

# **Instituto Tecnológico y de Estudios Superiores de Occidente**

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación del 29 de noviembre de 1976.

Departamento de Electrónica, Sistemas e Informática  
**Maestría en Sistemas Computacionales**



## **DISEÑO E IMPLEMENTACIÓN DE UNA BASE DE CONOCIMIENTO PARA EL DIAGNÓSTICO MÉDICO DEL ASMA Y SU SEVERIDAD**

---

**TRABAJO RECEPCIONAL** para obtener el **GRADO** de  
**MAESTRO EN SISTEMAS COMPUTACIONALES**

Presenta: **JORGE LUIS OROZCO RUIZ**

Asesora **NORMA ISABEL VILLANUEVA PAREDES**

Tlaquepaque, Jalisco. Diciembre de 2021.

## AGRADECIMIENTOS

El autor desea dar las gracias a Intel por todo el apoyo recibido tanto por parte de la empresa, como dentro del equipo de trabajo para poder llevar a cabo los estudios de la maestría y por todas las facilidades otorgadas para poder cumplir con todas las exigencias de las asignaturas cursadas.

Por otro lado, desea dar gracias a la Dra. Verónica Herrera Espinosa por todo el tiempo necesario otorgado para las entrevistas, validaciones y opiniones dadas. Así como por el conocimiento y experiencia compartida, ya que toda la información proporcionada, fue pieza medular para el desarrollo del presente trabajo.

También desea dar las gracias al ITESO por el apoyo otorgado por medio del descuento empresarial, el cual permitió continuar con los estudios de forma ininterrumpida hasta la conclusión del posgrado.

Se desea dar gracias al CONACYT por la Beca 999041 recibida la cual representó una gran ayuda para poder estudiar un posgrado dentro de esta casa de estudios.

Por último, pero no menos importante se desea agradecer a la asesora del presente trabajo por toda la dedicación y compromiso para el desarrollo del presente trabajo de obtención de grado.

## DEDICATORIA

El autor dedica este trabajo a la memoria de todas aquellas personas que me han ayudado a llegar hasta este momento. Así como a mi familia, amigos y compañeros que me apoyaron y acompañaron durante mis estudios de posgrado.

# RESUMEN

El presente trabajo se desarrolla en torno al diseño e implementación de una base de conocimiento para el diagnóstico del asma en adultos, teniendo como principal propósito, representar el conocimiento de un médico experto en el tema, aplicando técnicas de modelado conocidas como Reglas de Producción y Mapas Cognitivos Difusos, cuyo conocimiento almacenado sirva para determinar la probabilidad de asma y su severidad en un paciente.

Este trabajo propone una metodología para llevar a cabo el proceso de adquisición y representación del conocimiento, la cual inicia con la selección del dominio y la elección del experto humano, de quién será extraído el conocimiento necesario para el diagnóstico del asma, utilizando instrumentos de recolección de datos: entrevistas y cuestionarios. Una vez hecha la recolección y caracterizado el discurso médico del experto, la información obtenida es analizada y representada en la base de conocimiento, que incluye las reglas de producción, que ayudan a descartar la presencia de asma en un paciente y el mapa cognitivo difuso que determina lo contrario. Las reglas de producción son almacenadas en una base de datos no relacional (MongoDb) y el mapa cognitivo difuso se almacena dentro de una base de datos orientada a grafos (Neo4j). Ambas bases de datos conforman la base de conocimiento que usará el experto para diagnosticar el padecimiento.

La base de conocimiento creada es validada por el experto humano mediante la implementación de un prototipo que integra, un motor de búsqueda y una interfaz gráfica, desarrollado con tecnología web. Las pruebas realizadas por el experto, precisamente derivadas del proceso de validación y verificación permitieron hacer reajustes a la base de conocimiento, permitiendo elevar la fiabilidad de los resultados emitidos por el prototipo.

Durante el desarrollo de este trabajo se enfrentaron diferentes desafíos, siendo la representación del conocimiento el reto principal, por estar siempre presente la incertidumbre y la experiencia del experto en el proceso de toma de decisiones. Es por ello que se tomó la ventaja que ofrecen los mapas cognitivos difusos respecto a la flexibilidad de representar la incertidumbre, conjugando la teoría de grafos y la teoría difusa, para incrementar la asertividad en el diagnóstico médico.

Está claro que, las tecnologías de almacenamiento de datos utilizadas nos ayudan a crear una base de conocimiento más rápidamente e incluso ampliarla, permitiendo así, agregar fácilmente nuevos elementos de diagnóstico, que incluyan personas de cualquier rango de edad y no solo personas adultas como se consideraron en la base de conocimiento implementada. Además, como parte del trabajo futuro a desarrollar, se puede habilitar dentro del prototipo el manejo de estudios de laboratorio, lo que complementaría la evaluación y confirmación del diagnóstico.

# TABLA DE CONTENIDO

<b>MAESTRÍA EN SISTEMAS COMPUTACIONALES.....</b>	<b>1</b>
<b>1 INTRODUCCIÓN .....</b>	<b>11</b>
1.1 ANTECEDENTES.....	13
1.2 JUSTIFICACIÓN.....	13
1.3 PROBLEMA .....	14
1.4 OBJETIVOS.....	14
1.4.1 Objetivo general .....	14
1.4.2 Objetivos específicos.....	14
1.5 NOVEDAD CIENTÍFICA, TECNOLÓGICA O APORTACIÓN.....	15
<b>2 ESTADO DEL ARTE O DE LA TÉCNICA.....</b>	<b>16</b>
2.1 ADQUISICIÓN DEL CONOCIMIENTO .....	17
2.2 REPRESENTACIÓN DEL CONOCIMIENTO .....	17
2.3 IMPLEMENTACIÓN DE UNA BASE DE CONOCIMIENTO .....	19
2.4 COMPARATIVO ENTRE AUTORES.....	20
<b>3 MARCO TEÓRICO/CONCEPTUAL .....</b>	<b>22</b>
3.1 SISTEMA EXPERTO .....	23
3.2 SISTEMA EXPERTO BASADO EN CONOCIMIENTO MÉDICO .....	24
3.3 ADQUISICIÓN DEL CONOCIMIENTO .....	24
3.4 REPRESENTACIÓN DEL CONOCIMIENTO .....	25
3.5 BASE DE CONOCIMIENTO .....	25
3.6 REGLAS DE PRODUCCIÓN.....	26
3.7 MOTOR DE INFERENCIA .....	28
3.8 INTERFAZ GRÁFICA.....	29
3.9 TEORÍA DIFUSA.....	29
3.10 LÓGICA PROPOSICIONAL.....	30
3.11 TEORÍA DE GRAFOS.....	31
3.12 MAPAS COGNITIVOS DIFUSOS. ....	32
3.13 BASES DE DATOS NO RELACIONALES .....	32
3.13.1 MongoDB.....	33
3.13.2 Neo4j.....	33
3.14 DOMINIO .....	34
3.15 LINUX.....	34
3.16 UBUNTU .....	35
3.17 SERVIDOR WEB .....	35
3.18 APACHE.....	36
3.19 PYTHON.....	36
<b>4 DESARROLLO METODOLÓGICO .....</b>	<b>37</b>
4.1 METODOLOGÍA APLICADA.....	38

4.2	INGENIERÍA DEL CONOCIMIENTO .....	39
4.2.1	Adquisición del conocimiento .....	40
4.2.1.1	Selección del dominio .....	40
4.2.1.2	Selección de experto.....	40
4.2.1.3	Extracción del conocimiento .....	40
4.2.1.4	Análisis e interpretación del discurso del experto .....	42
4.2.1.5	Evaluación y selección de técnicas de representación del conocimiento.....	44
4.2.1.6	Validación y verificación .....	50
4.3	IMPLEMENTACIÓN .....	51
4.3.1	Entorno de desarrollo.....	51
4.3.1.1	Evaluación y selección de tecnologías para el almacenamiento de conocimiento.....	51
4.3.1.2	Instalación y configuración de herramientas para el desarrollo del prototipo.....	52
4.3.2	Prototipo .....	52
4.3.2.1	Levantamiento de requerimientos. ....	52
4.3.2.2	Análisis y diseño .....	53
4.3.2.3	Desarrollo de la base de datos .....	53
4.3.2.4	Desarrollo de interfaz .....	56
4.3.2.5	Pruebas de integración e implementación .....	59
4.3.2.6	Validación y verificación. ....	59
<b>5</b>	<b>RESULTADOS Y DISCUSIÓN.....</b>	<b>60</b>
5.1	RESULTADOS .....	61
5.1.1	Implementación del prototipo en un ambiente productivo .....	61
5.1.2	Experiencia del experto con la plataforma .....	61
5.1.3	Validación y calibración.....	61
5.1.3.1	Validación de resultados exitosos y fallidos.....	64
5.1.3.2	Calibración .....	68
5.2	DISCUSIÓN.....	69
<b>6</b>	<b>CONCLUSIONES.....</b>	<b>70</b>
6.1	CONCLUSIONES .....	71
6.2	TRABAJO FUTURO .....	71
<b>7</b>	<b>BIBLIOGRAFÍA.....</b>	<b>73</b>
<b>8</b>	<b>GLOSARIO DE TÉRMINOS MÉDICOS.....</b>	<b>74</b>
	<b>APÉNDICE A. ANEXOS.....</b>	<b>75</b>
	ANEXO 1.1 CUESTIONARIO PARA LA SELECCIÓN DE DOMINIO .....	75
	ANEXO 1.2 CUESTIONARIO PARA EL DIAGNÓSTICO DE ASMA EN ADULTOS .....	76
	ANEXO 1.3 CUESTIONARIO PARA ESTABLECER LA RELEVANCIA EN LA RELACIÓN ENTRE LOS ELEMENTOS DE DIAGNÓSTICO.....	78
	<b>APÉNDICE B. OBTENCIÓN DE REQUERIMIENTOS MEDIANTE EL ESTANDAR IEEE 830.....</b>	<b>79</b>
	1. FUNCIONALIDADES DEL PRODUCTO .....	79
	2. CARACTERÍSTICAS DEL USUARIO .....	79
	3. RESTRICCIONES.....	79
	4. SUPOSICIONES Y DEPENDENCIAS .....	79

5. ACTUALIZACIÓN DE REQUERIMIENTOS.....	80
6. REQUERIMIENTOS ESPECÍFICOS .....	80
6.1 Requerimientos de interfaz externa.....	80
6.1.1 Interfaz de usuario .....	80
6.1.2 Interfaz de hardware .....	80
6.1.3 Interfaz de software .....	80
6.2 Requerimientos funcionales .....	81
6.2.2 Ingresar información del paciente (RF1).....	81
6.2.3 Mostrar la clasificación de severidad asmática del paciente (RF2) .....	82

## LISTA DE FIGURAS

Fig. 3.1 Arquitectura de un sistema experto.	23
Fig 3.11 Representación de un grafo.	31
Fig. 4.2.1.4.1 Formula para el diagnóstico del asma.	42
Fig. 4.2.1.5.2.1 Mapa cognitivo difuso.	50
Fig. 4.3.2.2 Arquitectura propuesta para el desarrollo del prototipo.	53
Fig. 4.3.2.3.1.1 Catálogo de elementos que pueden indicar ausencia de asma.	54
Fig. 4.3.2.3.1.2 Representación de las reglas de producción dentro de la base de datos.	55
Fig. 4.3.2.4.2 Interfaz para capturar los antecedentes del paciente.	57
Fig. 4.3.2.4.3 Interfaz para capturar las predisposiciones del paciente.	57
Fig. 4.3.2.4.4 Interfaz para capturar las causas de la condición de salud del paciente.	57
Fig. 4.3.2.4.5 Interfaz para capturar los síntomas del paciente.	58
Fig. 4.3.2.4.6 Interfaz para presentar la confirmación de los datos ingresados por el paciente.	58
Fig. 4.3.2.4.7 Interfaz para presentar los resultados del diagnóstico del paciente.	59
Fig. 5.1.3.1.1 Caso de prueba exitoso Id No. 3	64
Fig. 5.1.3.1.2 Resultado caso de prueba Id No. 3	64
Fig. 5.1.3.1.3 Caso de prueba Id No.9	65
Fig. 5.1.3.1.4 Resultado caso de prueba Id No. 9	65
Fig. 5.1.3.1.5 Caso de prueba Id No. 12	66
Fig. 5.1.3.1.6 Resultado caso de prueba Id No. 12	66
Fig. 5.1.3.1.7 Caso de prueba no exitoso Id No. 1	67
Fig. 5.1.3.1.8 Resultado caso de prueba no exitoso Id No. 1	67
Fig. 5.1.3.1.9 Caso de prueba no exitoso Id No. 16	68
Fig. 5.1.3.1.10 Resultado caso de prueba no exitoso Id No. 16	68



## LISTA DE TABLAS

Tabla 2.4.1 Trabajos realizados por otros autores para la adquisición del conocimiento. ....	20
Tabla 2.4.2 Comparativa por autor respecto a la representación del conocimiento. ....	20
Tabla 2.4.3 Comparación entre metodologías para la representación del conocimiento .....	21
Tabla 2.4.4 Trabajos realizados relacionados a la implementación de una base de conocimiento. ....	21
Tabla 2.4.5 Comparación de la validación del trabajo realizado por autor. ....	21
Tabla 3.6 Ejemplo de tabla PER .....	28
Tabla 3.13.1 Equivalencia de conceptos entre bases de datos relacionales y Mongo DB .....	33
Tabla 3.13.2 Equivalencia entre las bases de datos relacionales y base de datos basada en grafos. ....	33
Tabla 4.2.1.4.1 Elementos para el diagnóstico del asma. ....	43
Tabla 4.2.1.4.2 Elementos de diagnóstico ponderados y agrupados por nivel de abstracción. ....	44
Tabla 4.2.1.5.2.1 Tabla de pesos asignados a cada elemento de diagnóstico. ....	48
Tabla 4.2.1.5.2.2 Valores difusos para el diagnóstico del asma .....	49
Tabla 4.3.1.1 Comparativo de base de datos para el almacenamiento de conocimiento. ....	51
Tabla 4.3.1.2 Herramientas utilizadas para el desarrollo del prototipo. ....	52

# LISTA DE ACRÓNIMOS Y ABREVIATURAS

UI	Interfaz de usuario
MVC	Módulo Vista Controlador
KB	Base de Conocimiento
HTTP	Protocolo de Transferencia de Hipertexto
SE	Sistema Experto

---

# 1 INTRODUCCIÓN

---

En la actualidad existen sistemas basados en conocimiento o también conocidos como sistemas expertos, capaces de emular el conocimiento y el razonamiento de un experto humano. Estos sistemas están orientados a cubrir un dominio o un tema en específico y están compuestos por una base de hechos, una base de conocimiento, un motor de inferencia y módulos de entrada-salida para el usuario.

Uno de los dominios de mayor interés en el diseño e implementación de un sistema experto, es el campo médico, porque el conocimiento puede ser almacenado en una base de conocimiento y a partir de este inferir y generar nuevo conocimiento, útil para el diagnóstico, prevención y tratamiento de enfermedades que menguan la calidad de vida de la sociedad, tal es el caso del asma, tema que acoge el presente trabajo.

Desde luego hay diferentes técnicas para adquirir y representar el conocimiento, en donde figuran, dos actores principales: el experto humano y el ingeniero del conocimiento. Este último encargado de recopilar la información que posee el experto humano a partir de su discurso, mediante la aplicación de instrumentos como entrevistas, y cuestionarios, dirigidos y semi-dirigidos, mientras que la elaboración de metáforas de pizarrón y la construcción de mapas conceptuales, ayudan a unificar y estructurar el conocimiento adquirido. Estas técnicas y herramientas facilitan la adquisición y representación del conocimiento sobre un dominio, el cual puede ser modelado en una base de conocimiento.

Una base de conocimiento almacena toda aquella experiencia que los expertos humanos generan con el paso del tiempo, así como su conocimiento en diferentes áreas. Es por esto que la información proporcionada por el experto requiere ser modelada dentro de una base de conocimiento, de tal forma que permita su almacenamiento y posterior uso por parte de los sistemas expertos.

Para el diseño de una base de conocimiento, algunos autores han modelado el conocimiento utilizando mapas de pizarrón, mapas conceptuales y reglas de producción. Por otro lado, existen autores que realizaron modelos usando redes de Petri difusas y mapas cognitivos difusos, los cuales toman todos los beneficios de la teoría difusa y la teoría de grafos para representar de manera más sencilla el conocimiento.

El almacenamiento del conocimiento se puede realizar de diferentes formas, en forma de reglas del modo Si – Entonces - en el caso de las reglas de producción -, las cuales permiten conformar una base de conocimiento ya que, al estar compuestas de comparaciones atómicas, permite que el sistema sea capaz de determinar un resultado con base en el conocimiento mapeado en dichas reglas. Por lo que se almacenan en programas donde pueden ser consultadas por un sistema experto y en forma de grafos en bases de datos no relacionales, - en el caso de los mapas cognitivos difusos -, manejan un esquema flexible, en el cual se pueden asignar valores difusos, los cuales ayudan a representar de una mejor manera el conocimiento y permiten almacenarlo fácilmente.

Respecto a los módulos de entrada y salida que constituyen un sistema experto, sin duda las nuevas tecnologías han favorecido el desarrollo de interfaces más intuitivas que permiten consumir el conocimiento mediante el uso de buscadores que procesan los algoritmos y proveen una respuesta al usuario final.

El presente trabajo contribuye al diseño, desarrollo e implementación de una base de conocimiento para el campo médico, cuyo dominio específico es el diagnóstico del asma en adultos y su severidad. El conocimiento es adquirido a través de entrevistas semi-dirigidas y elaboración de cuestionarios, previa selección de un médico experto en el tema, del cual se extrae el conocimiento relacionado al diagnóstico del asma.

Una vez extraída toda la información necesaria, esta se analiza y se representa mediante reglas de producción programadas, las cuales son almacenadas dentro de una base de datos en formato JSON y mapas cognitivos difusos, mismos que son almacenados dentro de una base de datos no relacional. También como parte del alcance se incluye el desarrollo de un motor de búsqueda y una interfaz gráfica de usuario para realizar consultas a la base de conocimiento. Utilizando la tecnología de Angular y Python se consigue desarrollar el prototipo de diagnóstico para la detección del padecimiento.

La motivación de este esfuerzo es brindar una herramienta tecnológica que permita determinar la probabilidad de que un paciente tenga asma y su nivel de severidad.

En este capítulo se presentan los antecedentes, la justificación y el planteamiento del problema, así como también se establecen los objetivos generales y específicos del proyecto.

## 1.1 Antecedentes

La ingeniería del conocimiento es una de las ramas de la inteligencia artificial, relacionada con la comprensión del lenguaje natural, la extracción y representación del conocimiento y por ende con los sistemas expertos.

La lógica difusa permite la formulación de sentencias y argumentos vagos o imprecisos, por lo que nos ayuda a trabajar con el pensamiento humano, dado que este es impreciso por naturaleza. Este tipo de lógica permite calcular predicados que sirven de base para los formalismos que tratan de modelar de manera procedural y declarativa, distintas formas de conocimiento.

La teoría de grafos, basada en un conjunto de puntos (nodos) unidos por líneas (aristas), ayuda a representar relaciones no consecuentes (grafos no dirigidos) y relaciones causales (grafos dirigidos) entre los elementos que componen un grafo. Una herramienta gráfica muy útil para representar conocimiento y sus relaciones.

La combinación de la teoría difusa con la teoría de grafos que se incluye en el trabajo realizado por Ashish Chandiok y D. K. Chaturvedi [2] los cuales realizaron una comparativa entre los modelos basados en técnicas de aprendizaje máquina y modelos basados en mapas cognitivos difusos. Concluyendo que gracias a que la teoría difusa proporciona mucha flexibilidad al momento de realizar el modelado, los mapas cognitivos que aprovechan estas características permiten modelar el conocimiento más adecuadamente.

Por otro lado, autores como Aldana y Velázquez [1] en su tesis desarrollaron una base de conocimiento para el diagnóstico del asma, en el cual se expone el proceso de obtención, procesamiento e implementación de las reglas de producción resultantes del experto médico.

## 1.2 Justificación

Es imprescindible para los sistemas expertos, contar con una base de conocimiento para su funcionamiento. Actualmente existen diferentes técnicas y metodologías para adquirir y representar el conocimiento. De esta manera, surge el interés por explorar y evaluar cuáles podrían ser las técnicas más apropiadas para diseñar e implementar una base de conocimiento para el diagnóstico médico del asma, y aunque el dominio médico es muy amplio, el asma es un padecimiento cuyos síntomas son fáciles de identificar comúnmente, esto hace que su diagnóstico sea más certero, permitiendo que con pocos estudios se pueda confirmar dicho padecimiento. Es por esto por lo que se decidió seleccionar este tema para desarrollar el presente trabajo.

## 1.3 Problema

En la actualidad, el modelado del conocimiento en el campo médico es un gran reto, ya que este se compone básicamente de aprendizaje adquirido mayormente de forma empírica y en donde las experiencias de cada médico definen la forma en la que se llega a un diagnóstico, por esto se incrementa la complejidad del diseño de una base de conocimiento, por lo que el uso de técnicas como las reglas de producción y los mapas cognitivos difusos pueden representar con mayor fidelidad el conocimiento.

## 1.4 Objetivos

### 1.4.1 Objetivo general

El objetivo principal es diseñar una base de conocimiento para el diagnóstico médico del asma, mediante técnicas de modelado reglas de producción y mapas cognitivos difusos, su implementación a través del almacenamiento de las reglas definidas en sentencia de – Si .... Entonces...- dentro de una base de datos no relacional, y desarrollar una interfaz gráfica con un motor de búsqueda integrado para la consulta de información del paciente, que permita estimar la probabilidad el padecimiento y su severidad.

### 1.4.2 Objetivos específicos

Los objetivos específicos fueron agrupados en 2 grandes áreas:

#### Adquisición y Representación

- Seleccionar un experto humano que tenga experiencia en el diagnóstico del asma.
- Evaluar y elegir las técnicas de adquisición y representación del conocimiento.
- Diseñar, elaborar y aplicar los instrumentos para la adquisición de información.
- Analizar e interpretar los datos recolectados.
- Validar la información analizada con el experto humano.
- Representar el conocimiento del experto con reglas de producción y mapas cognitivos difusos.

#### Implementación

- Evaluar y elegir las tecnologías con las que se realizará el prototipo.
- Diseñar la base de conocimiento.
- Almacenar el conocimiento obtenido del experto humano en forma de reglas de producción.
- Crear un motor de búsqueda que permita al usuario consultar la base de conocimiento.
- Diseñar y desarrollar la interfaz gráfica de usuario.

- Realizar pruebas unitarias y de integración del prototipo.
- Validar el diagnóstico generado por el prototipo con el experto humano.

## 1.5 Novedad científica, tecnológica o aportación

Se utilizarían herramientas de vanguardia como lo son las bases de datos no relacionales para almacenar las reglas de producción del conocimiento obtenido a partir del experto médico, los mapas cognitivos difusos ideales para modelar el conocimiento por su flexibilidad y adaptabilidad, ya que, al combinar la teoría difusa con la teoría de grafos, permite representar la incertidumbre que caracteriza al conocimiento, y las tecnologías para el desarrollo web como lo es Angular, Python, HTML y la metodología MVC, para agilizar el desarrollo de interfaces de usuario.

---

## 2 ESTADO DEL ARTE O DE LA TÉCNICA

---

En este capítulo se presenta un resumen de la literatura revisada y relacionada con el uso de diferentes técnicas de adquisición y representación del conocimiento, así como su implementación en una base de conocimiento para el diagnóstico médico del asma.

Hoy en día el conocimiento existente en el campo de la medicina es objeto de interés para los ingenieros del conocimiento, quienes logran extraer y representar información, para construir una base de conocimiento en este ámbito.

En esta revisión se podrá apreciar cómo algunos autores han definido el proceso de la adquisición del conocimiento, y los problemas que han enfrentado para representar el conocimiento extraído del experto humano.

Además, se expone como la lógica proposicional, los mapas conceptuales, las tecnologías cognitivas, la teoría difusa y la teoría de grafos, han ayudado a la creación de modelos de representación del conocimiento más apropiados, aplicando las técnicas de reglas de producción y mapas cognitivos difusos; teniendo la segunda técnica más ventajas de representación sobre la primera, principalmente porque las reglas de producción no son capaces de representar con precisión el mundo real, ya que tienen la limitante de adquirir nuevo conocimiento a partir de los errores y la complicación de obtener información consistente del campo experto. Por otro lado, los mapas cognitivos difusos ayudan a expresar la incertidumbre propia del conocimiento humano, lo que hace que con su uso se tenga mayor flexibilidad a la hora de representar el conocimiento. La aplicación de la teoría difusa en los mapas cognitivos permite el manejo de valores difusos, que aproximan una representación más hacia un modelo biológico y apegado al mundo real, lo que permite un modelado del conocimiento médico más preciso y en consecuencia un diagnóstico más certero.

También se menciona como diversos autores manejan la creación de Reglas de Producción, las cuales al ser almacenadas en una base de datos conforman la base de conocimiento y como éstas son consumidas por algunos programas especializados. Asimismo, se mostrará la forma en la que algunos autores hicieron uso de las bases de conocimiento generadas y algunos trabajos futuros que podrían realizarse en torno al tema.



## 2.1 Adquisición del conocimiento

La Ingeniería del Conocimiento tiene que ver con la adquisición, representación, validación, inferencia, explicación y mantenimiento del conocimiento, la cual se constituye como una tecnología cognitiva muy útil para la generación de una base de conocimiento de un sistema experto. Este tipo de ingeniería es considerada una de las ramas de la Inteligencia Artificial (IA).

El ingeniero del conocimiento es el encargado de extraer el conocimiento del experto a partir del discurso médico mediante la aplicación de diferentes técnicas, como lo son la realización de entrevistas, la aplicación de cuestionarios o la implementación de la metáfora del pizarrón.

Algunos medios para adquirir el conocimiento comentadas por Sánchez [16] son mediante libros, revistas, informes y directamente del experto humano. Considerando la adquisición del conocimiento como la tarea más importante dentro del desarrollo de sistemas basados en conocimiento.

Carlos [9] establece que la información puede obtenerse de distintas fuentes: dialogo directo con expertos, información escrita, datos empíricos proporcionados por aparatos de medida o por observaciones y por datos gráficos, por lo que deben de identificarse todos los elementos que intervienen en la solución del problema.

Por otro lado, Aldana y Vázquez [1] mencionan que para adquirir el conocimiento es necesario realizar entrevistas a profundidad, de manera dirigidas, así como también, aplicar la metáfora del pizarrón con la finalidad de unificar el vocabulario entre el experto y el ingeniero.

Por lo que, para llevar a cabo la adquisición del conocimiento de una manera más sistemática, Sánchez [16] define las siguientes fases o etapas:

1. Identificación del problema.
2. Conceptualización.
3. Formalización.
4. Implementación
5. Prueba

Todas estas herramientas utilizadas por el ingeniero del conocimiento son especialmente útiles, dado que son de gran ayuda para recolectar y ordenar la información que se obtiene del experto humano, lo cual ayuda a que sea más fácil el representar el conocimiento del experto.

## 2.2 Representación del conocimiento

El modelado del conocimiento es una parte esencial para el funcionamiento de los sistemas expertos ya que son la fuente para la toma de decisiones de un sistema experto y que permite que el conocimiento extraído del experto humano pueda ser consultada y usada por un sistema experto.

Carlos [9] propone las siguientes fases para llevar a cabo la representación del conocimiento:

- Elección de un formalismo de representación del conocimiento
- Elección de una arquitectura que permita coordinar y manejar los distintos tipos de elementos que intervienen en la solución del problema
- Creación de la base de conocimiento utilizando el formalismo y la arquitectura elegida.
- Diseño de la interfaz del SE con el usuario y con el resto del entorno lógico.

El conocimiento puede ser representado de diferentes maneras, siendo la aplicación de la técnica del pizarrón una de las más comúnmente utilizadas para modelar el conocimiento, así también las reglas de producción, las cuales al ser almacenadas en forma de hechos constituyen una base de conocimiento, y los mapas cognitivos difusos que utilizan la teoría difusa para representar el conocimiento con un grado de certeza y más próximos a modelar la realidad del conocimiento.

El uso de las reglas de producción o también llamadas reglas de Si-Entonces proveen un formalismo para representar el conocimiento, es una de las formas más comúnmente utilizadas debido a su sencillez ya que consta de un conjunto de acciones o efectos que son ciertas cuando se cumplen un conjunto de condiciones o causas, como lo menciona Sánchez [16] dentro de su trabajo.

Algunos de los trabajos previos exponen los retos que se tienen al representar el conocimiento, entre ellos la complejidad y la incertidumbre. No todas las técnicas facilitan la fiel representación del conocimiento, sin embargo, se ha propuesto la teoría difusa para afrontar el desafío de la incertidumbre a través de los mapas cognitivos difusos, los cuales combinan la teoría de grafos y la teoría difusa, de esta manera se representa de forma clara el conocimiento, permitiendo la definición de variables, valores y lingüística difusa.

Fillottrani [10] menciona dentro de su trabajo que el objetivo fundamental de la lógica difusa es la formalización de sentencias y argumentos vagos o imprecisos y asegura que los sistemas de lógica difusa pueden considerarse como lógica clásica.

Aldana y Vázquez [1], dentro de sus conclusiones, mencionan que la mayor dificultad para modelar el conocimiento es la ambigüedad que puede existir en el discurso médico, así como también lo es representar correctamente el conocimiento obtenido del experto médico. Ellos utilizaron la metáfora del pizarrón para unificar el vocabulario existente entre el experto humano y el ingeniero del conocimiento y junto con los mapas conceptuales, lograron mapear las reglas de producción desde los conceptos obtenidos a partir del pizarrón, para a partir de ahí representar el conocimiento mediante reglas de producción.

Por otro lado, Ashish Chandiok y Chaturvedi [2] exponen en su trabajo los beneficios de considerar las reglas de producción como base para la construcción de los mapas cognitivos difusos ya que estas el conocimiento del experto definido atómicamente permitiendo establecer de forma sencilla la relación entre los elementos de diagnóstico; Esta combinación ofrece mayor certeza para representar el conocimiento.

Incluso los autores compararon el modelado usando redes neuronales y mapas cognitivos difusos en el campo médico y concluyeron que el uso de mapas cognitivos difusos, permiten una mayor flexibilidad y adaptabilidad del modelo que representa el conocimiento Chandiok, Ashish y Chaturvedi, [2].

## 2.3 Implementación de una base de conocimiento

En cuanto a la implementación de las bases de conocimiento dentro del dominio médico, tenemos que algunos autores mencionan a las implementaciones como trabajos futuros a realizar, dado que la aplicación y desarrollo de un sistema experto quedaba fuera del alcance de algunos proyectos.

En el trabajo de Aldana y Vázquez [1] se menciona que se deben realizar implementaciones variadas para intentar averiguar más acerca del proceso de caracterización del discurso del experto y también se describe como estos autores implementaron un sistema experto llamado “ASMA” para llevar a cabo el proceso de validación, realizado por 6 médicos de instituciones de salud pública.

Después de terminar su papel de ingenieros del conocimiento Aldana y Vazquez [1] modelaron el conocimiento adquirido acerca del asma bronquial, utilizando mapas conceptuales, los cuales se construyeron a partir del discurso médico, que dieron origen a las reglas de producción que se codificaron en la herramienta CLIPSWIN desarrollada por la NASA. Esta herramienta permite el ingreso de reglas de producción con relaciones del tipo “OR” exclusivo.

Por otro lado, para los autores del artículo [11] mencionan que el proceso de diagnóstico posee una gran complejidad dado que, desde el punto de vista del cómputo, el problema requiere la utilización de un conocimiento médico profundo y extenso tanto en el ámbito enciclopédico como en el práctico

En dicho artículo [11], los autores después de analizar tres diferentes sistemas para diagnóstico de la salud (INFOTOXI, GERISOFT, SEAA), concluyeron que los SE dentro del dominio de la medicina se ha contemplado como un proceso de diseño de clasificadores, que, dependiendo de los datos disponibles y del conocimiento introducido, puede resultar más efectiva que la individual del profesional de la salud.

En lo que se refiere a la metodología de implementación Carlos [9] propone una serie de fases por las que se debe pasar durante el desarrollo de una base de conocimiento desde el punto de vista del desarrollo de software, las fases propuestas son las siguientes:

1. Elección de la aplicación adecuada para capturar el conocimiento usado por los expertos.
2. Elección de la herramienta apropiada para almacenar el conocimiento.
3. Transferencia de experiencia desde el experto humano hacia el SE el cual comprende:
4. Validación

## 2.4 Comparativo entre autores

A continuación, se presenta un comparativo del trabajo realizado por cada uno de los autores dentro de los elementos de revisión visto en este capítulo.

Elementos de revisión /Autor(es)	Adquisición del conocimiento
Aldana y Vázquez	<ul style="list-style-type: none"> <li>• Entrevistas de profundidad</li> <li>• Entrevistas dirigidas</li> <li>• Metáfora del pizarrón</li> </ul>
Carlos	<ul style="list-style-type: none"> <li>• Dialogo directo con expertos</li> <li>• Información escrita</li> <li>• Datos empíricos proporcionados por aparatos de medida o por observaciones y por datos gráficos.</li> </ul>
Sánchez	<ul style="list-style-type: none"> <li>• Libros</li> <li>• Revistas</li> <li>• Informes</li> <li>• Entrevistas con el experto</li> </ul>

Tabla 2.4.1 Trabajos realizados por otros autores para la adquisición del conocimiento.

Elementos de revisión /Autor(es)	Representación del conocimiento
Aldana y Vázquez	<ul style="list-style-type: none"> <li>• Uso de la metáfora del pizarrón</li> <li>• Uso de las reglas de producción</li> </ul>
Carlos	
Sánchez	<ul style="list-style-type: none"> <li>• Uso de las reglas de producción</li> </ul>
Fillottrani	<ul style="list-style-type: none"> <li>• Uso de la lógica difusa para la representación de argumentos vagos o imprecisos</li> </ul>
Ashish Chandiok y Chaturvedi	<ul style="list-style-type: none"> <li>• Consideran las reglas de producción como base de los mapas cognitivos</li> <li>• Aplicación de la teoría difusa en los mapas cognitivos para el manejo de la incertidumbre</li> </ul>

Tabla 2.4.2 Comparativa por autor respecto a la representación del conocimiento.

Elementos de revisión /Autor(es)	Metodología para la representación del conocimiento
Aldana y Vázquez	<ul style="list-style-type: none"> <li>• Elaboración de la metáfora del pizarrón</li> <li>• Creación de las reglas de producción</li> <li>• Validación de la representación del conocimiento</li> </ul>
Carlos	<ul style="list-style-type: none"> <li>• Elección de un formalismo de representación del conocimiento</li> <li>• Elección de una arquitectura que permita coordinar y manejar los distintos tipos de elementos que intervienen en la solución del problema</li> <li>• Creación de la base de conocimiento utilizando el formalismo y la arquitectura elegida.</li> </ul>
Sánchez	<ul style="list-style-type: none"> <li>• Identificación del problema</li> <li>• Conceptualización</li> <li>• Formalización</li> <li>• Implementación</li> <li>• Prueba</li> </ul>
Ashish Chandiok y Chaturvedi	<ul style="list-style-type: none"> <li>• Selección del experto médico</li> <li>• Selección de padecimiento</li> <li>• Combinación de síntomas y lingüística difusa</li> <li>• Definición de valores difusos para cada variable</li> <li>• Construcción de modelos de machine learning</li> <li>• Construcción del mapa cognitivo difuso</li> </ul>

Tabla 2.4.3 Comparación entre metodologías para la representación del conocimiento

Elementos de revisión /Autor(es)	Implementación de una base de conocimiento
Aldana y Vázquez	<ul style="list-style-type: none"> <li>• La implementación del sistema experto “ASMA”</li> </ul>
Carlos	<ul style="list-style-type: none"> <li>• Diseño de la interfaz del SE con el usuario y con el resto del entorno lógico</li> </ul>
Sánchez	<ul style="list-style-type: none"> <li>• Elección de la aplicación para capturar el conocimiento utilizado</li> <li>• Elección de la herramienta de apropiada para el desarrollo</li> <li>• Transferencia de experiencia</li> <li>• Validación</li> </ul>

Tabla 2.4.4 Trabajos realizados relacionados a la implementación de una base de conocimiento.

Elementos de revisión /Autor(es)	Validación del modelo
Aldana y Vázquez	<ul style="list-style-type: none"> <li>• Se validó mediante el sistema “ASMA” utilizando 97 historias clínicas de pacientes con asma proporcionadas por el médico, mediante corridas individuales dentro del sistema.</li> </ul>
Carlos	<ul style="list-style-type: none"> <li>• Sugiere llevar a cabo una validación del modelo para mejoras continuas</li> </ul>
Sánchez	<ul style="list-style-type: none"> <li>• Presenta tres experimentos con la finalidad validar el aprendizaje con un algoritmo de backpropagation y el algoritmo evolutivo.</li> </ul>
Ashish Chandiok y Chaturvedi	<ul style="list-style-type: none"> <li>• Evaluación del estado de salud de un paciente mediante un mapa cognitivo difuso</li> </ul>

Tabla 2.4.5 Comparación de la validación del trabajo realizado por autor.

---

### 3 MARCO TEÓRICO/CONCEPTUAL

---

En este capítulo se presentan las bases teóricas y conceptuales acerca de los medios y recursos para la adquisición y representación de conocimiento, así como para el diseño e implementación de una base de conocimiento.

Para la adquisición de información se utilizan instrumentos que ayudan a recolectar la información del experto humano, las de uso más común son las entrevistas y los cuestionarios.

En la representación del conocimiento se hace uso de los mapas cognitivos difusos, los cuales surgen de la aplicación de la teoría difusa y la teoría de grafos, ambos han ayudado a la conceptualización de la información, ya que permiten realizar inferencias en el dominio de la medicina. Para el diseño e implementación de los mapas cognitivos difusos han sido almacenados en una base de datos no relacional con apoyo de la herramienta NEO4J.

Las reglas de producción también son usadas frecuentemente para representar el conocimiento, así que éstas han sido almacenadas en Mongo, cuya base de datos es no relacional, conformando de esta manera la base de conocimiento, para ser posteriormente utilizadas por los sistemas expertos.

### 3.1 Sistema experto

Un sistema experto es también conocido como sistema basado en conocimiento, y es todo aquel sistema que simula el comportamiento de un experto humano en un ámbito o dominio de algún campo, su operación consiste principalmente en simular el proceso de aprendizaje, razonamiento y memorización de un experto humano especializado en algún tema en específico, de esta manera permite a los usuarios del sistema buscar una posible solución.

Alonso [17] menciona que este tipo de sistemas se caracterizan por ser programas capaces de llevar a cabo búsquedas y recuperación de información interrelacionada y almacenada masivamente, utilizando una relación simbólica para establecer los conocimientos del dominio sobre el que actúa, son capaces de dialogar con el usuario y plantear le preguntas sobre el razonamiento seguido y pueden complementarse como una ayuda para los usuarios que requieren de una buena interpretación de información compleja

Los sistemas expertos permiten tomar decisiones con base en conocimiento y hechos que son almacenados en forma de reglas en una base de conocimiento. Las distintas relaciones, conexiones y afinidades sobre un tema pueden ser compiladas dentro del sistema lo que permite establecer relaciones altamente complejas y con múltiples interacciones.

La principal diferencia y ventaja de un sistema experto respecto a su versión humana, es que su tiempo de vida es infinito, en tanto se mantenga en constante actualización, y su principal desventaja es que aplica solo a un campo del conocimiento.

En la siguiente figura se puede observar la arquitectura de un sistema experto.

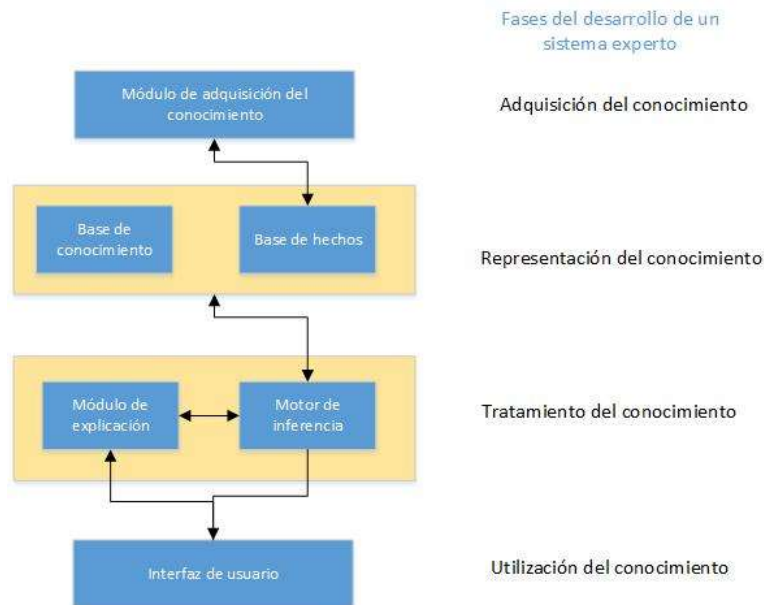


Fig. 3.1 Arquitectura de un sistema experto.

## 3.2 Sistema experto basado en conocimiento médico

Un sistema experto médico es aquel que es capaz de tomar conocimientos del dominio con base en hechos, conocimiento, inferencia y métodos de comunicación externa con el usuario mediante algún tipo de interfaz.

Se han aplicado estos sistemas en varias áreas de la medicina, particularmente en el diagnóstico de enfermedades con un alto grado de asertividad.

## 3.3 Adquisición del conocimiento

La adquisición del conocimiento es la parte inicial de un sistema experto, el ingeniero del conocimiento es el encargado de realizar la extracción y análisis de la información que obtiene del experto humano, a través de diversas herramientas como pueden ser los cuestionarios, las entrevistas dirigidas y semi-dirigidas y otras más que abstraen los principales conceptos presentes en el discurso del experto, como lo es el uso de la metáfora del pizarrón.

La adquisición del conocimiento es la fase inicial y la más importante del desarrollo de un sistema experto, ya que es el soporte sobre el que se construye la base de conocimiento del sistema. La base de conocimiento es usada posteriormente por un motor de inferencia dentro de la fase del tratamiento del conocimiento.

Velázquez [1] definió las siguientes cinco etapas para la adquisición del conocimiento:

1. **Identificación.** Es en esta etapa donde se conoce el problema y se crean sub-problemas, y también donde se identifican los recursos con los que se cuenta, justo es donde el Ingeniero del Conocimiento expone la información recolectada.
2. **Entendimiento.** Es en esta etapa donde se asimila la información y se presentan todos los cuestionamientos necesarios para poder determinar los conceptos necesarios y las relaciones que existen entre ellos para representar el conocimiento.
3. **Formalización.** En esta fase es donde se define como se deberá organizar la información con el fin de crear las reglas de producción, en esta etapa se representa toda la información extraída del experto humano en una base de conocimiento, es una de las fases más complicadas.
4. **Implementación.** Es en esta etapa donde se programa el conocimiento mediante la creación de un prototipo que almacena la información de las reglas de producción.
5. **Pruebas.** Es la última etapa de la adquisición del conocimiento y es aquí donde el ingeniero crea casos de prueba, cuyos resultados son mostrados y analizados en forma conjunta con el experto humano, quién le dará la validez a los mismos.



Estas etapas nos permiten diseñar la base de conocimiento de una forma más adecuada, ya que se define el proceso por el cual podemos representar la experiencia adquirida del experto humano y que ha sido recopilada por el Ingeniero del Conocimiento.

## 3.4 Representación del conocimiento

La representación del conocimiento es el mapeo de la información obtenida del experto humano de un estado a otro, su principal objetivo es expresar el conocimiento de una forma manejable para una computadora.

Esta etapa es la que permite aterrizar y crear un modelo basado en la adquisición del conocimiento realizada por el ingeniero del conocimiento y que representa de forma sistemática al experto humano.

Dicha representación se conforma de 2 componentes, el módulo de explicación y el motor de inferencia. El primer componente o módulo permite refinar el motor de inferencia y permite la construcción y validación de la base de conocimientos y también se encarga de explicar al usuario como es que el sistema experto ha llegado a esa solución. El segundo módulo, el cual está compuesto por el motor de inferencia el cual, es considerado la parte central de un sistema experto dado que es capaz de construir razonamientos haciendo uso de la base de conocimientos de forma determinística o probabilística.

Para llevar a cabo dicha representación se siguen los siguientes pasos:

- Decidir qué tipos de objetos y relaciones se deben representar.
- Elegir un vocabulario para codificar el conocimiento general del dominio.
- Codificar casos específicos del problema.
- Emplear procedimientos de inferencia para resolver los problemas codificados.

Existen diferentes maneras de representar el conocimiento, las cuales son usadas dependiendo del problema que se intenta resolver. Algunas de estas formas de representación son los marcos, mapas conceptuales, redes semánticas, lógica de predicados, ontologías y las reglas de producción. Estas últimas son adecuadas para el diagnóstico médico ya que generalmente se usan en problemas donde es necesario explicar el proceso necesario para alcanzar una solución.

## 3.5 Base de conocimiento

La base de conocimiento es toda aquella información referente a un tema en específico y que puede contener dos tipos de conocimiento, el efectivo que es todo aquel conocimiento basado en información certera como lo pueden ser los libros, revistas especializadas y el heurístico que es todo aquel conocimiento basado en las experiencias [5].

El tamaño y la calidad de la base de conocimientos está fuertemente relacionada a la calidad de la competencia del sistema experto y su funcionamiento.

El conocimiento contenido en la base se puede caracterizar, tomando como ejemplo la aplicación que un experto humano hace de su conocimiento para resolver de manera eficaz un problema de su ámbito de competencia. En tal búsqueda de la solución éste utiliza, por una parte, conocimiento teórico sobre el modelo a aplicar y, por otra parte, conocimiento práctico obtenido en el desarrollo de su actividad profesional, con el objetivo de ajustar el modelo teórico de resolución a la realidad del problema planteado.

Dentro de una base de conocimiento se encuentran almacenados los elementos, objetos, relaciones entre estos y los procesos involucrados en el discurso del experto humano, conteniendo de esta manera el conocimiento teórico y práctico.

El conocimiento almacenado en estas bases puede estar representado y modelado de diferentes formas como los son las reglas de producción o los mapas cognitivos difusos, estos modelos otorgan una flexibilidad dentro de los elementos que las componen y que son necesarias para que la información pueda ser utilizada de manera correcta por el sistema experto.

Las bases de conocimiento pueden existir en distintos dominios como lo pueden ser las que contienen conocimiento financiero, comercial o médicas. Y que dependiendo de los datos de las que están compuestas pueden ser categorizadas en 2 grandes categorías: las que contienen datos exactos se denominan bases de conocimientos factuales o precisas, mientras que aquellas cuyos datos no suele ser precisos como lo puede ser el pensamiento humano son clasificadas como bases de conocimiento heurístico.

Dado que en el presente trabajo se trata de una base de conocimientos sobre el asma, esta se considera una base de tipo heurística, ya que permite la realización del diagnóstico utilizando la experiencia adquirida del experto humano, la cual está compuesta con elementos exactos los cuales son combinados con la experiencia para poder llegar a un diagnóstico,

Las bases de conocimiento que contienen información heurística están fuertemente ligadas a los sistemas expertos médicos, ya que son la forma en la que estos son capaces de adquirir conocimiento para, dado que comprenden los hechos y el conjunto de reglas que definen el problema que se debe de resolver. Dichas reglas son también conocidas como reglas de producción, las cuales se encuentran almacenadas dentro de la base de conocimiento.

### 3.6 Reglas de producción

Las reglas de producción son un formalismo computacional, que permite representar el discurso del experto en un dominio. El proceso de construcción de dichas reglas se considera la parte medular de la ingeniería del conocimiento, ya que partiendo del uso de condiciones y consecuencias. Por ejemplo, SI existe un antecedente ENTONCES existe una consecuencia.

El modelado del conocimiento por medio de reglas de producción regularmente se usa cuando se debe de describir el proceso que se sigue para llegar a un resultado. Por lo que para lograr construir las reglas de producción se deben mapear todos los elementos involucrados en dicho proceso. De esta manera es posible

identificar adecuadamente cuales son los elementos considerados como antecedentes y cuáles como consecuentes.

Una vez que se obtienen son documentadas y almacenadas en tablas que contienen la regla de causalidad ya sean explícitas o implícitas y la regla o reglas que la componen. A este tipo de tablas se les conoce como tablas PER (Palabras Experto-Regla). A continuación, se muestra un ejemplo de una tabla PER.

Identificador del grupo de reglas	Determinación a partir de las propiedades requeridas de la soldadura de las propiedades que debe satisfacer los electrodos	Reglas
Palabras del experto	<p>"... En el problema de la determinación de electrodos para la soldadura eléctrica, juegan un papel importante distintas características finales que la soldadura debe tener, como ser: penetración, propiedades mecánicas, operatividad, contenido de hidrógeno, termina del cordón. Los tipos de electrodos a tener presente son: rutilico, básicos, celulósicos y rutilcelulósicos..."</p> <p>"... los rutilcelulósicos tiene penetración media, propiedades mecánicas entre regulares y malas, operatividad buena, contenido de hidrógeno entre medio y alto y terminación del cordón entre buena y regular..."</p> <p>"...Los rutilicos tienen penetración baja, propiedades mecánicas malas, operatividad buena , contenido de hidrógeno medio y terminación del cordón buena..."</p> <p>"...los celulósicos tienen penetración alta, propiedades mecánicas regulares, operatividad buena, contenido de hidrógeno medio y terminación del cordón mala..."</p> <p>"...los básicos tiene penetración media, propiedades mecánicas buenas, operatividad mala, contenido de hidrógeno bajo y terminación del cordón regular..."</p>	<p>SI LA SOLDADURA &lt;--PENETRACIÓN REQUERIDA=ALTA ENTONCES ELECTRODO&lt;--PENETRACIÓN DADA=ALTA</p> <p>SI SOLDADURA&lt;--PENETRACIÓN REQUERIDA=MEDIA ENTONCES ELECTRODO&lt;--PENETRACIÓN DADA=MEDIA</p> <p>SI SOLDADURA&lt;--PENETRACIÓN REQUERIDA=BAJA ENTONCES ELECTRODO&lt;-- PENETRACIÓN DADA=BAJA</p> <p>SI LA SOLDADURA &lt;--PROPIEDADES MECÁNICAS REQUERIDAS=BUENAS ENTONCES ELECTRODO&lt;--PROPIEDADES MECÁNICAS REQUERIDAS DADAS=BUENAS</p> <p>SI LA SOLDADURA &lt;--PROPIEDADES MECÁNICAS REQUERIDAS=REGULARES ENTONCES ELECTRODO&lt;--PROPIEDADES MECÁNICAS REQUERIDAS DADAS=REGULARES</p> <p>SI LA SOLDADURA &lt;--PROPIEDADES MECÁNICAS REQUERIDAS=MALAS ENTONCES ELECTRODO&lt;--PROPIEDADES MECÁNICAS REQUERIDAS DADAS=MALAS</p> <p>SI LA SOLDADURA &lt;--OPERATIVIDAD REQUERIDAS=BUENA ENTONCES ELECTRODO&lt;--OPERATIVIDAD DADA=BUENA</p>

	<p>SI LA SOLDADURA &lt;--OPERATIVIDAD REQUERIDAS=MALA ENTONCES ELECTRODO&lt;--OPERATIVIDAD DADA=MALA</p> <p>SI LA SOLDADURA &lt;--CONTENIDO DE HIDRÓGENO           REQUERIDO=MEDIO ENTONCES ELECTRODO&lt;--CONTENIDO DE HIDRÓGENO           DADO=MEDIO</p> <p>SI LA SOLDADURA &lt;--CONTENIDO DE HIDRÓGENO           REQUERIDO=BAJO ENTONCES ELECTRODO&lt;--CONTENIDO DE HIDRÓGENO           DADO=BAJO</p> <p>SI LA SOLDADURA &lt;--TERMINACIÓN DEL CORDÓN           REQUERIDO=BUENA ENTONCES           ELECTRODO&lt;-- TERMINACIÓNN   DEL   CORDÓN DADO=BUENA</p> <p>SI LA SOLDADURA &lt;--TERMINACIÓN DEL CORDÓN           REQUERIDO=REGULAR ENTONCES           ELECTRODO&lt;-- TERMINACIÓNN   DEL   CORDÓN DADO=REGULAR</p> <p>SI LA SOLDADURA &lt;--TERMINACIÓN DEL CORDÓN           REQUERIDO=MALA ENTONCES           ELECTRODO&lt;-- TERMINACIÓNN   DEL   CORDÓN DADO=MALA</p>
--	--

Tabla 3.6 Ejemplo de tabla PER

### 3.7 Motor de inferencia

El motor de inferencia es una componente de los sistemas expertos encargado de activar las reglas de producción en función a la información contenida en la base de datos, así como de proporcionar las reglas que determinaron determinada consulta al usuario y de esta forma trazar la justificación de esta.

Existen dos tipos de motores, los que usan una estrategia orientada por el objetivo o también conocida como búsqueda hacia atrás y los que su estrategia es orientada por los datos la cual es conocida como búsqueda hacia adelante. La primera estrategia toma como origen al objetiva para a partir de ahí construir un árbol hacia los datos conocidos en donde las reglas son asociadas a las ramas de este. Por otro lado,

aquellos motores de inferencia orientados a los datos toman como origen los datos para a partir de esto intentar construir un elemento que contenga al objetivo.

El motor de inferencia usa las evidencias, como las reglas aplicadas a la base de conocimiento para obtener nuevas conclusiones o hechos. Por ejemplo, si la premisa de una regla es cierta, entonces la conclusión de la regla debe ser también cierta. Los datos iniciales se incrementan incorporando las nuevas conclusiones. Es por ello, que tanto los hechos iniciales o datos de partida como las conclusiones derivadas de ellos forman parte de los hechos o datos de que se dispone en un instante dado. Para obtener conclusiones, los expertos utilizan diferentes tipos de reglas y estrategias de inferencia y control [6].

El rendimiento del motor de inferencia depende del conjunto de reglas en su base de conocimiento. Hay situaciones en las que el motor de inferencia puede concluir utilizando un conjunto de reglas, pero no puede, utilizando otro (aunque éstos sean lógicamente equivalentes).

### 3.8 Interfaz gráfica

La interfaz gráfica es una parte de la arquitectura del software en donde, su principal funcionamiento es permitir le al usuario interactuar con el sistema y el cual le permite realizar tareas de lectura/escritura de datos en el sistema.

Está compuesta de una serie de componentes gráficos que permiten la entrada y salida de datos en pantalla, mismos que poden ser accedidos por el sistema desde el backend.

La manipulación de una interfaz gráfica web es realizada en su mayoría mediante el uso de javascript, quien es el encargado de proveer a los sistemas los datos que se encuentran en la interfaz.

### 3.9 Teoría difusa

La teoría difusa fue introducida por primera vez por Lofti A. Zadeh permitiendo representar el conocimiento de una forma más cualitativa. Dicha teoría permite la representación de concepto basado en el grado de certeza y no en un límite preestablecido, lo que facilita el modelado del pensamiento humano.

Esta teoría usa conjuntos difusos, tomando conceptos básicos de la teoría de conjuntos solo en este caso en vez de definir la pertenencia de un elemento, es decir valores de 0 o 1, se define sobre valores entre 0 y 1. Permitiendo tener un espectro más amplio para clasificar cada elemento de este tipo de conjuntos [3].

El grado de ausencia o presencia de interacción entre elementos de dos conjuntos difusos distintos está definido por las relaciones difusas o también conocidas como reglas difusas, que son todas aquellas proposiciones del tipo SI-ENTONCES que modela algún problema que se desee resolver. Mismas que pueden ser llevadas a la lógica proposicional con el fin de aplicar los conceptos de está.

## 3.10 Lógica proposicional

La lógica proposicional se encarga de estudiar las frases simples que son consideradas como elementos básicos dentro de la transmisión del conocimiento humano y que pueden ser consideradas como Verdaderas o Falsas mediante el uso de la semántica proposicional.

Las dos partes que componen dicha lógica son las proposiciones, las cuales se caracterizan por ser enunciados cortos que pueden o no estar relacionados entre sí y las conclusiones que se interpreta como las consecuencias de un conjunto de proposiciones.

Las proposiciones o premisas son representadas con letras del alfabeto en mayúscula como O, P, R, S o también pueden ser representadas mediante una letra P seguida de un sufijo numérico

Al conjunto de proposiciones seguidas de una conclusión se le conoce como argumento o razonamiento. Un ejemplo de un razonamiento válido puede ser:

Premisa:

- P: La madre de María es Fátima
- Q: El padre de Fátima es inglés

Conclusión:

- R: María tiene ascendencia inglesa.

Mientras un argumento inválido se puede ejemplificar de la siguiente manera:

Premisa:

- P: Todos los limones son frutos redondos
- Q: Todos los frutos redondos crecen en los árboles

Conclusión:

- R: Los melones crecen en los árboles.

Esta lógica se apoya principalmente en conectores lógicos, los cuales son expresiones lingüísticas que conectan como, por ejemplo: y; o; no; si... entonces...; solo si... entonces..., que regularmente son utilizados dentro de los argumentos y que ayudan a expresar argumentos en los cuales las conclusiones pueden ser deducidas o inferidas.

Una deducción es aquella en la cual partiendo de las premisas se puede validar si la conclusión es verdadera o falso. Un ejemplo de esto puede ser:

**Si** adquiero un préstamo de 5000 pesos con un interés del 10% anual

**entonces** en 10 años pagare 500 pesos.

Si se realiza el cálculo del interés de la primera premisa, se podrá comprobar la conclusión es verdadera.

Por otro lado, la inferencia se caracteriza por no necesariamente comprobar la veracidad de la conclusión a partir de las premisas, pero que si pueden afirmar la probabilidad de que estas sean verdaderas y que regularmente se basan en hechos. A continuación, un ejemplo:

**Infero** que como este año ha llovido mucho durante los últimos meses, el invierno será muy frío igual que al año pasado.

La inferencia anterior se realiza basada en hechos anteriores como el año pasado llovió durante meses e hizo frío entonces este año también lo será, no es algo seguro, pero es una probabilidad.

El uso de la lógica proporcional en la construcción de las reglas de producción utilizadas para la representación del conocimiento es de suma importancia dado que son las bases sobre las que se construyen dichas reglas.

### 3.11 Teoría de grafos

La teoría de grafos forma parte de las matemáticas discretas y ha tenido mucho desarrollo en los últimos años debido a la gran cantidad de aplicaciones que pueden tener.

Un grafo se compone de un conjunto de vértices que pueden o no estar conectados entre sí, este tipo de conexiones representan las relaciones existentes entre los diferentes vértices. Al conjunto de nodos que se encuentran relacionados se les denomina aristas.

Los grafos dirigidos son un tipo de grafos en donde sus aristas van en una sola dirección y por lo tanto nos permite representar relaciones de tipo causa y efecto, por lo que nos permiten representar conceptos más complejos de una forma sencilla.

La representación de problemas mediante el uso de grafos permite representar cada elemento del problema como un vértice, el cual dependiendo de la influencia que tiene un vértice sobre otro, se le asigna un valor a su arista, este valor también es conocido como peso.

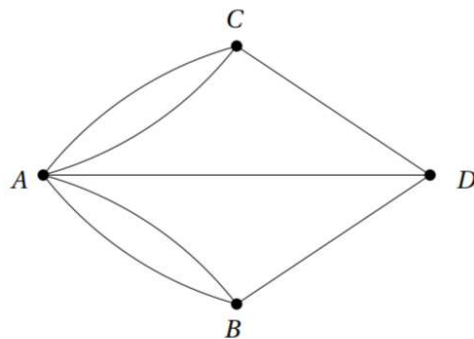


Fig 3.11 Representación de un grafo.

### 3.12 Mapas cognitivos difusos.

Los mapas cognitivos combinan la teoría de grafos y la teoría difusa, permitiendo modelar conceptos y problemas más abstractos. La construcción de dichos mapas se realiza mediante la representación de cada característica con nodos, y en donde cada valor en las aristas del grafo es representado por un valor difuso, es decir un valor de  $[0-1]$ .

Una de las ventajas que tienen los mapas cognitivos sobre las reglas de producción para modelar el conocimiento es que la teoría de grafos aporta al modelo mayor intuición, dado que se aproxima más a una forma biológica, en contra parte con la incompetencia de las reglas de producción para adquirir conocimiento de los errores, por lo que los mapas cognitivos difuso proporciona mayor flexibilidad al momento de representar el conocimiento. En cuanto a la principal desventaja de los mapas difusos es la incapacidad de explotar al máximo la lógica difusa durante el proceso de propagación de la información ya que es estos conceptos se utilizan principalmente durante su construcción.

### 3.13 Bases de datos no relacionales

Las bases de datos no relacionales son un nuevo paradigma en el diseño de base de datos, en donde se le da mayor peso a la flexibilidad y a la capacidad de almacenamiento de grandes volúmenes de datos, sacrificando la disponibilidad o la integridad de los datos.

Este tipo de tecnología surgió a raíz de la necesidad de almacenar toda la información que se está generando por el uso de nuevos dispositivos y dado a que dichos datos generados no comparten la misma estructura se hace imposible almacenarlos en una base de datos relacional.

La forma en que estas bases de datos almacenan la información es completamente diferente de las bases de datos relacionales, ya que en lugar de que los datos sean almacenados en forma de tablas con filas y columnas, estos son almacenados en un formato libre a forma de fichero, por lo cual cada uno de estos registros pueden contener información completamente diferente entre sí, por lo que de esta manera las bases no relacionales permiten mayor flexibilidad en el almacenamiento y manejo de los datos.

Las bases de datos no relacionales pueden almacenar de una mejor forma un modelo que represente el conocimiento de un experto humano, ya que cada registro puede ser tratado como elemento diferente y permite la creación de relaciones entre cada elemento.

Dichas relaciones pueden contener información adicional, permitiendo de esta manera generar relaciones causales entre cada elemento y de esta forma representar las reglas de producción dentro de la base de datos.

Al guardar las reglas de producción dentro de este tipo de bases de datos, permite que, cuando se realice una consulta, la información de las reglas pueda ser cargada y manipulada por el sistema más rápidamente, mejorando el tiempo de procesamiento y respuesta.



### 3.13.1 MongoDB

MongoDb es una base de datos no relacional basada en documentos, que permite almacenar información bajo un esquema y tener una alta consistencia de datos a través de los diferentes nodos que conforman las instancias de la base de datos.

Otra de sus características es su buen manejo de redundancia y reducción de errores. Algunos de sus beneficios son su escalabilidad, su facilidad para soportar grandes incrementos de datos, y se requiere de muy poco mantenimiento.

En la siguiente figura se puede observar una comparativa entre la estructura de las bases de datos relacionales y MongoDB.

<b>Base de datos Relacionales</b>	<b>Mongo DB</b>
Filas	Documento
Columna	Propiedades
Tabla	Colecciones

Tabla 3.13.1 Equivalencia de conceptos entre bases de datos relacionales y Mongo DB.

### 3.13.2 Neo4j

Neo4j es una base de datos no relacional, para el almacenamiento de grafos, la cual permite agregar propiedades a las diferentes partes de un grafo como los son: nodos y aristas,

Algunos de sus beneficios son el almacenamiento persistente, la independencia ya sea física o lógica, la integridad y la consistencia de datos [7].

A continuación, se presenta una tabla comparativa entre la estructura de una base de datos relacional y su equivalencia dentro de Neo4j:

<b>Base de Datos Relacionales</b>	<b>Base de Datos de Grafos</b>
Filas	Nodos
Columnas	Propiedades
Nombre de las Tablas	Etiquetas en Nodos/Aristas
Claves Foráneas	Aristas entre Nodos

Tabla 3.13.2 Equivalencia entre las bases de datos relacionales y base de datos basada en grafos.

El lenguaje de consultas que utiliza esta base de datos se llama CYPHER y permite realizar consultas a los grafos de una manera sencilla y práctica, su sintaxis está basada principalmente en instrucciones que contienen expresiones regulares.

## 3.14 Dominio

El dominio en términos de sistemas expertos se refiere a todo aquel espacio acotado en el que reside la base del conocimiento sobre la cual trabaja el sistema experto. Por lo que el espacio de búsqueda de información está limitado a un conjunto de temas muy específicos.

En gran medida el nivel del sistema experto está definido mediante el alcance del dominio, sin embargo, un dominio puede contener subdominios o temas que permitan enriquecer el conocimiento y facilitar que el sistema lleve a cabo el trabajo de inferencia de una manera más efectiva y segura.

## 3.15 Linux

Linux es un sistema operativo de código abierto, el cuál fue creado por múltiples desarrolladores para ser usado en equipos de cómputo, su desarrollo y código está basado en otro sistema operativo libre llamado UNIX.

Algunas de las características que diferenciaron a Linux de los sistemas de su tiempo y que siguen siendo aplicables, y otras heredadas de UNIX, podrían ser [12]:

- a) Sistema operativo de código abierto: cualquiera puede disponer de sus fuentes, modificarlas y crear nuevas versiones que poder compartir bajo la licencia GPL lo que lo convierte en un sistema operativo libre.
- b) Portabilidad: tal como el UNIX original, Linux está pensado para depender muy poco de una arquitectura concreta de máquina. Consecuentemente Linux es, en su mayor parte, independiente de la máquina de destino y puede portarse a casi cualquier arquitectura que disponga de un compilador C como el GNU gcc.
- c) Kernel de tipo monolítico: el diseño del kernel está unido en una sola pieza, pero es conceptualmente modular en las diferentes tareas. Linux se decidió como monolítico, porque es difícil extraer buen rendimiento de los microkernels.
- d) Módulos dinámicamente cargables: permiten poner partes del sistema operativo, como filesystems, o controladores de dispositivos, como porciones externas que se cargan (o enlazan) con el kernel en tiempo de ejecución bajo demanda. Esto permite simplificar el kernel y ofrecer estas funcionalidades como elementos que se pueden programar por separado. Con este uso de módulos, se podría considerar a Linux como un kernel mixto, ya que es monolítico, pero ofrece una serie de módulos que complementan el kernel.
- e) Desarrollo del sistema por una comunidad vinculada por Internet: los sistemas operativos nunca habían tenido un desarrollo tan amplio y disperso El fenómeno de la comunidad Linux permite que cada uno colabore en la medida que el tiempo y sus propios conocimientos se lo permitan.

Es por esto por lo que existen múltiples distribuciones de Linux, tanto de uso libre como de versiones con licenciamiento. Algunos ejemplos de estos son Fedora, Suse, Red Hat y Ubuntu.

### 3.16 Ubuntu

Ubuntu es una distribución GNU/Linux fácil la cual va orientada tanto al usuario de escritorio como al servidor. Y como se mencionó anteriormente, esta se encuentra mantenida por una comunidad de desarrolladores pertenecientes a una empresa la cual se encarga de distribuir y vender esta versión de Linux.

Sergio Blanco indica que actualmente Ubuntu soporta las arquitecturas: Intel x86 (IBM-compatible PC), AMD64 (Hammer) y PowerPC (Apple iBook y Powerbook, G4 y G5). En la distribución se incluyen más de 1000 paquetes que van desde el núcleo Linux hasta GNOME 2.12, cubriendo todas las aplicaciones necesarias para el escritorio, acceso a Internet, programación y servicios [13].

Ubuntu nació a partir de la distribución Debian. Esta es ampliamente conocida por su gestor de paquetes integrado que facilita la instalación de miles de aplicaciones de forma sencilla, rápida y eficiente. Sin embargo, Debian también presenta ciertos problemas en cuanto a su política de versionado. Debian ofrece versiones estables de la distribución, estas son altamente fiables y robustas, pero con aplicaciones bastante antiguas. Esto hace que los usuarios no puedan disfrutar cómodamente de los últimos avances en aplicaciones de escritorio o servidor [13].

Es por esto, que uno de los principales objetivos de Ubuntu es conseguir una distribución de escritorio que sea fácil de utilizar y accesible.

### 3.17 Servidor web

Un servidor Web es un programa que es capaz de interpretar código HTML por medio de un protocolo HTTP, dicho código generalmente se compone de textos, enlaces, multimedia y que son comúnmente conocidos como páginas web.

La arquitectura utilizada para un servidor web es cliente/servidor, es el cual, el equipo cliente hace una petición al equipo servidor, y éste atiende dicha solicitud. En el equipo cliente se ejecuta una aplicación que es conocida como “cliente web” y que sirve de interfaz con el usuario, proporcionando un conjunto de herramientas que facilitan su comunicación con el servidor.

En el equipo servidor la única tarea es atender las peticiones recibidas desde los navegadores o clientes web y hacerlo de forma eficiente y segura. Este es el caso de los servidores web seguros que solicitan un nombre de usuario y una contraseña para permitir el acceso sólo a usuarios registrados y por tanto, con permiso para visualizar la página/s.

## 3.18 Apache

El servidor HTTP Apache2 es un servidor web de software libre desarrollado por la Apache Software Foundation. El producto obtenido de este proyecto es un servidor de código fuente completo, descargable y gratuito.

Apache2 es robusto y con un ciclo de desarrollo muy rápido gracias a la gran cantidad de colaboradores voluntarios de que dispone. Es también un servidor estable, eficiente, extensible y multiplataforma [14]. Por lo que se dice que es:

**Estable:** es una consecuencia de su probada robustez que impide caídas o cambios en el servidor inesperados.

**Flexible y eficiente:** es capaz de trabajar con el estándar HTTP/1.1 y con la mayor parte de las extensiones web que existen en la actualidad, como son los módulos PHP, SSL, CGI, SSI, proxy, etc.

**Extensible:** dispone de gran cantidad de módulos que amplían su funcionalidad.

**Multiplataforma** ya que está disponible para diferentes plataformas como Linux, Windows, MacOS.

## 3.19 Python

Python es un lenguaje de programación de alto nivel, es decir las sintaxis que se suele usar es fácil de leer para un ser humano, a comparación de otros lenguajes como java y c++, ya que la filosofía del lenguaje es proporcionar una sintaxis muy limpia y que beneficie con código legible [15].

Es una de las herramientas tecnológicas que en los últimos años se ha hecho muy popular, gracias a varias razones:

- Contiene una amplia gama de librerías.
- Es posible crear programas de una forma rápida.
- Es multiplataforma ya que se puede desarrollar y ejecutar programas en varios sistemas operativos.
- Es libre de licencia desarrollo

El propósito de Python es que se pueden desarrollar todo tipo de programas como aplicaciones web o de escritorio, crear rutinas para servidores o servicios. Tiene una gran capacidad de procesamiento por lo que es utilizado frecuentemente en programas de inteligencia artificial.

Python es posible desarrollar aplicaciones sobre el paradigma de la programación orientado a objetos

En cuanto a la sintaxis, esta es muy clara es muy fácil de leer, gracias a su indentación de código lo cual evita errores que sean difíciles de encontrar, por lo que de esta manera es posible mantener el código en un futuro.

---

## 4 DESARROLLO METODOLÓGICO

---

Definir las áreas de trabajo y sus respectivas tareas o actividades mediante una metodología es de suma importancia para el desarrollo del presente trabajo, ya que marca los pasos a seguir para lograr construir una base de conocimiento, así como también de cómo llevar a cabo la implementación de un prototipo para consultas.

Por lo que en el presente capítulo se presenta la metodología utilizada para diseñar e implementar una base de conocimiento, que como se ha descrito en los capítulos previos, permitirá determinar la probabilidad del asma y su severidad en pacientes adultos, considerando sus antecedentes, causas, síntomas y predisposiciones.

Para el desarrollo de la base de conocimiento, se propuso definir y seguir una metodología basada en dos grandes módulos: ingeniería del conocimiento e implementación computable. Estos módulos son subdivididos en otros sub-módulos, los cuales permiten trabajar de una forma lógica y ordenada las actividades y tareas definidas.

La razón para tener esta división fue con la finalidad de tener 2 procesos distintos agrupados dentro de una misma metodología y dado que la implementación computable es dependiente de las actividades realizadas en el módulo de ingeniería del conocimiento, esto permite establecer una relación entre ambos módulos.

A continuación, se describe detalladamente cada uno de los procesos que componen la metodología, desde obtener el conocimiento del experto humano y su modelado, pasando por una revisión de las tecnologías usadas para la implementación de la base de conocimiento hasta el desarrollo de la interfaz gráfica que realiza la consulta a dicha base.

## 4.1 Metodología aplicada

La metodología propuesta es una guía de referencia para obtener y modelar conocimiento para el diagnóstico médico del asma. Esta se compone de varios módulos y sub.módulos los cuales agrupan funcionalidades muy específicas dentro del desarrollo del trabajo realizado.

La metodología establece una subdivisión del trabajo que realiza el ingeniero del conocimiento y el ingeniero de desarrollo de software, en donde dentro de la primera división se agrupan todas las tareas a realizar entorno a la adquisición del conocimiento, análisis y representación en términos de la construcción de la base de conocimiento y que van desde la selección de dominio, entrevistas con el experto hasta la evaluación de técnicas para la representación del conocimiento y la validación de este.

Con respecto a la segunda subdivisión del trabajo, correspondiente a al ingeniero de desarrollo de software en donde se engloban todas aquellas tareas relacionadas con la implementación del prototipo de consultas, como lo son la evaluación de tecnologías existentes, la instalación de herramientas necesarias para el desarrollo y el ciclo general de desarrollo de software.

En la figura 4.1 se presentan las tareas de cada subdivisión y que pertenecen a los dos módulos principales antes mencionados. La primera subdivisión enmarcada en color azul corresponde a la ingeniería del conocimiento, la cual detalla las actividades realizadas para adquirir el conocimiento del experto humano y las tareas para llevar a cabo la representación del conocimiento como las reglas de producción y mapas cognitivos difusos.

La segunda subdivisión enmarcada en color verde agrupa las tareas que se deben realizar para la implementación de la base de conocimiento, el desarrollo del motor de búsqueda y la interfaz de usuario.

En cada agrupación dentro de las subdivisiones se incorpora una tarea de validación y verificación de cada una de las tareas realizadas con la finalidad de asegurar la calidad y asertividad del desarrollo.

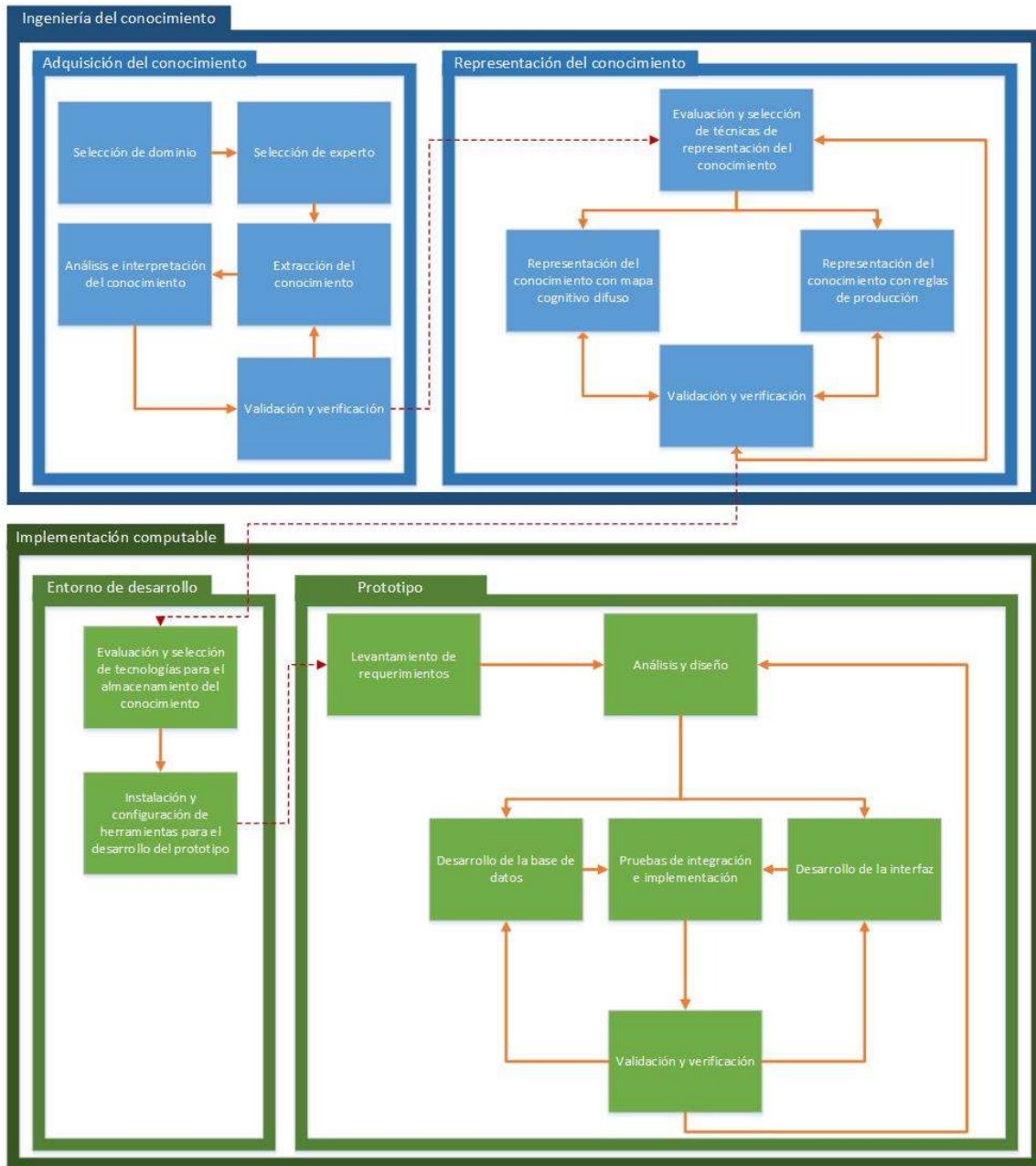


Fig. 4.1 Metodología aplicada para el diseño e implementación de una base de conocimiento.

## 4.2 Ingeniería del conocimiento

Para llevar a cabo la ingeniería del conocimiento, se realizaron una serie de tareas orientadas a la obtención del conocimiento por parte del experto humano, llevando a cabo primeramente la selección tanto del dominio, como del experto, para después adquirir el conocimiento necesario para el diagnóstico del asma,

sirviéndonos de algunas técnicas como la aplicación de entrevistas y cuestionarios para extraer el conocimiento.

Y una vez analizada la información obtenida y evaluadas las diferentes técnicas de representación, se dio a la tarea de representar el conocimiento del experto.

## 4.2.1 Adquisición del conocimiento

### 4.2.1.1 Selección del dominio

Dentro del campo de la medicina, la selección de dominio puede ser una tarea difícil cuando se requiere desarrollar una base de conocimiento, principalmente porque el conocimiento médico puede ser amplio y delicado, sobre todo para diagnóstico de enfermedades o padecimientos. Inicialmente se pensó en utilizar bases de conocimiento del campo médico existentes, sin embargo, en la búsqueda que se hizo se encontró que son recursos no públicos y no gratuitos. Es por esta razón que se decide crear una base de conocimiento para el diagnóstico del asma en adultos, y aunque este proceso puede ser confuso y complejo por su similitud con otras enfermedades, al final puede concretarse un diagnóstico asertivo, si se construye una base de conocimiento sólida y una herramienta de consulta que la explote.

Sin ser objeto de este trabajo los estudios de laboratorio también son importantes, ya que estos complementan y confirman el diagnóstico obtenido a través de la herramienta.

### 4.2.1.2 Selección de experto

Un elemento clave en el proceso de adquisición de la información es el experto humano. Su rol es fundamental para lograr comprender cuáles son los elementos que conjuga para realizar un diagnóstico del padecimiento. Para este trabajo la selección del experto implicó un acercamiento con un médico general, que frecuentemente atiende a pacientes con asma en sus consultas médicas y quien abiertamente compartió sus saberes, experiencias e información, además de facilitar la metodología que sigue para determinar el asma en adultos y la forma en la que se lleva a cabo el diagnóstico en un paciente.

### 4.2.1.3 Extracción del conocimiento

Los instrumentos que se utilizaron para la recopilación de la información fueron definidos con base en a la metodología que sigue el médico con sus pacientes, considerando cual era la manera óptima de extraer el conocimiento necesario para el diagnóstico del asma. Es por esta razón que se decidió aplicar entrevistas dirigidas y cuestionarios al experto humano, para aprovechar mejor el tiempo entre el médico y el



ingeniero del conocimiento y obtener información suficiente. Para conformar las reglas de producción necesarias para determinar cuando el paciente no presenta un caso de asma.

#### 4.2.1.3.1 Entrevistas con el experto humano

Una vez establecido el dominio y elegido el experto humano que colaboraría con el proyecto, el siguiente paso fue llevar a cabo una serie de entrevistas dirigidas, con la finalidad de conocer cómo es el procedimiento que se aplica durante la consulta de un paciente cuando este presenta un caso clínico de asma.

Para la entrevista guiada, se preparó un documento con una serie de preguntas abiertas y cerradas, buscando que las respuestas fueran lo más precisas para el análisis e interpretación de la información obtenida.

Gracias a estas entrevistas se logró recopilar la información necesaria para diagnosticar a un paciente con asma.

Entre la información recopilada, se encuentra la experiencia del médico y como es que se puede llegar a un diagnóstico asertivo mediante algunas preguntas claves, cuyas respuestas en su conjunto pueden definir la severidad del asma que pudiera estar presente en un paciente.

#### 4.2.1.3.2 Elaboración y aplicación de cuestionarios

Los cuestionarios realizados fueron la principal herramienta que se utilizó para adquirir el conocimiento del experto humano, y el mejor instrumento para validar la información recabada por el ingeniero del conocimiento con relación a los diferentes elementos que forman parte del diagnóstico.

Los cuestionarios fueron diseñados y definidos en conjunto con el experto humano, para organizar las preguntas que el médico sigue en el diagnóstico del asma, durante la consulta. Por lo que se creó un cuestionario con preguntas para el diagnóstico del asma en adultos, las cuales se encuentran disponibles en el Anexo 1.3.

Las preguntas contenidas en los cuestionarios aumentan o disminuyen la probabilidad de presentar asma, ya que la combinación de las diferentes respuestas es necesaria para el diagnóstico, éstas deberán de ser mapeadas en valores cualitativos o cuantitativos que permitan modelar el conocimiento de una manera más eficiente e incrementando el asertividad de las consultas.

Es por esto por lo que se elaboró un cuestionario con un formato de tabla en donde se le indico al médico asignara un valor 1 al 100 para representar la relevancia para cada elemento en relación con los demás elementos de diagnóstico, como se puede observar en el Anexo 1.4.

#### 4.2.1.4 Análisis e interpretación del discurso del experto

Dentro de este proceso, se realizó un análisis de la información recolectada, derivada de las entrevistas y cuestionarios, que permitió identificar los conceptos y las relaciones semánticas entre sí, así como definir los valores cualitativos y cuantitativos que se determinan en la valoración médica del paciente.

Para llevar a cabo el proceso de interpretación del discurso del experto, por lo que se realizó una caracterización del discurso médico en donde se clasificaron los elementos que componen el diagnóstico dentro de niveles de abstracción, para de esta manera, establecer valores cuantitativos a cada uno de los elementos de diagnóstico.

Una vez realizada dicha caracterización, se validó y verificó con el experto si los valores establecidos para cada elemento reflejaban correctamente el valor dentro del diagnóstico.

##### 4.2.1.4.1 Caracterización del discurso médico

El proceso mediante el cual el médico realiza el diagnóstico del asma se basa principalmente en el cuestionamiento al paciente sobre la presencia de los elementos característicos del asma, en donde el experto humano pondera la probabilidad de asma, dependiendo de la cantidad de respuestas positivas dadas por el paciente.

Por lo que el médico evalúa todas las respuestas de cada uno de los elementos que están relacionados con el asma, al obtener respuestas afirmativas, este incrementa la asertividad del diagnóstico y aunque el orden en las preguntas puede variar, esto no afecta el diagnóstico final dado que todos los elementos presentes le permiten al médico evaluar la probabilidad de la presencia del asma.

En el análisis se identificaron 4 principales conceptos, los cuales se categorizaron como niveles de abstracción: Antecedentes, Predisposiciones, Causas y Síntomas, como se puede observar en la figura 4.2.1.4.1, cuya clasificación es usada por Aldana y Velázquez [1], en su tesis como parte de la caracterización del discurso, misma que es representada mediante la teoría de conjuntos. De acuerdo con la propuesta de estos autores, las relaciones para el diagnóstico se agrupan e identifican de la siguiente manera:

A=Antecedentes      P=Predisposiciones      C=Causas      S=Síntomas

Por lo que el diagnóstico (D) se define así:

$$D = A \cup P \cup C \cup S$$

Fig. 4.2.1.4.1 Fórmula para el diagnóstico del asma.

<b>Diagnóstico del Asma</b>				
Niveles de Abstracción				
<b>Antecedentes</b>	<b>Predisposiciones</b>	<b>Causas</b>	<b>Síntomas</b>	<b>Diagnóstico (cualitativo)</b>
Padres asmáticos	Tabaquismo	Daño pulmonar	Tos	Asma severo
Alergias	Infecciones virales	Hipertensión	Hiperventilación	Asma moderado – severo
Obesidad	Contacto con animales	Sibilancias	Espasmos pulmonares	Asma moderado
Contacto con polen	Ansiedad	Disnea	Erge	Asma leve – moderado
Contacto con polvo	Sedentarismo	Taquicardia	Fiebre	Asma leve

Tabla 4.2.1.4.1 Elementos para el diagnóstico del asma.

A partir de esta clasificación las respuestas a cada una de las preguntas contenidas en los cuestionarios fueron mapeadas y ponderadas con valores cualitativos y cuantitativos en conjunto con el experto humano, con la finalidad de facilitar la interpretación del discurso.

Cada nivel de abstracción agrupa los elementos de diagnóstico que mantienen una fuerte cohesión entre sí y una clara relación semántica entre ellos. Dentro de dichos niveles se clasificaron los 20 elementos más críticos para el diagnóstico, permitiendo contar con los 5 elementos más significativos dentro de cada nivel de abstracción.

Por otro lado, las ponderaciones tanto cualitativas como cuantitativas de cada uno de los elementos fueron establecidas de acuerdo con el tipo de respuesta que se espera obtener en la evaluación médica del paciente y a la severidad del diagnóstico. En la siguiente tabla se puede observar el mapeo de los elementos de diagnóstico y sus ponderaciones.

Nivel de abstracción	Elementos de diagnóstico	Ponderación cuantitativa
Antecedentes	Padres asmáticos	Sí = 1, No = 0
	Alergia	Sí = 1, No = 0
	Obesidad	Sí = 1, No = 0
	Contacto con polen	Sí = 1, No = 0
	Contacto con polvo	Sí = 1, No = 0
Predisposiciones	Tabaquismo	Sí = 1, No = 0
	Infecciones virales	Sí = 1, No = 0
	Contacto con animales	Sí = 1, No = 0
	Ansiedad	Sí = 1, No = 0
	Sedentarismo	Sí = 1, No = 0
Causas	Daño pulmonar	Sí = 1, No = 0
	Hipertensión	Sí = 1, No = 0
	Sibilancias	Sí = 1, No = 0
	Disnea	Sí = 1, No = 0
	Taquicardia	Sí = 1, No = 0

Síntomas	Tos	Sí = 1, No = 0
	Hiperventilación	Sí = 1, No = 0
	Espasmos pulmonares	Sí = 1, No = 0
	Erge	Sí = 1, No = 0
	Fiebre	Sí = 1, No = 0

Tabla 4.2.1.4.2 Elementos de diagnóstico ponderados y agrupados por nivel de abstracción.

El experto humano, obtiene la respuesta para cada uno de los elementos al momento de realizar las preguntas al paciente, es importante resaltar que el médico no siempre realiza todas las preguntas del cuestionario en el mismo orden, ya que el orden no influye en el diagnóstico.

#### 4.2.1.4.2 Validación y verificación

Una vez analizada la información compilada dentro de la tabla de agrupación por nivel de abstracción, se validó con el experto humano si la ponderación cuantitativa asignada a cada elemento, estaban representadas correctamente para poder llevar a cabo el diagnóstico del asma.

La forma en la que se verificó si estos valores eran correctos, fue mediante la simulación de una consulta entre el médico y un paciente, asignado los valores correspondientes a cada elemento dependiendo de las respuestas obtenidas durante el interrogatorio médico.

Comprobando así que se podía llegar a un diagnóstico del asma basado en los valores cuantitativos de cada elemento.

#### 4.2.1.5 Evaluación y selección de técnicas de representación del conocimiento

Este proceso de evaluación y selección de técnicas de representación del conocimiento estuvo sujeto a 2 tipos de modelos: un modelo basado en reglas de producción y otro modelo representado con mapas cognitivos difusos. Ambos modelos funcionan de forma conjunta uno como complemento del otro en el diagnóstico del asma. De esta manera se aprovecha el poder de representación que tiene cada modelo, tratando de tomar las ventajas que ofrecen y cubrir las limitaciones de uno con el otro.

##### 4.2.1.5.1 Representación del conocimiento con reglas de producción

La creación de las reglas de producción se basó en la información obtenida durante el proceso de análisis e interpretación del discurso médico a través de los instrumentos implementados. Su construcción consistió en definir un conjunto de sentencias lógicas de cada una de las relaciones establecidas entre

antecedentes, predisposiciones, causas y síntomas y su efecto negativo en el diagnóstico final, es decir, se realizó el mapeo de cada uno de los elementos involucrados en el diagnóstico que dan como resultado un diagnóstico sin presencia de asma en el paciente.

Las reglas de producción se escriben como sentencias de la forma Si ... Entonces y en cuya estructura incluye el antecedente y el consecuente.

La siguiente tabla muestra la conformación de las reglas de producción que determinan la ausencia de asma en el paciente.

Reglas	Evaluación
AR1= Si el paciente tiene algún padre asmático entonces tiene antecedentes de padre(s) asmático(s).	Si SR5 entonces Asma=No Si SR4 entonces Asma=No
AR2= Si el paciente tiene alergias entonces es alérgico.	Si AR3=1 entonces Si PR5=1 entonces Si CR2=1 entonces
AR3= Si el paciente tiene obesidad entonces obesidad.	Si SR1=1 entonces Asma=No
AR4= Si el paciente tiene contacto constantemente con polvo entonces puede padecer asma.	Si AR3=1 entonces Si PR1=1 entonces Si CR1=1 entonces
AR5= Si el paciente está en contacto con el polen entonces puede padecer asma.	Si SR1=1 entonces Asma=No
PR1= Si el paciente fuma diariamente entonces padece tabaquismo.	Si AR3=1 entonces Si PR2=1 entonces Si CR2=1 entonces Si SR1=1 entonces Asma=No
PR2= Si el paciente tiene infecciones virales entonces puede padecer asma.	Si AR3=1 entonces Si PR2=1 entonces Si CR2=1 entonces Si SR2=1 entonces Asma=No
PR3= Si el paciente tiene contacto constantemente con epitelio de animales entonces puede padecer asma.	Si AR3=1 entonces Si PR3=1 entonces Si CR5=1 entonces Si SR1=1 entonces Asma=No
PR4= Si el paciente no hace ejercicio entonces es sedentario.	Si AR3=1 entonces Si PR5=1 entonces Si CR2=1 entonces Si SR2=1 entonces Asma=No
PR5= Si el paciente tiene ansiedad crónica entonces puede padecer asma.	Si AR3=1 entonces Si PR5=1 entonces

<p>CR1= Si el paciente tiene problemas en los pulmones entonces puede tener daño pulmonar.</p> <p>CR2= Si el paciente toma medicamentos para la presión entonces presenta hipertensión.</p> <p>CR3= Si el paciente tiene presión arterial mayor a 140/100 mmHg por varios días seguidos entonces presenta hipertensión.</p> <p>CR4= Si el paciente al respirar se escucha un silbido entonces presenta sibilancias.</p> <p>CR5= Si el paciente no puede respirar fácilmente entonces presenta disnea.</p> <p>CR6= Si el paciente presenta latidos del corazón rápidos no correspondientes a su edad entonces presenta taquicardia.</p> <p>CR7= Si el paciente presenta latidos del corazón rápidos no correspondientes a su actividad entonces presenta taquicardia.</p> <p>SR1= Si el paciente tose entonces presenta tos.</p> <p>SR2= Si el paciente inhala y exhala rápidamente entonces presenta hiperventilación.</p> <p>SR3= Si el paciente no puede respirar con facilidad entonces presenta espasmo pulmonar.</p> <p>SR4= Si el paciente presenta irritación de estómago por bilis entonces presenta Erge.</p> <p>SR5= Si el paciente tiene una temperatura mayor a 38 °C entonces tiene fiebre</p>	<p>Si CR5=1 entonces Si SR1=1 entonces Asma=No</p> <p>Si AR3=1 entonces Si PR5=1 entonces Si CR5=1 entonces Si SR2=1 entonces Asma=No</p> <p>Si AR3=1 entonces Si PR5=1 entonces Si CR5=1 entonces Si SR3=1 entonces Asma=No</p>
---	--

Tabla 4.2.1.5.1 Reglas de producción para el diagnóstico de no asma.

Para facilitar la evaluación de reglas, cada una fue etiquetada con un identificador según su categoría y el número de regla que le corresponde dentro del conjunto. Por ejemplo, AR3, representa la regla número 3 de la categoría Antecedentes.

A estas reglas se le aplica un formato JSON, previamente a ser almacenadas dentro de una colección en MongoDB, la cual, al ser una base de datos no relacional para almacenamiento de documentos, permite almacenar datos con dicho formato, conformando de esta manera parte de la base de conocimiento y que por sí sola indican ausencia de asma.

Si los signos que presenta el paciente no cumplieran con ninguna de las reglas mencionadas con anterioridad, entonces, se hace uso de los mapas cognitivos difusos para determinar la probabilidad de asma, así como su grado de severidad en un paciente.

#### 4.2.1.5.2 Representación del conocimiento con mapa cognitivo difuso

El modelado del conocimiento con mapas cognitivos difusos es una representación más gráfica que textual, el modelo es un grafo dirigido con nodos que representan los elementos de diagnóstico del asma y las aristas representan las relaciones entre ellos. Este grafo además de ser dirigido (tiene sentido y dirección), es etiquetado, lo que significa que cada arista tiene un peso asociado, cuyo valor representa los elementos que conforma el diagnóstico del asma.

Para construir el grafo, fue necesario identificar los principales elementos que influyen directamente en el diagnóstico del asma, excluyendo aquellos síntomas que determinan un diagnóstico negativo, como lo es el Erge o la Fiebre, mismos que son descartados del grafo, ya que estos están incorporados dentro de las reglas de producción previamente descritas.

Una vez mapeados los elementos de diagnóstico y establecidas las relaciones que dan estructura al mapa, se definieron valores difusos para la asignación de pesos ( $w$ ). Esta se determinó con base a la importancia de la relación que tiene cada elemento entre sí con los demás elementos y que en su conjunto puede ayudar a calcular la probabilidad de padecer asma. A continuación, se presenta la tabla de pesos correspondiente a los elementos de diagnóstico seleccionados para este estudio, y la cual conforma la matriz de adyacencia con la que se construyó el grafo. el que representa el conocimiento mediante un mapa cognitivo difuso.

	Padres asmáticos	Alergia	Obesidad	Contacto con polvo	Contacto con polen	Tabaquismo	Infecciones virales	Contacto con animales	Ansiedad	Sedentarismo	Daño pulmonar	Hipertensión	Sibilancias	Disnea	Taquicardia	Tos	Hiperventilación	Espasmo pulmonar
Padres asmáticos	0	.9	.8	.9	.9	.8	.6	.8	.8	.8	.9	.7	.9	.8	.7	.8	.9	.9
Alergia	.9	0	.6	.7	.7	.6	.5	.3	.7	.3	.7	.3	.5	.6	.3	.3	.6	.6
Obesidad	.8	.6	0	.7	.7	.6	.3	.7	.9	0	.7	0	.7	.6	0	.3	.3	.8
Contacto con polvo	.9	.7	.7	0	.7	.7	.3	.6	.8	.3	.3	0	.7	.8	.7	.9	.9	.9
Contacto con polen	.9	.7	.7	.7	0	.7	.3	.3	.6	0	0	0	0	.6	.3	.7	.8	.7
Tabaquismo	.8	.6	.6	.7	.7	0	0	.5	.7	0	0	0	.3	.3	0	0	.7	.5
Infecciones virales	.6	.5	.3	.3	.3	0	0	0	.3	0	.5	0	.3	.3	.3	0	.3	.3
Contacto con animales	.8	.3	.7	.6	.3	.5	0	0	.8	0	.3	0	.3	.5	0	.3	.8	.8
Ansiedad	.8	.7	.9	.8	.6	.7	.3	.8	0	.5	.6	0	.7	.9	.3	.3	.8	.8
Sedentarismo	.8	.3	0	.3	0	0	0	0	.5	0	0	0	.7	.5	0	.3	.6	.7
Daño pulmonar	.9	.7	.7	.3	0	0	.5	.3	.6	.6	0	0	.5	.5	.5	0	.7	.7
Hipertensión	.7	.3	0	0	0	0	0	0	0	0	0	0	.8	.8	0	0	.5	.7
Sibilancias	.9	.5	.7	.7	0	.3	.3	.3	.7	.7	.5	.8	0	.9	.8	.7	.9	.8
Disnea	.8	.6	.6	.8	.6	.3	.3	.5	.9	.9	.5	.8	.9	0	.7	.7	.8	.8
Taquicardia	.7	.3	0	.7	.3	0	.3	0	.3	.3	.5	0	.8	.7	0	0	.8	.8
Tos	.8	.3	.3	.9	.7	0	0	.3	.3	.3	0	0	.7	.7	0	0	.8	.9
Hiperventilación	.9	.6	.3	.9	.8	.7	.3	.8	.8	.8	.7	.5	.9	.8	.8	.8	0	.9
Espasmo pulmonar	.9	.6	.8	.9	.7	.5	.3	.8	.8	.8	.7	.7	.8	.8	.8	.9	.9	0

Tabla 4.2.1.5.2.1 Tabla de pesos asignados a cada elemento de diagnóstico.



Para asignar los pesos correctos se solicitó la ayuda al experto humano, quien se encargó de determinar los porcentajes de importancia para cada relación, basándose en un rango de valores de [0,1], donde los valores indican el grado de impacto o influencia de cada uno de los elementos presentes dentro de cada nivel de abstracción (antecedentes, predisposiciones, causas y síntomas), los cuales pueden aumentar la posibilidad de un diagnóstico positivo. El valor de 0 indica que no existe ninguna relación.

Estos valores fueron validados por el médico experto mediante un ejercicio de diagnóstico, para verificar de esta manera, si los valores difusos asignados a cada relación representaban de manera correcta las relaciones entre cada elemento.

Para la elaboración del mapa cognitivo difuso se utilizó una base de datos en Neo4J, la cual permite almacenar grafos directamente, así como la asignación de pesos en cada una de las aristas, por lo que se creó el mapa conceptual directamente en ella, permitiendo así al motor de búsqueda realizar consultas directamente.

Tomando los valores difusos obtenidos de la tabla de pesos por cada elemento descrito en la tabla 4.1.2, se creó el mapa conceptual de la figura 4.2.1.5.2.

Las relaciones taxonómicas para cada uno de los nodos pertenecientes a cada uno de los niveles de abstracción ya sean Antecedentes, Predisposiciones, Causas y Síntomas se representan en agrupaciones de cuatro colores dentro de este mismo grafo. Dichos grupos contribuyen a la estimación de la severidad del asma, ya que cada uno de estos representa un 25% del diagnóstico de severidad.

Por lo que tomando en cuenta los valores obtenidos de los pesos definidos para cada concepto en la matriz de adyacencia mencionada en el párrafo anterior, se define una tabla de parámetros lingüísticos en donde se le da significado a cada uno de estos parámetros mediante los valores difusos. En la tabla que se muestra a continuación, se presentan los valores para los parámetros lingüísticos:

Parámetros	Valores Difusos
+ Severa	$0.7 < \text{variable} < 1$
+ Moderado-Severa	$0.6 < \text{variable} < 0.7$
+ Moderado	$0.5 < \text{variable} < 0.6$
+ Leve Moderado	$0.3 < \text{variable} < 0.5$
+ Leve	$0.05 < \text{variable} < 0.3$
Ausente	0

Tabla 4.2.1.5.2.2 Valores difusos para el diagnóstico del asma

Todos aquellos conceptos o elementos que pueden ser síntomas que cancelan por completo el diagnóstico del asma, no son considerados como parte del modelo del presente mapa cognitivo difuso, ya que la finalidad del presente modelo es solo definir la probabilidad y severidad del asma cuando está presente en el paciente.

En la siguiente imagen se puede observar el mapa cognitivo difuso creado para el diagnóstico de la probabilidad de que una persona padezca asma.

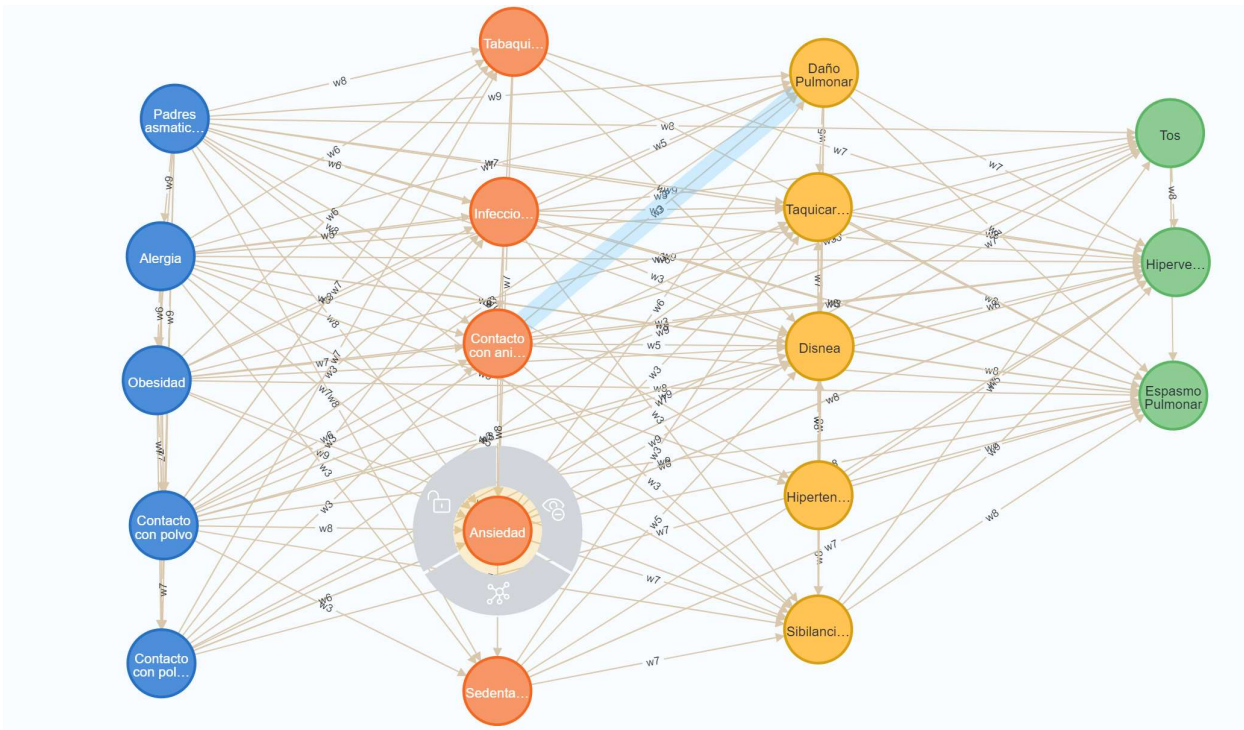


Fig. 4.2.1.5.2.1 Mapa cognitivo difuso.

#### 4.2.1.6 Validación y verificación

Una vez realizadas la extracción, el análisis e interpretación de la información, se llevó a cabo un proceso de verificación y validación con el médico, para revisar si la información recolectada era suficiente para el modelado del conocimiento.

Con base en la tabla con la caracterización de los elementos involucrados en el diagnóstico del asma, se validaron los valores cualitativos y cuantitativos con el experto, con la finalidad de avalar que dichas ponderaciones correspondían efectivamente a cada una de las respuestas a las preguntas establecidas dentro del cuestionario, mismas que van dirigida a dichos elementos del diagnóstico.

Después de tener el visto bueno por parte del médico, se inició con el proceso de representación, el cual permite modelar el conocimiento obtenido del experto humano, resultado del análisis e interpretación del discurso.

## 4.3 Implementación

### 4.3.1 Entorno de desarrollo

#### 4.3.1.1 Evaluación y selección de tecnologías para el almacenamiento de conocimiento

Para almacenar el conocimiento, fue necesario explorar diferentes tecnologías y realizar un análisis para determinar cuál de ellas, era la más adecuada, priorizando la flexibilidad para almacenar los datos y el rendimiento al momento de hacer la consulta de información a la base de datos. Los manejadores de bases de datos que se revisaron y se analizaron sus ventajas, se listan a continuación.

Base de datos	Tipo	Ventajas	Desventajas
MS SQL	Relacional	<ul style="list-style-type: none"><li>• Mantiene la integridad de los datos.</li><li>• Manejo de transacciones.</li><li>• Tiene soporte.</li><li>• Almacenamiento de datos de forma estructurada y relacional</li></ul>	<ul style="list-style-type: none"><li>• Difícil manejo de estructuras de datos dinámicas.</li><li>• Bajo rendimiento en el manejo de información masiva.</li><li>• Requiere licenciamiento.</li></ul>
MySQL	Relacional	<ul style="list-style-type: none"><li>• Manejo de distintos motores.</li><li>• Agrupación de transacciones.</li><li>• Almacenamiento de datos de forma estructurada y relacional</li></ul>	<ul style="list-style-type: none"><li>• Difícil manejo de estructuras de datos dinámicas.</li><li>• Bajo rendimiento en el manejo de información masiva.</li><li>• No tiene soporte</li></ul>
MongoDB	No Relacional	<ul style="list-style-type: none"><li>• Flexibilidad en almacenamiento de datos con diferentes estructuras.</li><li>• Alto rendimiento en manejo de información masiva.</li><li>• Fácil integración con APIs.</li></ul>	<ul style="list-style-type: none"><li>• No asegura la integridad de los datos.</li></ul>
Neo4j	No Relacional	<ul style="list-style-type: none"><li>• Infraestructura nativa para el almacenamiento de grafos.</li><li>• Alto desempeño en el manejo de información masiva.</li><li>• Permite agregar metadatos a los nodos y aristas.</li></ul>	<ul style="list-style-type: none"><li>• Uso de CLIP como lenguaje de consultas.</li><li>• No mantiene la integridad de los datos.</li></ul>
Cassandra	No Relacional	<ul style="list-style-type: none"><li>• Permite almacenamiento de datos relacionales</li><li>• Fácil manejo de estructuras de dato dinámicas.</li><li>• Buen rendimiento para manejo de información masiva.</li></ul>	<ul style="list-style-type: none"><li>• No permite el almacenamiento de grafos.</li><li>• No permite el manejo de estructuras de clave/valor anidadas.</li></ul>

Tabla 4.3.1.1 Comparativo de base de datos para el almacenamiento de conocimiento.

Después de analizar las diferentes ventajas y desventajas de cada una de las bases de datos, se seleccionó MongoDB como mejor opción para almacenar las reglas de producción ya que al permitir almacenar documentos en formato JSON, ofrece una gran flexibilidad para el manejo del conocimiento en cuanto a reglas de producción se refiere.

La segunda base de datos seleccionada fue la de Neo4j, ya que permite almacenar grafos y hacer consultas sobre estos de una manera más ágil, por lo que es idónea para almacenar el mapa cognitivo difuso que representa el conocimiento del experto.

Estas dos bases de datos permiten almacenar completamente los datos en la base de conocimiento y trabajar de forma conjunta en la consulta de la información para el diagnóstico médico del asma y su severidad.

### 4.3.1.2 Instalación y configuración de herramientas para el desarrollo del prototipo

El desarrollo del prototipo requirió instalar diversos entornos de desarrollo que permitieran codificar las instrucciones para el almacenamiento de la base de conocimiento y diseñar la interfaz de usuario.

<b>Componentes del prototipo</b>	<b>Herramientas</b>
Base de datos.	Instalación de Neo4j Instalación de MongoDB
Manejador de base de datos.	Neo4j Desktop MongoDB Compass
Interfaz gráfica de usuario.	Visual Studio Code Python 3.5

Tabla 4.3.1.2 Herramientas utilizadas para el desarrollo del prototipo.

## 4.3.2 Prototipo

### 4.3.2.1 Levantamiento de requerimientos.

El levantamiento de requerimientos del sistema se llevó a cabo con la finalidad de conocer cuáles son las funcionalidades mínimas que se esperan en la plataforma de consulta.

Se utilizó el estándar de la IEEE-830 para el levantamiento de dichos requerimientos. El documento completo que recolecta los requerimientos funcionales del sistema se encuentra en el Apéndice B.

Una vez definidos los requerimientos del prototipo se prosiguió a realizar el análisis y diseño del sistema.

### 4.3.2.2 Análisis y diseño

La arquitectura propuesta para el desarrollo del presente trabajo está compuesta por una página web que sirve como interfaz de usuario; una API la cual procesa la información proveniente de la interfaz gráfica y que funciona como backend y por último la base de conocimiento, la cual está compuesta por 2 bases de datos, las cuales son consumidas por la API. A continuación, se presenta dicha arquitectura:

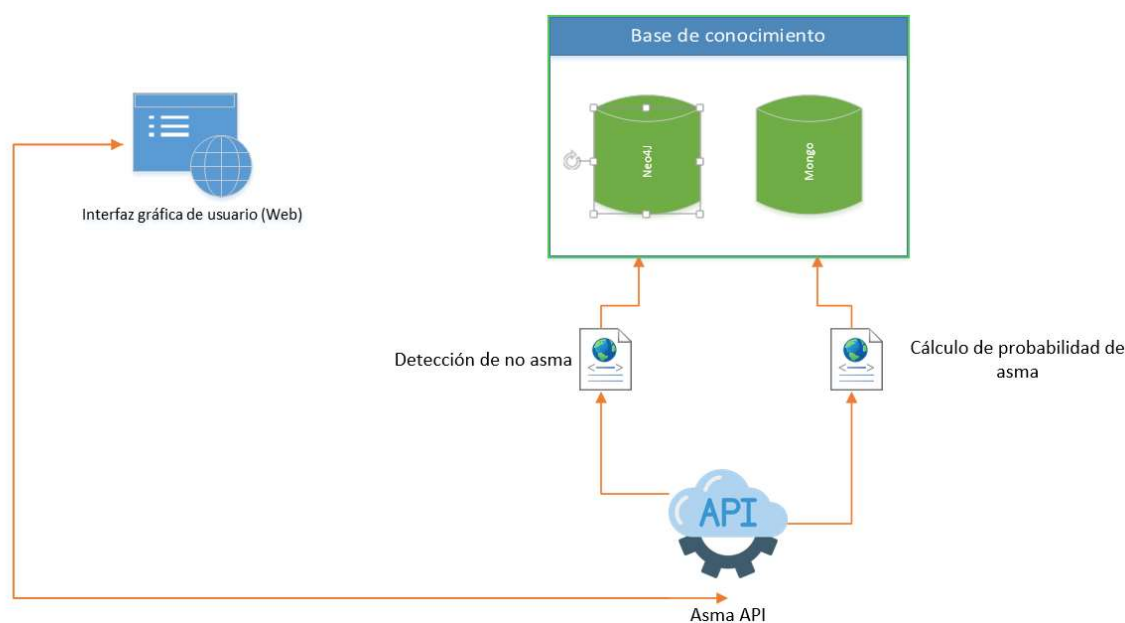


Fig. 4.3.2.2 Arquitectura propuesta para el desarrollo del prototipo.

### 4.3.2.3 Desarrollo de la base de datos

#### 4.3.2.3.1 Almacenamiento de las reglas de producción

Las reglas de producción obtenidas son almacenadas en una base de datos no relacional en formato JSON, en donde las propiedades del objeto que se mapea son los valores correspondientes a los antecedentes, predisposiciones, causas y síntomas, es decir los elementos de diagnóstico.

Para poder almacenar correctamente las reglas de producción, se creó primeramente un catálogo con la información completa de cada uno de los elementos de diagnóstico, donde se asigna una clave alfanumérica a cada uno de ellos, misma que es utilizada al momento de almacenar las reglas de producción como se puede observar en la siguiente imagen:

```

[
  {
    "Nombre" : "Padres Asmaticos",
    "Tipo" : "Antecedente",
    "Clave": "A1"
  },
  {
    "Nombre" : "Alergia",
    "Tipo" : "Antecedente",
    "Clave": "A2"
  },
  {
    "Nombre" : "Obesidad",
    "Tipo" : "Antecedente",
    "Clave": "A3"
  },
  {
    "Nombre" : "Contacto con Polvo",
    "Tipo" : "Antecedente",
    "Clave": "A4"
  },
  {
    "Nombre" : "Contacto con Polen",
    "Tipo" : "Antecedente",
    "Clave": "A5"
  },
  {
    "Nombre" : "Tabaquismo",
    "Tipo" : "Predisposicion",
    "Clave": "P1"
  },
  {
    "Nombre" : "Infecciones Virales",
    "Tipo" : "Predisposicion",
    "Clave": "P2"
  },
  {
    "Nombre" : "Contacto con Animales",
    "Tipo" : "Predisposicion",
    "Clave": "P3"
  },
  {
    "Nombre" : "Ansiedad",
    "Tipo" : "Predisposicion",
    "Clave": "P4"
  },
  {
    "Nombre" : "Sedentarismo",
    "Tipo" : "Predisposicion",
    "Clave": "P5"
  },
],

```

Fig. 4.3.2.3.1.1 Catálogo de elementos que pueden indicar ausencia de asma.

Un ejemplo de un objeto que representa algunas reglas de producción dentro de la base de datos se puede observar en la siguiente imagen:

```

{
  "No Asma": [
    {
      "Regla": ["A3", "P5", "C2", "S1"]
    },
    {
      "Regla": ["A3", "P1", "C1", "S1"]
    },
    {
      "Regla": ["A3", "P2", "C2", "S1"]
    },
    {
      "Regla": ["A3", "P2", "C2", "S2"]
    },
    {
      "Regla": ["A3", "P3", "C5", "S1"]
    },
    {
      "Regla": ["A3", "P5", "C2", "S2"]
    },
    {
      "Regla": ["A3", "P5", "C5", "S1"]
    },
    {
      "Regla": ["A3", "P5", "C5", "S2"]
    },
    {
      "Regla": ["A3", "P5", "C5", "S3"]
    }
  ]
}

```

Fig. 4.3.2.3.1.2 Representación de las reglas de producción dentro de la base de datos.

#### 4.3.2.3.2 Almacenamiento del mapa cognitivo difuso

El grafo que representa el mapa cognitivo difuso se almacena en la base de datos de Neo4j, la cual es una base de datos no relacional que permite la manipulación de grafos.

Para crear los nodos que componen el grafo, se utilizó el comando de Neo4j, en el siguiente ejemplo se puede observar el comando CREATE

**CREATE (AR2: Antecedente {name:'Alergia'})**

En este comando permite crear un nodo con el nombre de **Alergia** perteneciente a la clasificación de **Antecedente**. El AR2 es la clave alfanumérica que identifica al elemento dentro de la base de datos.

Una vez creados todos los nodos se crean las aristas con el siguiente comando, el cual permite agregar metadatos a las relaciones, en cuyo caso el peso es agregado como un metadato y la etiqueta que muestra la arista, es el valor establecido en el peso multiplicado por 10 dado que no está permitido tener caracteres como el “.” como nombre de etiquetas dentro de Neo4j, a continuación, se muestra un ejemplo de la sentencia para la creación de una relación:

```
MATCH (CR5: Causa {name:'Taquicardia'}) MATCH (SR2: Sintoma {name:'Hiperventilacion'})  
CREATE (CR5)-[cr5sr2: w80 {w:.8}]->(SR2);
```

En el ejemplo anterior se muestra la creación de una relación entre una causa y un síntoma.

#### 4.3.2.4 Desarrollo de interfaz

Una vez almacenada la información en la base de conocimiento dentro de las bases de datos se desarrolló la interfaz de usuario que se encargará de realizar las consultas a las respectivas bases de datos para obtener un diagnóstico de asma.

La interfaz permite al usuario ingresar los datos que son necesarios para hacer la evaluación de la condición del paciente.

Dicha interfaz está compuesta de cinco secciones para ingresar datos del paciente las cuales son: Datos Generales, Antecedentes, Predisposiciones, Causas, Síntomas, Confirmación y Diagnóstico. A continuación, se presentan las imágenes de cada una de las secciones que componen el sistema:

Nombre	JUAN JOSE
Apellido Paterno	TORIS
Apellido Materno	PALACIOS
Sexo	<input checked="" type="radio"/> Masculino <input type="radio"/> Femenino
Fecha de Nacimiento	05/04/1983
Edad	38
Tipo de Sangre	A+

[Siguiete](#)

Fig. 4.3.2.4.1 Interfaz para capturar los datos generales del paciente.



Consulta Inicio

Generales Antecedentes Predisposiciones Causas Síntomas Confirmación Diagnóstico

Padres Asmáticos  Sí  No

Alérgico  Sí  No

Obesidad  Sí  No

Contacto con Polvo  Sí  No

Contacto con Polen  Sí  No

Atrás Siguiete

Fig. 4.3.2.4.2 Interfaz para capturar los antecedentes del paciente.

Consulta Inicio

Generales Antecedentes Predisposiciones Causas Síntomas Confirmación Diagnóstico

Tabaquismo  Sí  No

Infecciones Virales  Sí  No

Contacto con Animales  Sí  No

Ansiedad  Sí  No

Sedentarismo  Sí  No

Atrás Siguiete

Fig. 4.3.2.4.3 Interfaz para capturar las predisposiciones del paciente.

Consulta Inicio

Generales Antecedentes Predisposiciones Causas Síntomas Confirmación Diagnóstico

Daño Pulmonar  Sí  No

Hipertensión  Sí  No

Sibilancias  Sí  No

Disnea  Sí  No

Taquicardia  Sí  No

Atrás Siguiete

Fig. 4.3.2.4.4 Interfaz para capturar las causas de la condición de salud del paciente.

Fig. 4.3.2.4.5 Interfaz para capturar los síntomas del paciente.

A continuación, se presenta la sección de Confirmación en donde el usuario puede verificar los datos ingresados previo a realizar la consulta.

Fig. 4.3.2.4.6 Interfaz para presentar la confirmación de los datos ingresados por el paciente.

Una vez ingresados todos los datos en el sistema se determina si el paciente tiene asma o no y en caso positivo cuál es la severidad en la que lo presenta, por lo que el sistema informa del resultado del diagnóstico en la última sección, misma que se presenta a continuación:



Fig. 4.3.2.4.7 Interfaz para presentar los resultados del diagnóstico del paciente.

### 4.3.2.5 Pruebas de integración e implementación

Una vez terminado el desarrollo del prototipo, se validó que el sistema y la API funcionaran correctamente de acuerdo con la especificación de los requerimientos, mediante la ejecución de las pruebas unitarias y de integración.

Las pruebas unitarias se realizaron básicamente en 2 áreas: las primeras pruebas se realizaron con el objetivo de probar el backend y la implementación de la aplicación conforme a los requerimientos establecidos, por lo que se ingresaron primeramente datos de entrada a los endpoints de la API para validar si los datos eran procesados correctamente y si la respuesta de las llamadas estaba en el formato correcto.

El segundo bloque de pruebas realizadas fue de integración, donde se pretendía probar el acoplamiento entre el Frontend y el Backend, por lo que se ingresaron datos de prueba desde la interfaz gráfica, con la finalidad de validar el flujo de los datos dentro del proceso de la consulta, así como la respuesta a la consulta para mostrar correctamente la sección de resultados.

### 4.3.2.6 Validación y verificación.

Durante el proceso de validación y verificación del prototipo se le presentó la interfaz gráfica al experto humano para la retroalimentación del sistema, con el fin de saber si era lo que él esperaba de acuerdo con los requerimientos establecidos.

El médico comentó que la herramienta le parecía útil dado que la interfaz gráfica era muy intuitiva y fácil de usar, también señaló que el tiempo de respuesta era aceptable.

A opinión del experto humano, propuso una mejora que podría complementar al prototipo y sería tener la opción de poder mandar por correo electrónico el resultado de la consulta.

---

## 5 RESULTADOS Y DISCUSIÓN

---

En este capítulo se presentan los resultados del trabajo realizado, el cual se soporta en una metodología de diseño y desarrollo de una base de conocimiento propuesta para el diagnóstico médico del asma, la cual es consultada a través de un motor de búsqueda y una interfaz gráfica (web), donde el usuario es capaz de introducir datos generales de un paciente, sus antecedentes, predisposiciones, causas y síntomas a un proceso de evaluación que permita determinar si el paciente presenta asma y en qué nivel de severidad se encuentra.

También se aborda la discusión acerca de las pruebas realizadas en la solución tecnológica implementada, utilizando diferentes casos médicos, provistos por el experto médico, para hacer una comparación conjunta de los resultados obtenidos, así como también se expone el por qué el dominio médico es uno de los más complejos al momento de representar el conocimiento, por el grado de incertidumbre presente en el diagnóstico médico.

Además, se expone cuál podría ser el trabajo futuro, relacionado al uso de una base de conocimiento como la propuesta y qué mejoras podrían realizarse al prototipo.

## 5.1 Resultados

Los resultados están enfocados a la implementación del prototipo en un ambiente productivo, para recoger la experiencia vivida por el experto al momento de usar la herramienta, y al proceso de validación y calibración para mejorar la precisión del diagnóstico con diferentes casos de prueba.

### 5.1.1 Implementación del prototipo en un ambiente productivo

Con la finalidad de proveer al experto el entorno propicio para uso y prueba del prototipo desarrollado, se preparó y configuró el ambiente productivo, tareas que implicaron la publicación de la página web de la herramienta, así como la API que permite la interacción ente la página y el motor de búsqueda.

Para esto se usó un servidor con un sistema operativo Linux (Distribución de Ubuntu) al cual se le instalaron un servidor apache, el motor de MongoDB, el motor de Neo4j, lenguaje de programación Python versión 3.7 y las librerías requeridas para realizar la conexión a la base de datos desde Python. Además, se configuraron los puertos http tanto del sitio web como de las bases de datos dentro del firewall, para el acceso remoto.

### 5.1.2 Experiencia del experto con la plataforma

Una vez montado el ambiente de producción para realizar las pruebas, se le proporcionó al experto el acceso al prototipo para evaluar la herramienta y su usabilidad.

También se recolectó la opinión por parte del experto relacionado a la usabilidad de la herramienta, se le preguntó al usuario cuál era su impresión y experiencia con la interfaz de usuario al momento de realizar las consultas. En palabras del experto “La herramienta es sencilla de usar ya que es clara y te lleva de la mano para ingresar los datos del paciente y visualizar de forma rápida la probabilidad de la presencia del asma”.

### 5.1.3 Validación y calibración

Para hacer la validación del motor de búsqueda y verificar si este trabaja adecuadamente y es capaz de proporcionar la respuesta correcta al experto, se elaboraron una batería de casos de prueba en conjunto con el experto, quien fue el encargado de ingresarlos en la herramienta y comparar la respuesta obtenida contra el diagnóstico del experto por cada caso.

Dicha batería de pruebas contiene casos que permiten cubrir las diferentes respuestas que se esperarían de la herramienta, dependiendo de los datos ingresados por el experto en la valoración del paciente, cuyo proceso de evaluación puede llevar a tres resultados diferentes: paciente sin asma o paciente con asma y sus diferentes niveles de severidad o sin diagnóstico.

A continuación, se puede observar la batería de casos de prueba ejecutados desde el prototipo en conjunto con el experto humano para validar el diagnóstico.

Elementos de diagnóstico por categoría	No. de casos de prueba																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
<b>Antecedentes</b>																		
Padres Asmáticos	Sí	Sí	Sí	Sí		Sí	Sí			Sí	Sí	Sí				Sí	Sí	
Alérgico					Sí		Sí	Sí	Sí					Sí	Sí			
Obesidad	Sí	Sí	Sí	Sí										Sí				Sí
Contacto con Polvo			Sí		Sí			Sí	Sí	Sí		Sí			Sí	Sí	Sí	Sí
Contacto con Polen			Sí	Sí	Sí			Sí	Sí	Sí				Sí	Sí			
<b>Predisposiciones</b>																		
Tabaquismo				Sí		Sí											Sí	
Infecciones Virales			Sí		Sí			Sí	Sí	Sí				Sí				
Contacto con Animales	Sí	Sí			Sí	Sí		Sí	Sí	Sí	Sí				Sí	Sí	Sí	Sí
Ansiedad	Sí	Sí	Sí	Sí		Sí	Sí	Sí	Sí	Sí		Sí	Sí			Sí	Sí	Sí
Sedentarismo	Sí	Sí			Sí									Sí	Sí			
<b>Causas</b>																		
Daño pulmonar			Sí		Sí		Sí			Sí		Sí	Sí				Sí	Sí
Hipertensión								Sí	Sí				Sí	Sí				
Sibilancias	Sí	Sí	Sí	Sí			Sí	Sí	Sí	Sí	Sí					Sí	Sí	Sí
Disnea		Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí								Sí
Taquicardia	Sí	Sí		Sí										Sí				
<b>Síntomas</b>																		
Tos	Sí	Sí	Sí	Sí	Sí		Sí	Sí	Sí	Sí	Sí				Sí	Sí	Sí	Sí
Hiperventilación	Sí	Sí	Sí		Sí	Sí		Sí	Sí	Sí				Sí	Sí			Sí
Espasmo pulmonar	Sí	Sí	Sí						Sí	Sí			Sí	Sí		Sí	Sí	Sí
Erge			Sí					Sí										
Fiebre			Sí				Sí											

Tabla 5.1.3.1 Casos de prueba ejecutados para validar el funcionamiento del motor de búsqueda.

No.	Asma Esperado	Intensidad Esperada	Asma Obtenido	Intensidad Obtenida
1	Sí	Moderado	Sí	Grave
2	Sí	Grave	Sí	Grave
3	No		No	
4	Sí	Grave	Sí	Grave
5	Sí	Leve	Sí	Moderado
6	Sí	Moderado	Sí	Moderado
7	No		No	
8	No		No	
9	Sí	Moderado	Sí	Moderado
10	Sí	Grave	Sí	Grave
11	Sí	Leve	Sí	Moderado
12	No es posible		No es posible	
13	No es posible		No es posible	
14	No		No	
15	No es posible		No es posible	
16	Sí	Leve	Sí	Moderado
17	Sí	Moderado	Sí	Moderado
18	Sí	Grave	Sí	Grave

Tabla 5.1.3.2 Resultados de los casos de prueba ejecutados.

En la tabla anterior se pueden observar los resultados obtenidos una vez de haber ejecutado cada caso de prueba dentro del prototipo, aquellos resultados con fondo verde representan las respuestas correctas dadas por el buscador y que concuerdan con lo que el experto esperaba.

Mientras que los resultados señalados en color rojo son en los que existió una discrepancia entre la respuesta obtenida de herramienta y la respuesta esperada por el médico.

### 5.1.3.1 Validación de resultados exitosos y fallidos

Una vez ejecutada a batería de pruebas y como se puede observar en la tabla anterior, no en todos los casos se obtuvieron la respuesta esperada, ya que algunos de estos casos su severidad se estimaba de forma correcta, al estar muy cerca del límite entre 2 niveles.

A continuación, se mostrarán algunos casos exitosos, en donde una vez introducidos los datos en la herramienta esta fue capaz de determinar la severidad presente en cada caso de forma correcta:

Caso exitoso – Id No. 3

The screenshot shows a web-based diagnostic tool interface. At the top, there is a navigation bar with tabs: 'Generales', 'Antecedentes', 'Predisposiciones', 'Causas', 'Síntomas', 'Confirmación', and 'Diagnóstico'. The 'Diagnóstico' tab is currently selected. Below the navigation bar, the patient's general information is displayed: 'Datos Generales', 'Nombre : Alma Camacho Ramos', 'Sexo : F', 'Fecha de Nacimiento : 11-01-83', 'Edad : 38', and 'Tipo de Sangre : A+'. Below this, there are several categories of medical conditions listed with minus signs: 'Antecedentes' (Padres asmáticos, Obesidad, Contacto con polvo, Contacto con polen), 'Predisposiciones' (Infecciones virales, Ansiedad), 'Causas' (Daño pulmonar, Sibilancias, Disnea), and 'Síntomas' (Erge, Fiebre). At the bottom right of the form, there are two blue buttons: 'Atrás' and 'Diagnosticar'.

Fig. 5.1.3.1.1 Caso de prueba exitoso Id No. 3

Resultado

The screenshot shows the result of the diagnostic test. The navigation bar is the same as in the previous figure, but the 'Diagnóstico' tab is now highlighted. The result is displayed as follows: 'Asma', 'No', 'Probabilidad', and '0%'. The 'Diagnóstico' tab is the only one active in the navigation bar.

Fig. 5.1.3.1.2 Resultado caso de prueba Id No. 3



## Caso exitoso - Id No. 9

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico	
<b>Datos Generales</b>							
Nombre :	Martin Del Campo Covarrubias						
Sexo :	M						
Fecha de Nacimiento :	14-05-85						
Edad	36						
Tipo de Sangre :							
<b>Antecedentes</b>							
- Alergia							
- Contacto con polvo							
- Contacto con polen							
<b>Predisposiciones</b>							
- Infecciones virales							
- Contacto con animales							
- Ansiedad							
<b>Causas</b>							
- Hipertensión							
- Sibilancias							
- Disnea							
<b>Síntomas</b>							
- Tos							
- Hiperventilación							
- Espasmo Pulmonar							
						<a href="#">Atrás</a>	<a href="#">Diagnosticar</a>

Fig. 5.1.3.1.3 Caso de prueba Id No.9

## Resultado

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
<b>Asma</b>						
<b>Sí</b>						
Probabilidad						
<b>68%</b>						
Severidad						
<b>Moderada-Grave</b>						

Fig. 5.1.3.1.4 Resultado caso de prueba Id No. 9

## Caso exitoso – Id No. 12

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
Datos Generales						
Nombre :	Maria Rangel Porraz					
Sexo :	F					
Fecha de Nacimiento :	20-01-75					
Edad	47					
Tipo de Sangre :	O+					
Antecedentes						
- Padres asmáticos						
- Contacto con polvo						
Predisposiciones						
- Ansiedad						
Causas						
- Daño pulmonar						
Síntomas						
						<a href="#">Atrás</a>
						<a href="#">Diagnosticar</a>

Fig. 5.1.3.1.5 Caso de prueba Id No. 12

## Resultado

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
Asma						
<b>Si</b>						
Probabilidad						
-						
Severidad						
<b>No se puede diagnosticar</b>						

Fig. 5.1.3.1.6 Resultado caso de prueba Id No. 12

En seguida se mostrarán algunos casos fallidos, en donde la herramienta no determinó correctamente el nivel de severidad que el experto humano esperaría como resultado.

## Caso no exitoso – Id No. 1

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
<b>Datos Generales</b>						
Nombre :	Arturo Cuevas Torres					
Sexo :	M					
Fecha de Nacimiento :	28-08-92					
Edad :	29					
Tipo de Sangre :	B+					
<b>Antecedentes</b>						
- Padres asmáticos						
- Obesidad						
<b>Predisposiciones</b>						
- Contacto con animales						
- Ansiedad						
- Sedentarismo						
<b>Causas</b>						
- Sibilancias						
- Taquicardia						
<b>Síntomas</b>						
- Tos						
- Hiperventilación						
- Espasmo Pulmonar						
						<a href="#">Atrás</a>
						<a href="#">Diagnosticar</a>

Fig. 5.1.3.1.7 Caso de prueba no exitoso Id No. 1

## Resultado

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
<b>Asma</b>						
<b>Sí</b>						
Probabilidad						
<b>58%</b>						
Severidad						
<b>Grave</b>						

Fig. 5.1.3.1.8 Resultado caso de prueba no exitoso Id No. 1

## Caso no exitoso – Id No. 16

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
Datos Generales						
Nombre :	Mariana Fernandez Castrejon					
Sexo :	F					
Fecha de Nacimiento :	26-05-72					
Edad :	49					
Tipo de Sangre :	O-					
Antecedentes						
- Padres asmaticos						
- Contacto con polvo						
Predisposiciones						
- Contacto con animales						
- Ansiedad						
Causas						
- Sibilancias						
Síntomas						
- Tos						
- Espasmo Pulmonar						
						<a href="#">Atrás</a>
						<a href="#">Diagnosticar</a>

Fig. 5.1.3.1.9 Caso de prueba no exitoso Id No. 16

## Resultado

Generales	Antecedentes	Predisposiciones	Causas	Síntomas	Confirmación	Diagnóstico
Asma						
<b>Sí</b>						
Probabilidad						
<b>14%</b>						
Severidad						
<b>Moderada</b>						

Fig. 5.1.3.1.10 Resultado caso de prueba no exitoso Id No. 16

## 5.1.3.2 Calibración

Con base en los resultados de las pruebas realizadas, fue necesario realizar una calibración del prototipo, dado que la transición entre los niveles de severidad era imprecisa y no coincidían con el resultado esperado por el experto. Por lo que fue necesario agregar 2 niveles más de severidad dentro del diagnóstico, con la finalidad de incrementar la precisión en el diagnóstico que genera el motor de búsqueda. Quedando los niveles de severidad definidos de la siguiente manera: Leve, Leve-Moderada, Moderada, Moderada-Grave y Grave.

Una vez realizada la calibración del prototipo el experto humano volvió a ejecutar la misma batería de pruebas, para verificar que los resultados obtenidos fueran más cercanos al diagnóstico que el médico esperaría obtener.

## 5.2 Discusión

Basándonos en los resultados obtenidos se demuestra la complejidad que representa el modelado del conocimiento de un experto dentro del campo de la medicina, ya que, a pesar del uso de las reglas de producción y los mapas cognitivos difusos, la dificultad para diagnosticar de forma correcta se incrementa exponencialmente en relación con los elementos involucrados en el diagnóstico, puesto que las relaciones entre cada uno de estos también aumentan en un número significativo. Por lo que es posible establecer puntos de mejora al presente prototipo como lo puede ser, la implementación de un módulo de almacenamiento de expedientes, en donde se puedan cargar los estudios clínicos confirmatorios de cada paciente. También se puede mejorar la base de conocimiento agregando más elementos de diagnóstico para cada categoría establecida, ya que la arquitectura provee de flexibilidad para este tipo de tareas. El incremento de estos elementos permite incrementar la precisión del diagnóstico en el motor de búsqueda, al momento de determinar el nivel de severidad que presenta el paciente.

---

## 6 CONCLUSIONES

---

En este capítulo final se expondrán los aprendizajes que se tuvieron durante la realización de la base de conocimiento y el prototipo, las implicaciones que tiene la ingeniería del conocimiento para la adquisición y representación del conocimiento de un experto dentro de una base de conocimiento. También se hablará sobre los diferentes desafíos enfrentados durante las distintas etapas de desarrollo del presente trabajo, como lo fue el representar el conocimiento para el diagnóstico del asma, dado que el dominio médico, es uno de los más difíciles de modelar por la misma naturaleza de los elementos de diagnóstico, que involucran factores como la exploración del paciente, la realización de estudios clínicos confirmatorios y por supuesto la experiencia del médico.

Por último, se presentará el trabajo futuro a desarrollar dentro del dominio elegido, como lo puede ser el extender el uso del motor de búsqueda agregando más elementos de diagnóstico, para de esta manera incrementar la precisión del diagnóstico.

## 6.1 Conclusiones

El desarrollo de una base de conocimiento que permitiera el diagnóstico del asma presentó grandes retos al momento de elegir cual sería la mejor forma de representar un conocimiento del dominio médico. Desde la elección del experto y el uso de instrumentos para la recopilación de la información proporcionada por el experto, hasta el análisis e interpretación de los datos obtenidos.

El primer gran reto es el modelar adecuadamente el conocimiento del experto humano, ya que los procesos que involucran la toma de decisiones basadas en la experiencia médica son difíciles de modelar con cualquier técnica de representación del conocimiento. Otro gran reto al momento de representar el conocimiento es el manejo de la incertidumbre durante el proceso de diagnóstico, puesto que, para llegar a este, no se trata solo de cumplir con un conjunto de condiciones muy específicas, sino de considerar que también la experiencia del médico se encuentra involucrada. Es por ello que el uso de las reglas de producción y la aplicación del mapa cognitivo difuso, permitieron una gran flexibilidad para representar el conocimiento del médico de una forma más certera, entendiendo que, durante el proceso de diagnóstico, el experto no siempre sigue los mismos pasos o realiza las preguntas en el mismo orden, sino que va realizando preguntas al paciente, basándose en su experiencia.

El uso de mapas cognitivos difusos, permitieron hacer frente a este reto, ya que, al representar la relación entre los diferentes elementos del diagnóstico mediante el uso de la teoría de grafos, y el manejo de la incertidumbre haciendo uso de la teoría difusa para la asignación de pesos, fue posible realizar recorridos entre los nodos y aristas del grafo, recolectando así, todos los pesos especificados dentro de cada relación, con la finalidad de ayudar a calcular la probabilidad y severidad del asma de una forma más sencilla.

El principal aprendizaje que se tuvo fue que el uso de la teoría difusa para la representación del conocimiento es de mucha ayuda al momento de manejar la incertidumbre inherente al diagnóstico médico y como las tecnologías actuales como Neo4J permiten, almacenar de forma mucho más práctica, un mapa cognitivo difuso, hicieron que el desarrollo de un prototipo fuera más rápido, que si se hubiera intentado hacer con bases de datos relacionales.

## 6.2 Trabajo Futuro

Como parte del trabajo futuro que se puede realizar, podemos establecer la posibilidad de ampliar la base de conocimiento, ya que al agregar nuevos elementos de diagnóstico permite mejorar su asertividad.

Por otro lado, la expansión de la base de conocimiento para incluir diagnóstico infantil, que permita robustecer el prototipo completo y que sea capaz de realizar consultas a personas de cualquier rango de edad, habilitando el uso del prototipo a una población de médicos pediatras.

Por último, la implementación de un módulo dentro del prototipo que permita complementar el expediente del usuario con estudios de laboratorio para confirmar el diagnóstico. Esto en definitiva aumentaría

significativamente el alcance del prototipo, permitiendo de esta manera pasar de un prototipo como prueba de concepto al desarrollo de un sistema capaz de realizar el diagnóstico del asma con mayor precisión.



---

## 7 BIBLIOGRAFÍA

---

- [1] Aldana, Rosario y Vázquez Roberto, "Sistema Experto: Diagnóstico y Tratamiento Genérico de Asma Bronquial", 2019
- [2] Chandiok, Ashish y Chaturvedi, D.k., "Cognitive Decision Support System for medical diagnosis", 2005
- [3] Álvarez, Elena. "Lógica difusa con Geogebra", 2014
- [4] Pena Ayala, Alejandro. "Sistemas basados en Conocimiento: Una Base para su Concepción y Desarrollo". Instituto Politécnico Nacional, México, D.F., Centro de Investigación en Computación del IPN, 2006.
- [5] Cervantes, Jair. "Representación y aprendizaje de conocimiento con redes de Petri difusas. Tesis de maestría". Centro de investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México 2005
- [6] Gonzalez-Moreno, Diego. "Introducción a la teoría de las gráficas", 2017.
- [7] Vera, Christina y Migani Silvina, "Introducción a las bases de datos de grafos: Experiencias en Neo4j, Departamento de informática/ UNSJ
- [8] Zadeh, L. A., "Fuzzy sets information and Control", 1965.
- [9] Carlos, Marlene, "Sistema experto de diagnóstico médico del Síndrome de Guillian Barre".
- [10] Fillotrani, Pablo, "Un Análisis sobre la Lógica Difusa como Herramienta para a Representación del Conocimiento.
- [11] Cabrera, Mirna, "Aplicaciones médicas como ayuda al diagnóstico en la medicina. Experiencia SOFTEL-MINSAP", 2012
- [12] Esteve, Josep. "Introduccion al sistema operativo GNU/Linux". Universidad Oberta de Catalunya, 2010.
- [13] Blanco, Sergio. "Manual básico de GNU/Linux", 2005.
- [14] Mifsuf, Elvira, "Apache". Ministerio de Educación, Cultura y Deporte del Gobierno de España.
- [15] Condor, Enrique. "Algoritmos resueltos con Python". Eidec, 2020.
- [16] Sánchez, Christian. "Un enfoque evolutivo para el aprendizaje de conocimiento con redes de Petri difusas". CINVESTAV, 2013
- [17] Alomso, Margarita. "Conocimiento y Base de Datos: Una propuesta de integración inteligente". Departamento de organización de empresas/Universidad de Cantabria

---

## 8 GLOSARIO DE TÉRMINOS MÉDICOS

---

**Asma** - Afección en la que las vías respiratorias de una persona se inflaman, estrechan y producen mayores cantidades de mucosa de lo normal, lo que dificulta la respiración.

**ERGE** - Enfermedad crónica que ocurre cuando el ácido estomacal o la bilis ascienden por el esófago, lo que irrita su revestimiento. Cuando el reflujo ácido y la acidez estomacal suceden más de dos veces por semana, pueden ser un indicio de reflujo gastroesofágico.

**Sibilancias** - Sonido silbante y chillón al respirar.

**Disnea** - Dificultad para respirar.

**Alergia** - Afección en la que el sistema inmunológico reacciona de forma anormal frente a una sustancia extraña.

**Sedentarismo** - Aquellas actividades que realizamos las personas sentadas o reclinadas, mientras estamos despiertas, y que utilizan muy poca energía.

**Hipertensión** - Afección en la que la presión de la sangre hacia las paredes de la arteria es demasiado alta.

**Hiperventilación** - Es una respiración rápida o profunda.

---

## APÉNDICE A. ANEXOS

---

### Anexo 1.1 Cuestionario para la selección de dominio

- 1 ¿Qué tipo de padecimientos existen cuyo diagnóstico sea muy metódico?
- 2 ¿Cuál de estos padecimientos pueden ser diagnosticados de con mayor certeza sin necesidad de estudios de laboratorio?
- 3 ¿Qué preguntas se le realizan al paciente con la finalidad de poder diagnosticar el asma?
- 4 ¿Siempre se realizan las mismas preguntas?
- 5 ¿El orden de las respuestas a estas preguntas importa?
- 6 ¿Cuál es el porcentaje de certeza dan las respuestas a estas preguntas?
- 7 ¿Cuáles son las preguntas que tienen más peso en el diagnóstico del asma?
- 8 ¿Qué sucede si la fiebre está presente en el paciente al momento del diagnóstico?
- 9 ¿Qué tipo de elementos presentes en la valoración médica indican que el asma no está presente en el paciente?
- 10 ¿En una escala del 1 al 100 cuál es el impacto en la relación entre cada uno de los elementos?
- 11 ¿La asignación de valores cuantitativos reflejan el valor que se le da al momento de la consulta médica?

## Anexo 1.2 Cuestionario para el diagnóstico de asma en adultos

Datos Generales	
Nombre	
Apellido paterno	
Apellido materno	
Edad	
Sexo	

Datos Hereditarios	Sí	No
Padre o madre asmático		
Padre o madre alérgico		
Familiares con cáncer		

Datos Personales	Sí	No	
Diabetes			
Hipertensión			
Alergias			
Fuma			
Drogas			
Esta alfombrada la habitación			
Alérgico a algún medicamento			
Fumadores en casa			

Sintomatología	Sí	No									
Tos de foga			Frecuencia		Diaria		Semanal		Mensual		
Presenta congestión nasal			Frecuencia		Diaria		Semanal		Mensual		
Presenta silbidos de día o al dormir			Frecuencia		Diaria		Semanal		Mensual		
La tos se presenta con actividad física			Frecuencia		Diaria		Semanal		Mensual		
Gorgoteos al respirar			Frecuencia		Diaria		Semanal		Mensual		
Exacerbaciones y crisis frecuentes nocturnas			Frecuencia		Diaria		Semanal		Mensual		
Disnea continua			Frecuencia		Diaria		Semanal		Mensual		
Limitación de la actividad y del sueño											
Basal del 60%-80%											
Se encuentra en tratamiento			Aspirinas		Frecuencia		Diaria		Semanal		Mensual
			Desinflamatorios		Frecuencia		Diaria		Semanal		Mensual

Observación en la entrevista	Sí	No		
Puede hablar frases				
Se encuentra agitado				
Frecuencia respiratoria aumentada				
Músculos accesorio y retracción supraesternal				
Frecuencia cardiaca 100-120				
PEF Después del broncodilatador inicial entre 60% y 80%				
Disminución o aumento del trabajo respiratorio			Hablando	
			Prefiere sentarse	

### Anexo 1.3 diagnóstico.

### Cuestionario para establecer la relevancia en la relación entre los elementos de

	Padres asmáticos	Alergia	Obesidad	Contacto con polvo	Contacto con polen	Tabaquismo	Infecciones virales	Contacto con animales	Ansiedad	Sedentarismo	Daño pulmonar	Hipertensión	Sibilancias	Disnea	Taquicardia	Tos	Hiperventilación	Espasmo pulmonar	
Padres asmáticos																			
Alergia																			
Obesidad																			
Contacto con polvo																			
Contacto con polen																			
Tabaquismo																			
Infecciones virales																			
Contacto con animales																			
Ansiedad																			
Sedentarismo																			
Daño pulmonar																			
Hipertensión																			
Sibilancias																			
Disnea																			
Taquicardia																			
Tos																			
Hiperventilación																			
Espasmo pulmonar																			

---

## APÉNDICE B. OBTENCIÓN DE REQUERIMIENTOS MEDIANTE EL ESTANDAR IEEE 830

---

### 1. Funcionalidades del producto

El motor de búsqueda será capaz de consultar la base de conocimiento para calcular la probabilidad de que un paciente padezca de asma, tomando la información de una base de conocimientos para asma.

El prototipo de búsqueda podrá clasificar al paciente dentro de la severidad asmática a la que puede pertenecer de acuerdo con la información ingresada por el mismo.

El prototipo será capaz de almacenar la información referente a los pacientes dentro de una base de datos, con la finalidad de que pueda ser utilizada como fuente de reforzamiento de aprendizaje para versiones de mejoramiento del algoritmo en versiones futuras.

### 2. Características del usuario

El prototipo será capaz de dar acceso a un solo perfil de usuario, el cual solo podrá realizar consultas y decidir si la información ingresada del paciente debe ser almacenada o no previa autorización del paciente.

Este prototipo también tendrá la opción dentro del sistema de descargar el resultado de la consulta en un archivo PDF.

### 3. Restricciones

El prototipo solo presentara el resultado de la consulta a la base del conocimiento para el asma, presentando al usuario probabilidad de que el paciente tenga asma, basado en la información ingresada por el usuario.

El prototipo será capaz de clasificar la severidad asmática del paciente, sin embargo, no dará un diagnóstico definitivo.

Toda la información del paciente será encriptada y almacenada en la base de datos previa autorización de mismo.

### 4. Suposiciones y dependencias

El prototipo realizar consultas a dos bases de datos No-SQL para extraer el conocimiento del experto almacenado en dicha base de conocimiento, permitiendo al algoritmo del prototipo predecir la severidad del asma presentada por el paciente. Por lo que se deberá contar con una instancia instalada de MongoDB, así como una instancia de Neo4j.

Es necesario también contar con un servidor Linux, mismos que deberá tener instalado la versión 3.5 o superior de Python con la finalidad de hospedar el prototipo del sistema de consultas.

## 5. Actualización de requerimientos

Entre las posibilidades de mejora del prototipo podemos definir 2 grandes oportunidades, la primera sería la implementación de un base de hechos y un motor de inferencia con la finalidad de convertir el prototipo en un sistema experto, la segunda área de oportunidad sería mejorar la interfaz gráfica para que el usuario tenga una mejor experiencia al hacer uso del buscador.

## 6. Requerimientos específicos

### 6.1 Requerimientos de interfaz externa

#### 6.1.1 Interfaz de usuario

El usuario deberá contar con conocimiento básicos de navegación WEB, así como del uso de equipo de cómputo como el manejo de dispositivos de entrada y de salida, ya que esto es necesario para el uso del buscador.

#### 6.1.2 Interfaz de hardware

La interfaz de hardware con la que contará el prototipo estará principalmente relacionada con el acceso a una red local si el prototipo en un servidor dentro de una red local o de la conexión a internet si toda la infraestructura se encuentra instalada en servicios en la nube.

#### 6.1.3 Interfaz de software

La interfaz de software se realizará mediante las llamadas de la interfaz gráfica a los puntos de acceso al core del prototipo, así como también las transacciones que se realizará entre el core del sistema y las bases de datos, permitiendo la recuperación del conocimiento almacenado en las mismas.



Las únicas dos transacciones que se realizarán serán cuando el usuario decida realizar una consulta o cuando decida almacenar los datos ingresados para la consulta en la base de datos del prototipo

## 6.2 Requerimientos funcionales

A continuación, se enlistarán los requerimientos necesarios para el prototipo de consultas de manera de que se posible conocer los detalles de estos.

### 6.2.2 Ingresar información del paciente (RF1)

#### 6.2.2.1 Objetivo

Permitir el ingreso completo de la información referente al paciente tales como datos generales, antecedentes, exploración física y síntomas.

#### 6.2.2.2 Descripción

Este requerimiento se presentará una página web en donde se le pedirá al usuario ingresar toda la información referente al paciente dentro de cada una de las sub-secciones del formulario de consulta.

#### 6.2.2.3 Resultado esperado

Que el usuario pueda llenar cada una de las secciones del formulario de consulta mediante la interfaz de usuario presentada por el prototipo.

#### 6.2.2.4 Datos requeridos

Los datos requeridos por la interfaz de usuario son los siguientes:

- Generales:
  - Nombre: Se refiere al nombre del paciente.
  - Apellido Paterno: Se refiere al apellido materno del paciente.
  - Apellido Materno: Se refiere al apellido paterno del paciente.

- Sexo: Se refiere al género del paciente.
- Fecha de nacimiento: Se refiere a la fecha de nacimiento del paciente.
- Tipo de sangre: Se refiere a la familia sanguínea a la que pertenece el paciente.
- Hipertenso: Se refiere a que si el paciente es hipertenso o no.
- Antecedentes:
  - Padres asmáticos: Se refiere a si alguno de sus padres es asmático.
  - Fumador: Se refiere a si es fumador.
  - Alérgico: Se refieren a si el paciente es alérgico.
  - Obesidad: Se refieren a si el paciente tiene sobrepeso.
  - Contacto con polvo: Se refiere a si el paciente está en contacto con el polvo
  - Contacto con polvo: Se refiere a si el paciente está en contacto con el polen
- Predisposiciones:
  - Tabaquismo: Se refiere a si el paciente fuma frecuentemente.
  - Infecciones virales: Se refiere a si el paciente tiene alguna infección viral.
  - Contacto con animales: Se refiere a si el paciente convive con animales.
  - Ansiedad: Se refiere a si el paciente tiene ansiedad.
  - Sedentarismo: Se refiere a si el paciente tiene baja o nula actividad física.
- Causas:
  - Daño pulmonar: Se refiere a si el paciente presenta daño pulmonar.
  - Hipertensión: Se refiere a si el paciente es hipertenso.
  - Sibilancia: Se refiere a si el paciente presenta silbidos al respirar.
  - Disnea: Se refiere a si el paciente padece de disnea
  - Taquicardia: Se refiere a si el paciente presenta un ritmo de palpitaciones acelerado.
- Síntomas:
  - Tos: Se refiere a si el paciente presenta tos.
  - Hiperventilación: Se refiere a si el paciente respira de forma rápida y agitada.
  - Espasmo pulmonar: Se refiere a si el paciente presenta espasmos en los pulmones.
  - Erge: Se refiere a si el paciente tiene o tuvo ERGE
  - Fiebre: Se refiere a si el paciente a si el paciente supere los 37 grados de temperatura.

## 6.2.3 Mostrar la clasificación de severidad asmática del paciente (RF2)

### 6.2.3.1 Objetivo

El prototipo deberá de clasificar correctamente la severidad que presenta el paciente de acuerdo la información ingresada en el formulario de consulta. Y Mostrar la clasificación obtenida de a raíz de consultar la base de conocimiento del prototipo.

### 6.2.3.2 Descripción

Este requerimiento deberá de presentarse cuando el prototipo ejecute la consulta a la base de conocimientos y determine la severidad del caso presentado con base en la información ingresada del paciente mediante el formulario de consulta.

### 6.2.3.3 Resultado esperado

Que el prototipo sea capaz de mostrar la probabilidad de que padezca asma de acuerdo con la base de conocimiento sobre la que se realizó la consulta.

### 6.2.3.4 Datos requeridos

Los datos necesarios para este requerimiento es la base de conocimiento, compuesta por las reglas de producción y los mapas cognitivos difusos, los cuales están almacenados en Mongo y en Neo4j respectivamente.