar', 'Mas', 'En', 'Los', 'Demas', 'Que',
'En', 'Mi', 'Misma', 'Mi', 'Cuerpo', 'Mi', 'Malgenio', 'Y', 'Ansiedad', 'Si', 'No', 'Me', 'Dan',
'Ganas', 'De', 'Cocinar', 'Ni', 'De', 'Comer', 'No', 'Me', 'Da', 'Hambre']

Al relacionar estas palabras con la Ontología encontramos que presenta correlación con:
La sintomatología de : []
Si anteriormente no se ve ninguna relación es porque no se encontraron correlaciones.

Las palabras mas significativas dadas por el consultante son:
['Mal', 'La Falta De Trabajo', 'Solo', 'La Parte Económica', 'Emocional', 'Se', 'Vea', 'Afectada',
'El Desanimo', 'Me', 'Den Ganas', 'Hacer', 'Nada', 'Ser', 'Invisible', 'Su Fortaleza Todo', 'Me',
'Estabilidad Económica Felicidad', 'La Sensacion', 'Querer', 'Absolutamente', 'Nada', 'Te',
'Motive', 'Pesar', 'Las Dificultades', 'Mi', 'Mi Fe', 'Dios', 'Es', 'Mas Grande', 'Debilidad',
'Igual', 'El Amor', 'Mi', 'Mi Familia', 'Me', 'Gustan', 'Mis', 'Mis Malas Decisiones', 'Que', 'Ne',
'Cometer', 'Muchos', 'Muchos Errores', 'Pensar Mas', 'Los Demas', 'Que', 'Mi', 'Mi Misma', 'Mi',
'Cuerpo', 'Mi', 'Malgenio', 'Ansiedad', 'Me', 'Dan Ganas', 'Cocinar', 'Comer', 'Me', 'Hambre']

Los adjetivos que este utilizo fueron:
['Grande']

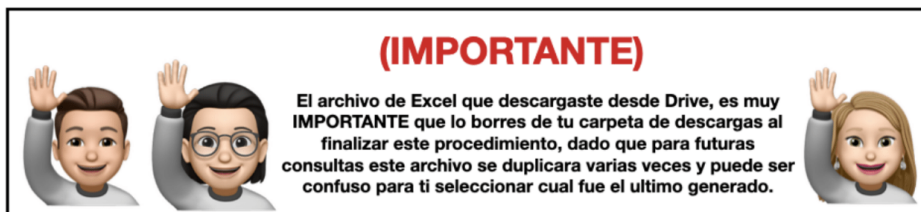
Los verbos que este utilizo fueron:
['Llevado', 'Hace', 'Gusta', 'Llevan', 'Da']

Las palabras mas usadas por el consultante son:
[('Si', 9), ('No', 8), ('Que', 7), ('De', 6), ('Me', 5), ('Mi', 5), ('La', 3), ('A', 3), ('En', 3),
('Económica', 2), ('El', 2), ('Ganas', 2), ('Nada', 2), ('Es', 2), ('Y', 2), ('Mas', 2), ('Ni', 2),
('Mal', 1), ('Falta', 1), ('Trabajo', 1)]

Si en alguno de los anteriores, te genero vacío, es porque el consultante no utilizo ningún concepto
relacionado a esa área específica.
```

**NOTA:** El anterior resultado es trabajado con una prueba que fue realizada no es una consulta real.

## NOTAS:



Para mayor información pueden consultar el siguiente video tutorial:

Link: ( )



**UNIVERSIDAD CATÓLICA**  
de Colombia

**MÉTODO AUTOMÁTICO PARA EL APOYO EN LA  
IDENTIFICACIÓN DE SÍNTOMAS DE DEPRESIÓN A  
TRAVÉS DE NARRATIVA ESCRITA**

**ANEXO 15  
MÉTODO AUTOMÁTICO SINTOMATOLOGÍA DEPRESIVA  
(ESCRITA).**

**ANDRÉS DAVID DOMÍNGUEZ ROZO  
CRISTIAN ANDRÉS ROBAYO MERCHAN**

**UNIVERSIDAD CATÓLICA DE COLOMBIA  
FACULTAD DE INGENIERÍA  
PROGRAMA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN  
BOGOTÁ  
2021**

```
#@title <- Activa las librerías
import spacy
import requests
from bs4 import BeautifulSoup
```

<- Activa las librerías

```
!pip install deplacy
!python -m spacy download es_core_news_sm
import pkg_resources,imp
imp.reload(pkg_resources)
import spacy
nlp=spacy.load("es_core_news_sm")
from collections import Counter
#!pip install textacy
#import textacy

import numpy as np
import matplotlib.pyplot as plt
!pip install monkeylearn
```

```
#@title <- Conecta con Drive
from google.colab import drive
drive.mount('/content/drive')
```

<- Conecta con Drive

Mounted at /content/drive

```
#@title <- Sube el archivo descargaste de Google Drive (Respuestas dadas)
from google.colab import files
files.upload()
```

<- Sube el archivo descargaste de Google Drive (Respuestas dadas)

```
#@title <- Valida que el archivo que cargaste se llama "Preguntas para identificar tu es
import openpyxl
doc = openpyxl.load_workbook("Preguntas para identificar tu estado de ánimo (Respuestas
hoja = doc.get_sheet_by_name("Respuestas de formulario 1")
preguntas = hoja.delete_rows(1)
```

<- Valida que el archivo que cargaste se llama "Preguntas para identificar tu estado de ánimo (respuestas).xlsx"

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:4: DeprecationWarning: Call to deprecated function get_sheet_by_name (Use wb[sheetname]).
after removing the cwd from sys.path.
```

```
#@title <- Agregue ID de la consulta del consultante
id = int(input("Ingrese el ID de la consulta del paciente = "))
```

<- Agregue ID de la consulta del consultante

Ingrese el ID de la consulta del paciente = 3

```

#@title <- Analisis de respuestas con lenguaje natural y repeticion de palabras
print("Análisis de respuestas, del paciente ID ",id,".")
#####
import spacy
import unicodedata
from spacy.lang.es.examples import sentences
import unicodedata
#Inp-spacy.load("es_core_news_sm")
#####
#region
respuestas=[]
palabras=[]
respuesta=None
##### Información correspondiente al usuario desde hoja de calculo #####
celdaInicial = "C(8)"
celdaFinalF = celdaInicial.format(id)
celdaFinal = "M(8)"
celdaFinalF = celdaFinal.format(id)

multiple_cells = hoja[celdaInicialF:celdaFinalF]
for row in multiple_cells:
    for cell in row:
        respuesta=None.append(cell.value)
##### Limpio de texto #####
respuestas = list(filter(None, respuesta))
Respuestascompletas = " ".join(respuestas)

output_clean = Respuestascompletas.replace('\n', ' ')
output_clean2 = output_clean.replace(' ', '')
output_clean3 = output_clean2.replace('.', ' ')
trans_tab = dict.fromkeys(range(ord, u'\u0001\u0008'), None)
output_clean4 = unicodedata.normalize('NFKC', unicodedata.normalize('NFKD', output_clean3).translate(trans_tab))
output_clean5 = output_clean4.title()
doc = nlp(output_clean5)
palabras = output_clean5.split()
##### Conteo de palabras #####
contador = Counter(palabras)
buscar = 'depression'
import re
reges="(".*?)"
listaCadena = " ".join(palabras)
encontrados = re.findall(reges, listaCadena, re.MULTILINE)
cantIdad=len([matchNum for matchNum, match in enumerate(encontrados, start=1)])
print("
#####
VA SE PUEDEN EJECUTAR
#####
LOS SIGUIENTES ANALISIS
#####

/usr/local/lib/python3.7/dist-packages/catalogue.py:138: DeprecationWarning: SelectableGroups dict interface is deprecated. Use select.
for entry_point in AVAILABLE_ENTRY_POINTS.get(self.entry_point_namespace, []):

```

<- Analisis de respuestas con lenguaje natural y repeticion de palabras

```

#@title <- Sujetos detectados de las repuestas dadas por el consultante.
#####SUJETOS#####
fig = plt.figure(figsize=(10,5))
ax = fig.gca()

Esencial = [chunk.text for chunk in doc.noun_chunks]
contador = Counter(Esencial)

labels, values = zip(*contador.most_common(20))

indexes = np.arange(len(labels))
width = 1

ax.bar(indexes, values, 0.5)
ax.set_xticks(indexes)
ax.set_xticklabels( labels, rotation=90, fontsize=20)
ax.set_title("Sujetos detectados de las repuestas dadas por el paciente en la consulta.", fontsize=20)
plt.show()
#####

```

<- Sujetos detectados de las repuestas dadas por el consultante.

```

#@title <- Adjetivos detectados de las repuestas dadas por el consultante.
#####Adjetivos#####
fig = plt.figure(figsize=(10,5))
ax = fig.gca()

ADJ_F = []

for token in doc:
    if token.pos_ == "ADJ":
        ADJ_F.append(token.text)

#print("Adjetivos = ",ADJ_F)
contador = Counter(ADJ_F)

labels, values = zip(*contador.most_common(20))

indexes = np.arange(len(labels))
width = 1

ax.bar(indexes, values, 0.5)
ax.set_xticks(indexes)
ax.set_xticklabels( labels, rotation=90, fontsize=20)
ax.set_title("Adjetivos detectados de las repuestas dadas por el paciente en la consulta.", fontsize=20)
plt.show()
#####

```

<- Adjetivos detectados de las repuestas dadas por el consultante.

```

#@title <- Verbos detectados de las repuestas dadas por el consultante.
#####Verbos#####
fig = plt.figure(figsize=(10,5))
ax = fig.gca()

VERB_F = []

for token in doc:
    if token.pos_ == "VERB":
        VERB_F.append(token.text)

#print("Verbos = ",VERB_F)
contador = Counter(VERB_F)

labels, values = zip(*contador.most_common(20))

indexes = np.arange(len(labels))
width = 1

ax.bar(indexes, values, 0.5)
ax.set_xticks(indexes)
ax.set_xticklabels( labels, rotation=90, fontsize=20)
ax.set_title("Verbos detectados de las repuestas dadas por el paciente en la consulta.", fontsize=20)
plt.show()
#####

```

<- Verbos detectados de las repuestas dadas por el consultante.

```

#@title <- Palabras mas usadas por el consultante.
#####POR PALABRAS MAS COMUNES#####
fig = plt.figure(figsize=(10,5))
ax = fig.gca()

contador = Counter(palabras)

labels, values = zip(*contador.most_common(20))

indexes = np.arange(len(labels))
width = 1

ax.bar(indexes, values, 0.5)
ax.set_xticks(indexes)
ax.set_xticklabels( labels, rotation=90, fontsize=20)
ax.set_title("Las palabras más frecuentemente usadas por el paciente", fontsize=20)
plt.show()
#####

```

<- Palabras mas usadas por el consultante.

```

#@title <- Cuantas veces se encontro el concepto de Depresion:
print("El concepto de 'Depresion' se encontro {} vez en las repuestas".format(cantidad))
#####
El concepto de 'Depresion' se encontro 3 vez en las repuestas

```

<- Cuantas veces se encontro el concepto de Depresion:

```

#@title <- Carga la Ontologia
Animo=["Perezoso", "Desalentada", "Desalentadas", "Decaído", "Desmotivado", "Desalentado", "Desalentados", "Desalentó", "
Apatia=["No Tiro", "No Me Dan Ganas", "Pasivo", "No Siento ", "Nada De Sexo", "No Quiero Nada", "Dejadedz", "Dejado", "Vac
Alimentacion=["Indiferente", "Yo", "Apetito", "Desinterés", "Emoción", "Abatimiento", "Dejadedz", "Dejado", "Vacío", "Aband
Descanso=["Débil", "Débiles", "Debilité", "Raquítico", "Debilitado", "Enclenque", "Flojo", "Apagado", "Debilitada", "Mamad
Actividad_fisica=["Desaliento", "Frustración", "Trauma", "Abatimiento", "Extenuación", "Incapacidad", "Agitado", "Débil"
Desaliento=["Ausente", "Ausentes", "Ausentado", "Ausencia", "Alejado", "Lejano", "Ausentada", "Ausentó", "Ausentará", "Aus
Cargo_de_conciencia=["Culpable", "Culpado", "Culpada", "Culpa", "Culpé", "Culpar", "Culpables", "Errores", "Embarrarla", "
Atencion_dispensa=["Ausente", "Ausentes", "Ausentado", "Ausentados", "Ausentada", "Ausentadas", "Ausentó", "Ausentará", "
Ideas_suicidas=["Desvelo", "Sensible", "Frágil", "Miedo", "Intranquilo", "Desconsolado ", "Desilusionado ", "Decepciona

```

<- Carga la Ontologia

```

#title <- Relaciona la Ontologia de sintomatologia depresiva, con las diferentes respuestas generadas por el pac
ramaf = []

print("Palabras relacionadas con la siotomatologia: Animo")
rama = []
palabras_estudiar=set(palabras)
Animo=set(Animo)
Animo= palabras_estudiar & Animo
if len(Animo) > 0 :
    print("Hay {} elementos coincidentes".format(len(Animo)))
    print(Animo)
    rama.append("Animo")
    ramaf.append("Animo")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la siotomatologia: Apatia")
rama = []
palabras_estudiar=set(palabras)
Apatia=set(Apatia)
Apatia= palabras_estudiar & Apatia
if len(Apatia) > 0 :
    print("Hay {} elementos coincidentes".format(len(Apatia)))
    print(Apatia)
    rama.append("Apatia")
    ramaf.append("Apatia")
else:
    print("No hay repeticiones")
print("")
#####

```

<- Relaciona la Ontologia de sintomatologia depresiva, con las diferentes respuestas generadas por el paciente.

```

print("Palabras relacionadas con la sintomatologia: Alimentacion")
rama = []
palabras_estudiar=set(palabras)
Alimentacion=set(Alimentacion)
Alimentacion= palabras_estudiar & Alimentacion
if len(Alimentacion) > 0 :
    print("Hay {} elementos coincidentes".format(len(Alimentacion)))
    print(Alimentacion)
    rama.append("Alimentacion")
    ramaf.append("Alimentacion")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la sintomatologia: Descanso")
rama = []
palabras_estudiar=set(palabras)
Descanso=set(Descanso)
Descanso= palabras_estudiar & Descanso
if len(Descanso) > 0 :
    print("Hay {} elementos coincidentes".format(len(Descanso)))
    print(Descanso)
    rama.append("Descanso")
    ramaf.append("Descanso")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la sintomatologia: Actividad_fisica")
rama = []
palabras_estudiar=set(palabras)
Actividad_fisica=set(Actividad_fisica)

```



```

Actividad_fisica= palabras_estudiar & Actividad_fisica
if len(Actividad_fisica) > 0 :
    print("Hay {} elementos coincidentes".format(len(Actividad_fisica)))
    print(Actividad_fisica)
    rama.append("Actividad_fisica")
    ramaf.append("Actividad_fisica")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la sintomatologia: Desaliento")
rama = []
palabras_estudiar=set(palabras)
Desaliento=set(Desaliento)
Desaliento= palabras_estudiar & Desaliento
if len(Desaliento) > 0 :
    print("Hay {} elementos coincidentes".format(len(Desaliento)))
    print(Desaliento)
    rama.append("Desaliento")
    ramaf.append("Desaliento")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la sintomatologia: Cargo_de_conciencia")
rama = []
palabras_estudiar=set(palabras)
Cargo_de_conciencia=set(Cargo_de_conciencia)
Cargo_de_conciencia = palabras_estudiar & Cargo_de_conciencia
if len(Cargo_de_conciencia) > 0 :
    print("Hay {} elementos coincidentes".format(len(Cargo_de_conciencia)))

```

```

print(Cargo_de_conciencia)
rama.append("Cargo_de_conciencia")
ramaf.append("Cargo_de_conciencia")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la sintomatologia: Atencion_dispersa")
rama = []
palabras_estudiar=set(palabras)
Atencion_dispersa=set(Atencion_dispersa)
Atencion_dispersa = palabras_estudiar & Atencion_dispersa
if len(Atencion_dispersa) > 0 :
    print("Hay {} elementos coincidentes".format(len(Atencion_dispersa)))
    print(Atencion_dispersa)
    rama.append("Atencion_dispersa")
    ramaf.append("Atencion_dispersa")
else:
    print("No hay repeticiones")
print("")
#####

print("Palabras relacionadas con la sintomatologia: Ideas_suicidas")
rama = []
palabras_estudiar=set(palabras)
Ideas_suicidas=set(Ideas_suicidas)
Ideas_suicidas = palabras_estudiar & Ideas_suicidas
if len(Ideas_suicidas) > 0 :
    print("Hay {} elementos coincidentes".format(len(Ideas_suicidas)))
    print(Ideas_suicidas)
    rama.append("Ideas_suicidas")
    ramaf.append("Ideas_suicidas")

```

```

    ramaf.append("Ideas_suicidas")
else:
    print("No hay repeticiones")

```

```

#@title <- Analisis generado con aprendizaje de maquina
from monkeylearn import MonkeyLearn
celdaML = "L{0}"
celdaMLF = celdaML.format(id)

respuestasML = []

multiple_cells = hoja[celdaMLF:celdaMLF]
for row in multiple_cells:
    for cell in row:
        respuestasML.append(cell.value)
#print(respuestasML)
m1 = MonkeyLearn('6443587718d4ce3259255d93d0cbb4b61d4ab3e0')
model_id = 'cl_prjHdwSn'
result = m1.classifiers.classify(model_id, respuestasML)
modelo=result.body
StrModelo = "".join(map(str, modelo))
StrModelo2 = "".join(StrModelo)
Modelo3 = StrModelo2.split()
#print("POS",Modelo3.index(['tag_name:']))
POS1 = Modelo3.index(['tag_name:'])
ModeloClas=(Modelo3[POS1:POS1+4])

StrClasificacion = "".join(map(str, ModeloClas))
StrClasificacion2 = "".join(StrClasificacion)

StrClasificacion2=StrClasificacion2.replace("[", "")
StrClasificacion2=StrClasificacion2.replace("]", "")
StrClasificacion2=StrClasificacion2.replace("{", "")
StrClasificacion2=StrClasificacion2.replace("}", " ")
StrClasificacion2=StrClasificacion2.replace("tag_name", "Clasificacion_ML")
print(StrClasificacion2)
#print("POS",Modelo3.index(['confidence:']))
#POS2 = Modelo3.index(['confidence:'])
#probabilidad_clas=(Modelo3[POS2:POS2+2])
probabilidad_clas=(Modelo3[POS1+5:POS1+6])

```

<- Analisis generado con aprendizaje de maquina

```

StrProbabilidad = "".join(map(str, probabilidad_clas))
StrProbabilidad2 = "".join(StrProbabilidad)
StrProbabilidad2=StrProbabilidad2.replace("[", "")
StrProbabilidad2=StrProbabilidad2.replace("]", "")
StrProbabilidad2=StrProbabilidad2.replace("{", "")
StrProbabilidad2=StrProbabilidad2.replace("}", " ")
StrProbabilidad2=StrProbabilidad2.replace("]", " ")
StrProbabilidad2=StrProbabilidad2.replace("}", " ")
StrProbabilidad2=StrProbabilidad2.replace(" ", " ")
pro=(Modelo3[POS1+5])
#StrProbabilidad2=StrProbabilidad2.replace("confidence", "% ")
print("Probabilidad_Clasificacion: ",pro[1],pro[2],".",pro[3],pro[4], "%")

```

```

Clasificacion_ML:Cargodeconciencia
Probabilidad_Clasificacion: 2 3 . 6 0 %

```

```

Cargar_informe=open("Informe.txt", "a")
Cargar_informe.write(informeF)
Cargar_informe.close()

```

```

"""
"Atencion_dispersa: {10}"
"\n"
"Ideas_suicidas: {11}"
"\n"
"Si anteriormente no se ve ninguna relación es porque no se encontraron correlaciones.\n"
"\n"
"Las palabras mas significativas dadas por el consultante son: \n"
"{12}"
"\n"
"\nLos adjetivos que este utilizo fueron: \n"
"{13}"
"\n"
"\nLos verbos que este utilizo fueron: \n"
"{14}"
"\n"
"\nLas palabras mas usadas por el consultante son: \n"
"{15}"
"\n"
"\n"
"\nLa clasificacion dada por el modelo de aprendizaje fue : \n"
"{16}"
"\n"
"\n"
"\nY su probabilidad es : \n"
"{17}"
"\n"
"\n"
"\nSi en alguno de los anteriores, te genere vacío, es porque el consultante no utilizo ningún concept
"relacionado a esa área especifica.\n"
)
informeP = informe.format(id,palabras,ramaf,Animo,Apatia,Alimentacion,Descanso,Actividad_fisica,Desaliento,Cargo_
informeF = "".join(informeP)

```

```

Cargar_informe=open("Informe.txt","a")
Cargar_informe.write(informeF)
Cargar_informe.close()

```