

PI14104

**MODELO INTELIGENTE DE RECOMENDACIÓN DE CAMPAÑAS, BASADO EN
PERFILAMIENTO DE HÁBITOS DE CONSUMO**

WILLIAM HERNÁN SUÁREZ MOLANO

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN
BOGOTÁ, D.C.
2015**

PI141-04

**MODELO INTELIGENTE DE RECOMENDACIÓN DE CAMPAÑAS,
BASADO EN PERFILAMIENTO DE HÁBITOS DE CONSUMO**

Autor:

William Hernán Suárez Molano

MEMORIA DEL TRABAJO DE GRADO REALIZADO PARA CUMPLIR UNO
DE LOS REQUISITOS PARA OPTAR AL TÍTULO DE
MAGÍSTER EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

Director

LUIS IGNACIO SUÁREZ SANTAMARIA

Comité de Evaluación del Trabajo de Grado

Página web del Trabajo de Grado

<http://pegasus.javeriana.edu.co/~PI141-04-RecomProdSecFinanz>

PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN INGENIERIA DE SISTEMAS Y COMPUTACIÓN
BOGOTÁ, D.C.
Noviembre, 2015

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
CARRERA DE INGENIERIA DE SISTEMAS**

Rector Magnífico

Jorge Humberto Peláez Piedrahita, S.J.

Decano Académico Facultad de Ingeniería

Ingeniero Jorge Luis Sánchez Téllez

Director de la Carrera de Ingeniería de Sistemas

Ingeniero Germán Alberto Chavarro Flórez

Director Maestría en Ingeniería de Sistemas y Computación

Ingeniera Angela Cristina Carrillo Ramos. Phd

Director Departamento de Ingeniería de Sistemas

Ingeniero Rafael Andrés González Rivera Phd.

Artículo 23 de la Resolución No. 1 de Junio de 1946

“La Universidad no se hace responsable de los conceptos emitidos por sus alumnos en sus proyectos de grado. Sólo velará porque no se publique nada contrario al dogma y la moral católica y porque no contengan ataques o polémicas puramente personales. Antes bien, que se vean en ellos el anhelo de buscar la verdad y la Justicia”

Contenido

I INTRODUCCIÓN	1
II – OBJETIVOS	4
1. OBJETIVO GENERAL.....	4
2. OBJETIVOS ESPECÍFICOS	4
3. METODOLOGÍA	4
III – ESTADO DEL ARTE.....	5
1. CENTRALIZACIÓN DE CLIENTES.....	5
2. SEGMENTACIÓN DE CLIENTES.	6
3. PERFILAMIENTO DE CLIENTES	6
3.1 CARACTERIZACIÓN DE PERFILES	7
3.2 DESCRIPCIÓN DE SEGMENTOS	8
3.3 ADQUISICIÓN DE DATOS DE CLIENTES	9
3.4 PERFILES Y PERSONAS	10
4. ANÁLISIS DE DATOS	10
5. MODELOS DE PERFILAMIENTO	11
6. SOCIAL MEDIA	12
7. MOTORES DE REGLAS.....	13
8. ONTOLOGÍAS EN REDES SOCIALES.....	15
9. TRABAJOS RELACIONADOS.....	17
10. CONCLUSIONES	19
IV - CONTRIBUCIONES	20
1. MODELO INTELIGENTE DE RECOMENDACIÓN DE CAMPAÑAS, BASADO EN PERFILAMIENTO DE HÁBITOS DE CONSUMO.	20
2. COMPONENTE DE SEGMENTACIÓN.....	25
2.1 SELECCIÓN DE VARIABLES	25
2.2 DEFINICIÓN DE SEGMENTOS	28

3.	CARGA DE DATOS DEMOGRÁFICOS Y PROCESAMIENTO DE EVENTOS TRANSACCIONALES.	32
4.	COMPONENTE DE PROCESAMIENTO DE EVENTOS.....	33
4.1	COMPONENTES, PROCESOS Y TÉCNICAS DEL PROCESADOR DE EVENTOS DE UBICACIÓN GEOGRÁFICA.....	33
4.2	PROCESADOR DE EVENTOS SOCIALES	34
4.2.1	<i>Procesamiento de comentarios</i>	34
4.2.2	<i>Red de ontologías</i>	39
5.	COMPONENTE DE INFERENCIA.....	44
5.1.1	<i>Motor de reglas:</i>	44
V - DISEÑO	45
1.	VISTA CASO DE USO ARQUITECTÓNICO	45
2.	DIAGRAMA DE ACTIVIDADES	53
3.	DIAGRAMA DE SECUENCIA FACEBOOK.....	56
4.	DIAGRAMA DE DESPLIEGUE.....	57
5.	DIAGRAMA DE COMPONENTES	58
VI – CONCLUSIONES Y TRABAJO FUTURO	59

ABSTRACT

Segmentation and sent to mass system campaigns, aren't based on a thorough analysis of individual preferences and interests expressed and inferred. Which can be performed throw data analyzing on different sources. Allowing to improve don't only knowledge about customer, also to turn react to new likes and improve effectiveness level on recommendation. Hence we arise creation of a profiling and consumption habits intelligent system based, which react providing personalized offers, against individual preferences and needs, through events processing on physical and real world, like events on networks, geographical location and financial transaction.

RESUMEN

Sistemas de segmentación y envío de campañas masivas, no se basan en un análisis exhaustivo a nivel individual de preferencias e intereses expresados e inferidos, que son establecidos mediante el análisis de varias fuentes de datos. Permitiendo mejorar no solo el conocimiento acerca del cliente, si no a su vez reaccionar a nuevos intereses, mejorando el nivel de efectividad en la recomendación. Para esto, se plantea la creación de un modelo inteligente de recomendación de campañas, basado en perfilamiento de hábitos de consumo. De manera que, reaccione proveyendo oferta adaptada a preferencias y necesidades del individuo, por medio del análisis y procesamiento de eventos del mundo físico y virtual como: Eventos en redes sociales, eventos de ubicación geográfica y eventos provenientes de transacciones financieras.

RESUMEN EJECUTIVO

Los procesos de envío de campañas masivas junto con los sistemas de recomendación, no han tenido los resultados esperados, a pesar de ser una prioridad estratégica en la época actual. Lo anterior, debido a varios factores, entre los cuales se encuentran: los métodos asociados a la definición del perfil del cliente, en los cuales se describe los intereses de los mismos a nivel individual, sin reflejar nuevas preferencias [1]; sistemas de envío de campañas masivas con bajo grado de individualización; falta de alimentación con otras fuentes de datos para mejorar los procesos de perfilamiento y de recomendación; falta de reacción a eventos del mundo físico y virtual, que pueden influir en la decisión de compra de los individuos. Todo lo descrito sustenta la importancia de utilizar otras fuentes de información para el proceso de perfilamiento, como las encontradas en la social media, implementando un enfoque de segmentación psicográfica, que incorpore características relacionadas con la actitud y percepción junto con las variables demográficas. De igual manera es relevante, la construcción de modelos de perfilamiento híbridos, que contemplen información individual y grupal [2].

El perfilamiento de clientes incluye todas las técnicas, algoritmos y procesos involucrados para la captura de datos acerca de clientes y su análisis, con el fin de crear información acerca del comportamiento. Esto incluye gustos y otras características, que son consolidadas como parte de la información mencionada, para ser presentada a los tomadores de decisiones, que podrán con base en la misma, impulsar acciones beneficiosas hacia el negocio [2]. Perfilamiento, que a su vez permite mejorar las estrategias enfocadas a centralización de clientes, estrategia que busca focalizar todas las operaciones de la compañía alrededor de estos, como un método para aumentar lealtad y reducir sus tasas de pérdida.

Tratando de mejorar la eficacia y calidad de la recomendación de campañas masivas, llevándolas a un mayor nivel de individualización, se creó un modelo inteligente de recomendación de campañas, basado el perfilamiento de hábitos de consumo, mediante cuatro fases, enmarcadas en la metodología la Ciencia Basada en el Diseño: la primera dedicada al análisis de las diferentes características demográficas, psicográficas y comportamentales de mayor relevancia para la definición de segmentos basados en consumo. Una vez identificadas estas variables, se tomaron un conjunto de segmentos previamente definidos por organizaciones reconocidas en el campo de investigación de mercado en el contexto colombiano, como el caso de la empresa Yanhaas y la cámara de comercio de Bogotá; instituciones que identificaron segmentos como: Super Mujeres, Nuevos ricos, Open minded, Abuelos activos, Independientes, Intelectuales, Rebeldes, Recién casados, Sabelotodos, Saludables, Sensibles, entre otros. Los anteriores segmentos fueron especificados en términos de variables demográficas, psicográficas y geográficas, algunas obtenidas y otras inferidas, con el propósito de ser utilizados como criterios de segmentación. En la segunda fase, se realizó un ciclo de rigor donde se seleccionaron las variables y segmentos a utilizar, que describieran y expresaran en términos demográficos y psicográficos las preferencias e intereses de los usuarios. Además de lo anterior, se identificaron elementos de entrada para el ciclo de diseño como: información del sector financiero transaccional y de social media para categorizar clientes dentro de segmentos previos; extracción de variables demográficas como la edad, género, estrato y estado civil; así mismo, psicográficos provenientes de redes sociales, y geográficos obtenidos mediante la identificación de estable-

cimientos de comercio cercanos a la ubicación del cliente. También, se detectaron errores comunes en la planeación de campañas de marketing efectivas, en las cuales no se tiene en cuenta las barreras socioculturales enfrentadas, por lo cual este proyecto toma de referencia información proveniente del mercado del consumidor colombiano[3]. Por otra parte, se identificó una tendencia creciente en la era actual, en que las compañías interactúan con sus clientes por diferentes canales, de forma directa e indirecta, en busca de mejorar su relación con estos, ofreciéndoles mayores niveles de personalización (*Multichannel Communication marketing*). Herramientas que permiten entender no solo quién es su cliente, si no a la vez, identificar su demanda, en el momento de la oferta [4]. Lo anterior se materializa a modo de recomendación de promociones en tiempo cercano al real, detectando y procesando eventos que evidencien una oportunidad o intencionalidad de compra. Lo mencionado permite generar aprendizaje acerca de los intereses manifestados, a ser implementados en campañas futuras.

Posteriormente se identificaron una serie de herramientas *Multichannel Communication Marketing*, las cuales aplicadas correctamente, ayudan en la consecución de objetivos propuestos en el proyecto. Sin embargo, dichas herramientas son de carácter *software* propietario, y no ofrecían una herramienta tangible que sirviera para implementar y demostrar la funcionalidad del modelo propuesto. Por tanto, se decidió realizar un nuevo ciclo de rigor dando como resultado la decisión de implementar un motor de reglas con capacidad de encadenamiento hacia delante (*Forward chaining*). Una vez cumplido el ciclo anterior, se dio paso al diseño de la solución y la construcción del artefacto. En este nuevo ciclo, se definió la arquitectura del sistema y sus principales componentes. Arquitectura que parte de un motor de reglas, un componente de segmentación y de procesamiento de eventos; para correlacionar una o más campañas con un conjunto de reglas descritas en términos de hechos y originados a partir del procesamiento de los siguientes tipos de eventos: ubicación geográfica (almacena los establecimientos cercanos al cliente a partir de su ubicación), eventos sociales (almacena los nuevos interés y las categorías de consumo inferidas a partir de los comentarios realizados por el usuario) y un procesador de eventos transaccionales (captura e identifica la categoría de comercio proveniente de una transacción financiera).

Finalmente para la validación del sistema, se realizaron pruebas funcionales y de concepto, partiendo de un conjunto de hipótesis con datos ficticios, se comprobó su funcionalidad en cada uno de los componentes: componente de segmentación, componente de procesamiento de eventos sociales, componente de procesamiento de eventos de geolocalización, componente de procesamiento de eventos transaccionales y un componente central compuesto por un motor de reglas, el cuál notifica al usuario vía correo electrónico la campaña(s) que se adhiere a los intereses manifestados, mediante la evaluación de un conjunto de reglas previamente establecidas. Entre las pruebas funcionales desarrolladas se encuentra un usuario ficticio con la siguiente información: persona de 26 años de edad, interesado por deportes y gimnasios. Usuario el cual una vez autenticado, fue clasificado en el segmento de “Solos”, posteriormente se realizó la simulación de un evento social, mediante el comentario en el muro de Facebook: “deseo ingresar al gimnasio”, clasificando la intencionalidad de consumo manifestada como un hecho “Saludable”. Además de lo anterior, se simuló un evento de ubicación geográfica, que detectó un establecimiento de comercio de la categoría “Gimnasios” a menos de 500Mts de la ubicación del cliente. Por último, el motor de reglas, ejecuto la siguiente regla, a partir de la correlación de esta con los hechos almacenados por los eventos procesados: personas que

pertenecen al segmento “Solos” y realizan un comentario social “Saludable”, y además se encuentran cerca de un almacén con categoría de comercio “Gimnasios”, aplican para promoción de 40% de suscripción por un año a gimnasios. La anterior regla efectúa envío de dicha promoción al correo electrónico del usuario.

I INTRODUCCIÓN

Entre los problemas que afectan al sector financiero, están los relacionados con la pérdida de lealtad empresarial, la escasa permanencia y las altas tasas de fuga de clientes [1]. Lo anterior se evidencia en diversos estudios que demuestran las bajas tasas de adquisición de nuevos productos por parte de clientes financieros con su banco primario (entidad donde un cliente posee su cuenta principal), los cuales prefieren la adquisición de nuevos productos con terceros [5]. Adicional a lo descrito anteriormente y a causa del incremento en la propagación y penetración de tecnologías como la banda ancha y el internet, ha surgido el nacimiento de nuevos servicios, que han provocado una demanda de usuarios cada vez más exigentes, aumentando la dificultad en las campañas de marketing [6] [7]. Por lo anteriormente descrito, han brotado nuevas herramientas, prácticas y métodos analíticos que buscan contrarrestar este tipo de problemáticas, además de una alta fidelización, aumentar las ventas e incrementar la satisfacción del cliente, entre estas herramientas se encuentran: los Multichannel communication marketing [8], el *cross selling*[9] y el perfilamiento de usuarios ;[10] este último, referido al proceso de capturar datos acerca de clientes y su análisis para obtener y crear información referente a necesidades, comportamientos, gustos, y otras características. Herramientas y técnicas que ofrezcan características de envió de ofertas adaptadas a las necesidades propias del cliente.

Entre las principales características de perfilamiento, se encuentran las que dependen de tres (3) conjuntos de variables, a saber:

Demográficas: atributos que son típicamente usados para describir individuos dentro de una población, como: edad, género, etnia, región de origen, raza, estado marital, nivel de educación, nivel económico, tipo de vivienda, entre otros.

Comportamental: atributos de valoración de datos provenientes de transacciones pasadas, que reflejan patrones en común; dentro de dichos patrones se pueden considerar: patrones de uso de productos, frecuencia de compra, volumen de compra, entre otros.

Psicográfica: variables que describen aspectos cualitativos adquiridos, atributos asociados a la personalidad, opiniones, creencias, intereses, estilo de vida, entre otros; generalmente usados para categorizar datos de compradores acorde con sus motivadores de compra

En estos procesos de caracterización, se pueden llegar a presentar atributos correlacionados; un ejemplo de correlación de atributos psicográficos con atributos demográficos, de cara a una segmentación que expresa propensión a consumo, puede ser el siguiente: clientes hombres entre los 40 y 60 años de edad con altos niveles de ingresos pueden llegar a conducir vehículos de lujo. [8] [9] [10].

El perfilamiento de usuarios también es usado en las organizaciones para validar que los clientes y productos estén acorde con la audiencia objetivo, al igual de ser implementados en estrategias de social media, centralización de clientes y en estrategias de Multichannel communication marketing. Para comprender mejor el último término, este se define como aquel que otorga

a las empresas la capacidad de definir, deducir, orquestar y comunicar ofertas a segmentos de clientes, por medio de un entorno multicanal, como emails, sitios web, call center, sms, entre otros; de igual modo selecciona ofertas que le son significativas al cliente, de acuerdo a una micro segmentación, y le permite generar aprendizaje para campañas futuras [11] [12]. Sin embargo, en el caso particular del sector financiero, los procesos de envío de campañas masivas junto con los sistemas de recomendación, no tienen los resultados esperados a pesar de ser una prioridad estratégica en la época actual. Lo anterior, debido a varios factores entre los cuales se identifican: los métodos asociados a la definición del perfil del cliente, en los cuales se describe sus intereses a nivel individual, sin reflejar nuevas preferencias [13]; sistemas de envío de campañas masivas con bajo grado de individualización; falta de alimentación con otras fuentes de datos para mejorar los procesos de perfilamiento y no reacción a eventos del mundo físico y virtual, que pueden influir en la decisión de compra de los individuos. Todo lo descrito sustenta la importancia de utilizar otras fuentes de información para el proceso de perfilamiento, como las encontradas en la social media, implementando un enfoque de segmentación psicográfica, que incorpore características relacionadas con la actitud y percepción junto con las variables demográficas. De igual manera es relevante, la construcción de modelos de perfilamiento híbridos, que contemplen información individual y grupal [12] [14] [15].

Con la llegada de la social media¹, se tiene el poder de acceder a grandes volúmenes de información, cada vez más variada, como la que se encuentra en redes sociales como Facebook y que relaciona gustos, deseos personales y campos de interés de los clientes como: likes, conexiones, eventos, grupos, insights, links, page_fan, estados, actividades, libros, localización, educación, ciudad, intereses, películas, música, deportes, televisión, trabajo, entre otros. Dicha información sirve como fuente de datos para el desarrollo de estrategias e implementación de técnicas y procesos, para la construcción de perfiles a partir de un examen exhaustivo de los intereses de los clientes. Lo expuesto hasta el momento, con el fin de impulsar los procesos de recomendación y generación de campañas, adaptadas a las características, necesidades e intereses de cada tipo de individuo [16]. Información social, que es aún más poderosa al ser combinada con información transaccional y otros tipos de eventos relacionados con intereses, para obtener una vista comprensible entre las necesidades del cliente y la oferta de la compañía [17].

Además del uso dado a la información proveniente de la social media para la creación de perfiles y segmentos de clientes por medio de su análisis, ésta misma puede igualmente contribuir a la mejora y fortalecimiento de las relaciones con estos, por medio del énfasis en tres tipos de información: compartida, conversaciones, relaciones y grupos [18]. Información que igualmente puede ser utilizada para la promoción de nuevos productos y para el incremento de las ventas en las organizaciones, ofreciendo una alternativa para enganchar nuevos clientes, a partir de la personalización en la atención al mismo. Como resultado, surge la necesidad de las organizaciones en incorporar este tipo de información [19] [20].

Este documento presenta un modelo inteligente de recomendación de campañas, el cual incorpora una vista estática del cliente (proveniente de un proceso de segmentación basado en datos

¹ [1] Sitio web y aplicaciones que permiten a los usuarios crear y compartir contenido o participar en redes sociales.

demográficos y psicográficos) y una vista dinámica del mismo (correspondiente al monitoreo de la actividad realizada por el usuario en redes sociales a nivel de likes y comentarios en la red social Facebook), para la generación de recomendaciones individualizadas en el marco de campañas masivas. Las campañas concebidas dentro del modelo propuesto, tienen en cuenta nuevos intereses y necesidades obtenidas e inferidas, mediante el procesamiento de eventos sociales, de geolocalización y de transacciones financieras. Lo anterior con base en la gestión de eventos virtuales y reales, y en la instanciación de reglas en un motor de inferencia.

La estructura de este documento se compone de: primera y segunda parte, contextualización acerca de la problemática que abarca el proyecto, objetivos y metodología. Tercera parte, análisis del estado del arte y trabajos relacionados. Cuarta parte, descripción de las principales contribuciones del proyecto y cómo se aborda la problemática planteada al inicio del proyecto. Quinta parte, descripción de los componentes e integraciones que hacen parte de la solución resultante. Finalmente, se presentan conclusiones y trabajo futuro.

II – OBJETIVOS

En la presente sección se describen los objetivos del proyecto.

1. Objetivo General

Crear un sistema inteligente para deducción y recomendación de campañas, basado en perfilamiento de hábitos de consumo.

2. Objetivos Específicos

- Identificar investigaciones de mercado que caracterizan segmentos basados en atributos demográficos, psicográficos y comportamentales, que determinan propensión a consumo.
- Establecer variables y segmentar usuarios basado en las investigaciones de mercado previamente establecidas, identificando cuáles son asequibles sobre las fuentes de información.
- Generar un modelo inteligente de deducción y recomendación de campañas, mediante la descripción de reglas que representen los segmentos y las propensiones a consumo establecidas.
- Construir y validar el modelo inteligente de recomendación de campañas mediante el diseño y construcción de una herramienta de software, y su implementación en un caso de estudio en el sector financiero, para la recomendación inteligente de campañas.

3. Metodología

Respecto a la metodología usada para el desarrollo del proyecto, en el Anexo I, se evidencian sus diferentes ciclos inmersos al interior del proceso, al igual que las diferentes etapas que lo conforman; el cual parte de la solución a una problemática del mundo real en la que se evidencian problemas asociados a la recomendación masiva de campañas, las cuales no realizan un análisis exhaustivo de las preferencias y necesidades de los clientes, que pueden ofrecer significancia a las campañas para el consumidor, en cuanto interpretan sus propios gustos y preferencias. Por otra parte, bajo esta metodología se crea contribución al conocimiento mediante la generación de un artefacto (construcción de un modelo inteligente de recomendación de campañas y la generación de un sistema de información que implementa dicho modelo).

III – ESTADO DEL ARTE

En la presente sección, se realiza un análisis del estado del arte, mediante la descripción de los principales conceptos asociados al contexto del proyecto, entre los que encontramos: centralización de clientes, perfilamiento y segmentación, descripción de métodos, técnicas y procesos implementados para tal fin, uso del social media para la extracción de conocimiento, problemáticas en las que se implementa motores de reglas y el uso de ontologías para la representación de conocimiento. Por último, se presenta un análisis comparativo de trabajos relacionados al presente proyecto.

1. Centralización de clientes

La centralización de clientes provee una introducción al concepto de enfocar la atención en la experiencia e interacción de los clientes.

A lo largo de los últimos años, el volumen de la interacción con los clientes ha aumentado de forma considerablemente acelerada.; como resultado se ha producido una despersonalización en la comunicación con el cliente, originando el nacimiento de iniciativas al interior de las organizaciones, respecto al desarrollo de nuevas estrategias de negocio para mejorar la satisfacción del cliente, concepto al que se le conoce como “centralización de clientes”. La centralización de clientes incorpora ideas, enfoques, estrategias, y tácticas que han evolucionado con el tiempo en alineación con iniciativas de gestión orientadas al cliente o con foco en este [8]. Todas estas iniciativas comparten un punto clave en común: enfocar las operaciones del negocio alrededor de los clientes de la compañía, como una forma de incrementar la lealtad de estos, reducir sus tasas de pérdida e incrementar las ventas; lo que resulta en la entrega de productos y servicios de mejor calidad. A modo de contextualización, se retorna a la era del comercio electrónico, donde se suponía que la atención al cliente era algo relativamente fácil de proveer, dado que los negocios eran menos complejos, y comparativamente pequeños si se confronta con los estándares actuales, por lo cual, establecer y mantener relaciones con los clientes era algo sencillo. Por lo anterior, en entornos pequeños los propietarios o trabajadores de productos al por menor o servicios de negocios, mantenían una relación directa con los clientes, al ser conocedores de varios de sus factores y características, como: el nombre de sus hijos, la fecha de sus cumpleaños e incluso la de su aniversario; como resultado este conocimiento ponía al cliente en una posición de carácter personal, ayudando a la compañía en la construcción de relaciones duraderas y proveyéndoles atención personalizada [21].

Para tratar de mantener la calidad de las relaciones, a medida que el negocio crecía, se veía la necesidad de aumentar la cantidad de personal para satisfacer las necesidades de una tendencia incremental de clientes. Por otra parte, la ampliación de horas de trabajo entre empleados ocasionaba la reducción de la interacción personal y afectaba el relacionamiento con los clientes, por ello a medida que los negocios crecían, mantener la calidad de sus relaciones con estos, era cada vez más complejo, de manera que las compañías empezaron a centrar sus esfuerzos en el uso de slogans de marketing orientados al cliente como: “El cliente es el número 1”, al igual que la implementación de varias estrategias enfocadas a este. A pesar de lo descrito anteriormente, entre las compañías más trataban de focalizarse en el cliente, su habilidad para proveer

ese “toque personal” como punto diferenciador, seguía siendo más difícil de mantener. Actualmente diferentes herramientas tecnológicas tratan de solventar este tipo de problemáticas, como el caso de los sistemas CRM’S [21].

Customer Relation management (CRM): Los sistemas CRM han nacido para ofrecer a las organizaciones la capacidad de mantener y mejorar las relaciones con sus clientes, al igual que el proveer un conjunto de estrategias para ayudar a las organizaciones con un servicio al cliente escalable, para una comunidad de clientes cada vez mayor [22]. CRM también es considerado como un enfoque de negocio relevante, el cual influencia el comportamiento del cliente por medio de una comunicación significativa para conseguir mejorar su adquisición, retención, fidelización e incrementar así la rentabilidad [23]. Sin embargo, las compañías han empezado a notar que con el uso de estos sistemas se ha perdido ese “toque personal” diferenciador querido por su clientela, al proveer una vista estática de la información, al momento de iniciar la relación con este último. Lo anteriormente expuesto, ocasiona una mercantilización de la atención, que afecta la relación con los clientes, los cuales cada día demandan por productos y servicios acorde con sus necesidades y preferencias. Por lo anterior, han surgido diferentes técnicas que buscan subsanar este tipo de problemáticas entre las que se pueden destacar: el perfilamiento de clientes, el análisis predictivo, entre otros, explicados a lo largo de este documento.

2. Segmentación de clientes.

Uno de sus principales objetivos, es lograr una comunicación objetiva con los clientes, en el cual su proceso se realiza mediante la descripción de las características de los individuos pertenecientes a ciertos grupos, dentro de los datos recolectados; es decir, agrupar a los clientes dependiendo de características e intereses similares a los demás miembros del grupo para poderlos accionar [23].

Dentro de los tipos de segmentación también se encuentra la segmentación de mercado, que involucra ver un mercado heterogéneo como un conjunto reducido de mercados homogéneos, que responden a cierto conjunto de preferencias y atributos relacionados con sus deseos, para de esta forma satisfacer mejor sus necesidades. Dicho lo anterior, deben existir criterios de homogeneidad, distinción y reacción, para clasificar a las personas en grupos homogéneos [24]. Sin embargo, el proceso de segmentación no es una tarea sencilla, ya que presenta varias dificultades como el no contar con suficiente información significativa y de calidad con actualización continua [23].

3. Perfilamiento de clientes

El perfilamiento de clientes incluye todas las técnicas, algoritmos y procesos involucrados para la captura de datos acerca de clientes y su análisis, con el fin de crear información acerca del comportamiento. Esto incluye gustos y otras características, que son consolidadas como parte de la información mencionada, para ser presentada a los tomadores de decisiones, que podrán con base en la misma, impulsar acciones beneficiosas hacia el negocio [2]. Fuentes de infor-

mación como las provenientes de la social media, también han sido utilizadas en algunos estudios, con el propósito de identificar hábitos e intereses [25]. De igual manera, dicha información social, ha sido combinada con las características de los usuarios, con la finalidad de lograr efectuar segmentaciones. Las características o variables incorporadas en cada segmento, permiten generar hipótesis acerca del comportamiento esperado de todo el grupo que pertenece a dicho segmento. Un ejemplo sencillo es la segmentación de hombres y mujeres, que permite a los analistas identificar diferentes patrones de compra, basados en el género. De ahí que si existen diferencias significativas a este respecto, se pueden asegurar cambios en la interacción con los clientes dependiendo de su género. Es por esta misma consideración que cuando un cliente interactúa en los servicios de atención telefónica, recibe una atención más rápida dependiendo de su valor para la compañía, con lo cual es asignada una mayor prelación en dicha atención cuando el usuario es clasificado como VIP frente a uno regular.

En términos generales, una segmentación de clientes es un ejercicio de modelado que incorpora los métodos para [2]:

- ✓ Decidir el conjunto de atributos más relevante para caracterizar a los clientes.
- ✓ Analizar los datos y aislarlos dependiente de dichos atributos.
- ✓ Organizar la población de clientes.
- ✓ Clasificación de segmentos.
- ✓ Uso de modelos para segmentación.

Gran parte de lo descrito hasta el momento depende del proceso de construcción del perfil del cliente, lo cual incluye las actividades de análisis, segmentación y clasificación.

3.1 Caracterización de perfiles

El perfilamiento de clientes busca un modelo de comunicación con clientes existentes, que permita ofrecerles mejores productos y servicios, para lo cual se requiere el uso de diversas fuentes de información, con el fin de poder efectuar las actividades de segmentación. Entre la información necesaria para la realización de este tipo de procesos encontramos [23] [24]:

Demográfica: aquellos atributos que son típicamente usados para describir a los individuos, dentro de una población. Estos atributos pueden ser combinados para caracterizar grupos. Algunos ejemplos de este tipo de variables son: edad, género, etnia, región de origen, raza, estado marital, nivel de educación, nivel económico, tipo de vivienda, número de adultos por casa, número de niños por casa, número de propietarios de autos, ciclo de vida familiar, nivel de ingresos, entre otros.

Psicográfica: aquella que describe aspectos cualitativos adquiridos y cuyo objetivo es obtener un mayor entendimiento del cliente. A manera de ejemplo, entre los posibles atributos asociados a la personalidad, podemos encontrar: ella es una persona feliz, opiniones (libertario), creencias (cree en astrología), intereses (le gusta montar en bicicleta), estilo de vida (saludable)

y gustos, entre otros. De igual manera, la información psicográfica puede verse correlacionada con una o múltiples características demográficas.

Comportamental: son aquellos atributos de valoración de datos históricos que reflejan patrones en común. Ejemplos incluyen: patrones de uso de productos, frecuencias de compra, volumen de compra o compra de ciertos productos juntos [2] [23] [24].

Es razonable asumir que para desarrollar un perfilamiento adecuado, se debería modelar y recolectar de forma variada y combinada estos diferentes tipos de atributos.

3.2 Descripción de segmentos

Como se mencionaba en un apartado anterior y según la definición de Smith (1956), la segmentación involucra ver un mercado heterogéneo, como un conjunto de mercados homogéneos, en respuesta a diferentes preferencias y atribuible a los deseos manifestados por los clientes.

Estos segmentos están compuestos por características que describen:

- Segmentos demográficos: pueden organizarse alrededor de cualquier tipo de atributo. Por ejemplo, un segmento demográfico puede incluir todas las mujeres entre los 20 y 30 años de edad que sean casadas y tengan dos o más hijos.
- Segmentos psicográficos: pueden incluir simplemente clientes que les gusta el golf, música clásica, manejar vehículos de lujo o les encanta la aventura.
- Segmentos basados en el comportamiento: pueden incluir aquellos que realizan compras con frecuencia, o personas que llaman al área de atención al cliente más de dos veces por semana, y todo tipo de valoración de datos históricos [12].

Los segmentos de clientes pueden llegar a ser más interesantes cuando están basados en una combinación de atributos demográficos, psicográficos y comportamentales. Un ejemplo puede ser: una mujer entre los 20 y 30 años de edad, con dos o más hijos que le gusta el golf y manejar carros lujosos y llama al servicio al cliente más de dos veces por semana.

Ciertos atributos psicográficos pueden estar correlacionados a ciertos atributos demográficos. Por ejemplo, un grupo de clientes hombres entre los 40 y 60 años de edad con altos niveles de ingresos puede llegar a conducir vehículos lujosos. [8]

Las variables correlacionales más utilizadas son aquellas en las que se relacionan la edad con el género y el nivel socioeconómico del individuo, lo que influye en sus hábitos de consumo e intereses de compra.

Algunos ejemplos de correlaciones son los siguientes [8]:

- ✓ Hombre y mujeres amantes de vivir solos, se refugian en deporte y tecnología.
- ✓ Solos consumen más deporte y televisión.
- ✓ Acompañados presentan poca actividad física, más siestas, más televisión.
- ✓ Acompañados consumen más redes sociales que los solos.

Clientes con una alta posibilidad de pertenecer al grupo de clientes de alta rentabilidad:

- ✓ Hombres con estilo de vida casados y miembros de una familia, con edad 43 o inferior.
- ✓ Hombres solteros, que viven con la familia o solos con una edad entre los 36 y 50 años.
- ✓ Hombres solteros, que viven solos o con la familia y con más de 56 años.

Presumir relación causal a la correlación:

- ✓ Hombres teniendo una crisis de la edad media con medios financieros, comprarán artículos de deporte o autos de lujo.

Cerca de la mitad de los colombianos son usuarios de las redes sociales. Sin embargo, tan solo un 15% de estos usuarios es fan de una marca en Facebook o Twitter. Por consiguiente, las compañías que desean incursionar en redes sociales, deben buscar la manera de fidelizar clientes, ofreciendo experiencias acorde con los intereses y estilo de vida de los consumidores conectados [26].

3.3 Adquisición de datos de clientes

El perfilamiento de usuarios requiere la colección, adquisición y también el análisis de información, por tanto demanda del uso de técnicas relacionadas a cada uno de esos procesos, al igual que analistas de información capacitados para identificar información valiosa para el proceso.

En el proceso de recolección se implementa la acumulación incremental de datos que pueden ser usados para formular atributos comportamentales como frecuencia de compras. Sin embargo, no todos deben ser datos extraídos de fuentes internas; debido a la existencia de numerosas compañías distribuidores y recolectores de datos que obtienen información de diferentes fuentes y proveedores, almacenando estos dentro de formatos que facilitan su recuperación [23] [24].

Los procesos de recolección pueden estar fuera del alcance de las denominadas *start-up*, que según Neil Blumenthal, son “compañías trabajando para solventar problemas, donde la solución no es obvia, y el éxito no está garantizado” [27]. Las cuales no cuentan con la suficiente información o con los suficientes clientes para realizar el proceso de perfilamiento; para estas y otras compañías existen técnicas como perfiles y personas.

3.4 Perfiles y personas

Una de las formas de iniciar con la implementación de técnicas asociadas a perfilamiento de usuarios, es el especular acerca de las características de clientes potenciales, identificando sus personalidades y demás características expresadas en términos demográficos, psicográficos y comportamentales; como resultado se originan las denominadas “Personas”. Posterior a lo anterior, se compara la base de clientes actuales para la definición de estos y se observa qué tan bien los valores de sus atributos y los valores de la "Persona" definida concuerdan, discrepancias entre los valores de los atributos, pueden indicar fallas potenciales en los procesos de negocio que involucran interacción con el cliente.

Perfiles y Personas son usados por el negocio, para validar que los clientes y sus productos estén acorde con la audiencia objetivo. Posteriormente y después de validar los clientes existentes contra las diferentes Personas, se usa el perfilamiento para identificar oportunidades de centralización de clientes que hubieran pasado de forma inadvertida antes de iniciar el proceso [2].

Las técnicas anteriormente descritas requieren del uso de otra serie de procesos analíticos especificados en la siguiente sección.

4. Análisis de datos

El análisis de datos de clientes busca entender la relación entre los valores del negocio, los datos analíticos y las diferentes técnicas usadas, entre las que se encuentran: segmentación, estimación, predicción y grupos de afinidad. Estas técnicas confían en métodos como link análisis, detención de clusters y árboles de decisión, para ayudar en el desarrollo de modelos usables [28].

Muchas personas tienen la falsa impresión de que el análisis de datos de clientes, resultará de forma indudable en la creación de valor por medio de una estrategia de centralización de clientes, pero la realidad muestra que hay dos prerequisites para que esto ocurra:

- El entendimiento de cómo el análisis soporta las oportunidades de negocio para crear valor.
- Manejar el cambio en las organizaciones, para tomar la mejor ventaja de los resultados analíticos.

Las tareas necesarias para cumplir con estos prerequisites son descritas a continuación:

- Identificar el problema de negocio que debe ser resuelto basado en el conocimiento de una oportunidad, junto con el deseo de tomar ventaja de ésta.
- Analizar los datos: diferentes técnicas pueden ser usadas para desarrollar métodos de segmentación como también modelos predictivos. Este paso incluye: recolección de datos, organización, preparación de plataforma, análisis, pruebas, validación e interpretación del resultado para determinar si el análisis entregado supone alguna acción;

en caso afirmativo, alternativas sugeridas son integradas dentro de los procesos de negocio existentes o mediante la creación de nuevos.

- Tomar la acción: sugerido por los descubrimientos durante el proceso de análisis; incrementa los procesos existentes relacionados con atención al cliente e informa a los tomadores de decisiones.
- Medir los resultados: para evaluar el grado en que las modificaciones de los procesos de negocio informados, aborda los objetivos originales.

5. Modelos de perfilamiento

El desarrollo de modelos de perfilamiento de clientes eficientes, es crucial para el mejoramiento de los sistemas que hacen uso de ellos, como es el caso de los sistemas de recomendación; en los cuales, se ha detectado utilizan modelos de perfilamiento que no generan los resultados esperados, debido a los métodos utilizados para definir perfiles de clientes que no describen los intereses del mismo, ni realizan identificación de preferencias, por medio de un análisis exhaustivo y su alimentación con otras fuentes de datos [29]. Modelos que pueden ofrecer una solución no sólo a los problemas presentados con los sistemas de recomendación, si no a su vez los presentados en la sobrecarga de información a lo largo y ancho de la internet, que demanda sitios web adaptables para cada tipo de visitante.

Existen modelos que buscan mejorar el rendimiento de los sistemas de recomendación, basados en el análisis de información comportamental tanto a nivel individual como grupal, analizando clics, inserciones a la canasta, compras y campos de interés. Sin embargo no se alimentan de otras fuentes de datos como las encontradas en la social media, para realizar un análisis exhaustivo de las preferencias de los clientes. En cuanto a los sistemas de recomendación que hacen uso de estos modelos podemos destacar que utilizan información filtrada aplicando técnicas de análisis de datos, como método de solución al problema de ayudar a los clientes a encontrar productos que quieren comprar, originando una lista de predicciones por puntuación de semejanza para un cliente dado.

Las técnicas utilizadas hoy en día en los sistemas de recomendación se dividen en dos categorías [14]:

- Filtración basada en contenido: el sistema provee ítems que son similares a los que el usuario se ha interesado en el pasado.
- Filtrado colaborativo: identifica otros usuarios que han mostrado preferencias similares y provee una lista de productos con interés potencial.

Sin embargo cada uno de los dos métodos tiene sus fortalezas y debilidades; por lo cual, algunos investigadores han optado por realizar combinaciones entre filtración basada en contenido, y filtrado colaborativo, para mejorar la precisión de la recomendación.

A pesar de lo dicho hasta el momento, se identificó que algunas falencias en los sistemas de recomendación aún son causadas por los métodos que definen el perfil del cliente. En el filtrado basado en contenido, únicos o diferentes productos no pueden ser recomendados al cliente debido a que el método usado para crear el perfil del cliente, parte de la información del comportamiento individual y luego recomienda los productos basados en la semejanza entre el perfil y el producto, sin valorar similitudes a nivel grupal, ni analizar transacciones históricas. En el filtrado colaborativo, la naturaleza de los productos no puede ser analizada debido a que al momento de crear el perfil del cliente, parte de los registros de transacciones de compras del cliente y recomienda los productos basado en la similitud entre el perfil y las preferencias de grupos de clientes con características similares, sin tomar en cuenta las particularidades individuales.

Para proceder con la creación de modelos de perfilamiento que partan de la identificación de características grupales e individuales, además que respondan en tiempo cercano al real frente a eventos externos del mundo físico y virtual, detectando y aprovechando intencionalidades de consumo, surge la necesidad del uso de tecnológicas como los motores de inferencia y la alimentación del modelo con fuentes de datos externas como las encontradas en social media.

6. Social media

Las redes sociales se han convertido en el medio propicio, bajo el cual, las personas expresan su inconformidad y opinan frente a productos y servicios ofrecidos por diversas marcas y establecimientos comerciales. Por lo anteriormente descrito, estas acciones se han transformado en una fuente primordial de información por parte de las organizaciones, que desean conocer mejor a sus clientes actuales y futuros. Dicha fuente de información social es más poderosa al ser combinada con información transaccional y de otros orígenes de datos, que permiten obtener una vista comprensible entre las condiciones del cliente y las ofertas del negocio [13]. Con la llegada de estas fuentes de información y de tecnologías como Big data, definida esta última, como aquella involucrada en la solución de problemas asociados a tres tipos diferentes de atributos: volumen, variedad y velocidad, se tiene el poder de acceder a grandes cantidades de información cada vez más variada [15]. Dentro de esta información se encuentran: gustos, deseos personales y campos de interés, entre otros aspectos. Dichos aspectos sirven como fuente de datos, para el desarrollo de estrategias e implementación de técnicas y procesos para la construcción de perfiles que se construyen a partir de un examen exhaustivo de los intereses de los clientes. Perfiles estos, que a su vez pueden servir como fuente primaria de información para impulsar los sistemas de recomendación de campañas, pues pueden adaptarse a las características, necesidades e intereses propios de cada tipo de individuo. De igual manera, esta información puede apoyar los procesos de toma de decisiones, por medio de su análisis [18] [30].

A partir de la información extraída en redes sociales, se efectúa una clasificación del usuario dentro de los segmentos de mercado previamente identificados, a partir de las investigaciones de mercado consultadas. Dichos segmentos serán tenidos en cuenta junto con el monitoreo de eventos sociales, como los “Me gusta” a páginas, y comentarios de tipo “post” dentro de la red

social Facebook, para la generación de hechos, que permitan instanciar y procesar reglas por parte de un motor de inferencia de recomendación de campañas.

7. Motores de reglas

Los motores de reglas, fueron diseñados como una manera de desligar y disminuir la complejidad de los sistemas de información, compuestos por un gran número de particularidades que pueden ser expresadas en términos de reglas de negocio; es decir, separar las decisiones del negocio del código de la aplicación.

Sin embargo, el término motor de reglas puede tender a ser confuso, debido a la existencia de varias definiciones, existiendo evidencia de aplicaciones donde no se capturan las reglas de forma correcta, obligando a los desarrolladores a construir toda la funcionalidad. Como resultado, se presentan aplicaciones no diferenciables frente a desarrollos tradicionales. En particular, una de las definiciones existentes acerca de una regla de negocio viene dado por su propio grupo de usuarios, donde nos habla de “una sentencia que define o restringe algún aspecto del negocio, que tiene la intención de estructurar las afirmaciones del negocio o controlar o influenciar el comportamiento del negocio” [31]. Por otro lado, los motores de reglas de negocio, son aplicaciones de software que contienen definiciones de reglas de negocio.

No obstante, existen algunos autores que realizan distinciones respecto de los motores de reglas, un ejemplo es Ronald G. Ross que los define como aplicaciones de computadora, que son primariamente usadas por las organizaciones para administrar información [32]. Por ello, en su gran mayoría poseen características de conexión con bases de datos, refiriéndose a este último como cualquier mecanismo de almacenamiento.

Algunos motores son utilizados para realizar inferencia. Así por ejemplo, en el caso de que un usuario desee obtener una respuesta frente a algo, éste podrá preguntar al motor de inferencia, el cual a su vez, generará una respuesta posible con base a las reglas del negocio y a las entradas recibidas. El anterior tipo de inferencia es conocido como de encadenamiento hacia atrás o Backward Chaining. Por otro lado, el usuario también puede necesitar obtener una o varias respuestas deducidas a partir de un conjunto de datos ingresados, sin la necesidad de realizar una pregunta específica al motor. En este caso, el tipo de inferencia aplicado es conocido como de encadenamiento hacia adelante o Forward Chaining.

Con respecto a las reglas de negocio implementadas por los motores de inferencia, desde la perspectiva de sistemas de información, son definidas según el Business Rules Group²², como una declaración que define o restringe algunos aspectos del negocio.

²² Organización cuyo propósito es fomentar el entendimiento y estandarización del concepto de regla de negocio.

Entre los motores de reglas encontramos algunos como Drools, el cual es un motor de reglas de negocio que ofrece capacidades de representación y razonamiento, además permite disminuir la complejidad de las aplicaciones compuestas por un gran número de condiciones de negocio, separándolas del código de la aplicación. [33].

La máquina de inferencia puede ser descrita como el programa de computadora que ofrece funcionalidades de representación de conocimiento y razonamiento (*KKR*). La cual se encuentra compuesto de tres principales componentes: Ontología, reglas y datos.

Drools se originó como un tipo específico de motor de inferencia, denominado Sistema de producción de reglas (*PRS*), y se encuentra basado en el algoritmo de *RETE*. Este algoritmo es el centro de razonamiento del motor, fue desarrollado por Charles Forgy por el año de 1974, y es capaz de escalar gran número de reglas y hechos.

Una regla está conformada por una estructura de dos partes (ver figura 1): El motor que encuentra coincidencias, entre hechos y datos contra las reglas definidas en la parte izquierda *LHS*, para deducir conclusiones que terminan en acciones descritas en la parte derecha *RHS*. Este proceso es conocido como *pattern matching*.

```
when
  <conditions>
then
  <actions>;
```

Figura 1. Rules drools (Tomado de la referencia [33]).

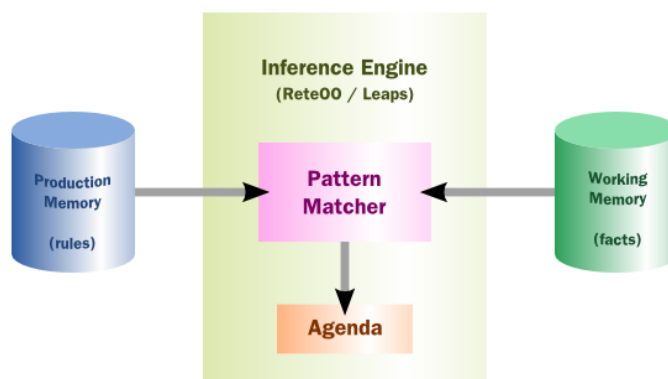


Figura 2. Drools (Tomado de la referencia [33]).

En la gráfica anterior (ver figura.2) observamos los principales componentes del motor de inferencia, que ejecutan acciones basadas en las coincidencias existentes entre los hechos almacenados en la memoria de trabajo (*Working Memory*) y las reglas almacenadas en la memoria de producción (*Production Memory*), proceso conocido como “*Pattern Matcher*”. El orden de ejecución de las consecuencias de las reglas, está dado por la “*Agenda*” [33].

Drools ofrece dos métodos de razonamiento: encadenamiento hacia adelante “*Forward chaining*” (basado en datos) o encadenamiento hacia atrás “*Backward chaining*”. Dadas las características del proyecto, al contar con datos acerca de los usuarios previamente segmentados y con la captura, procesamiento y almacenamiento de eventos sociales y transaccionales, en forma de hecho en el motor de reglas, será implementado el método de razonamiento *Forward chaining* (ver figura.3) [33]. Por otra parte, también serán incluidas a manera de hecho en el motor, las salidas provenientes del procesamiento de eventos sociales, que realiza uso de ontologías para analizar los comentarios del usuario, extraídos de la red social.

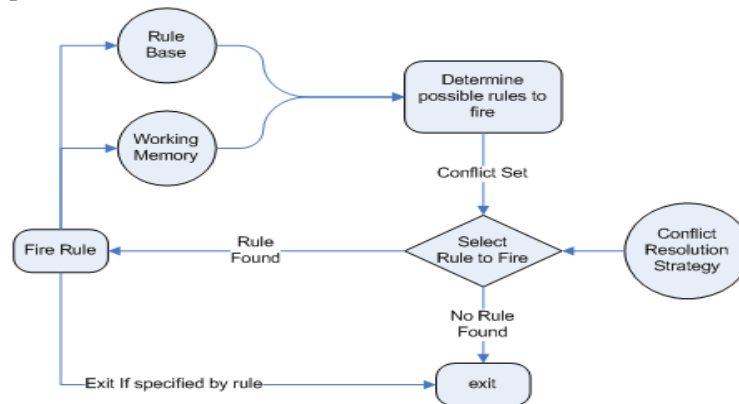


Figura.3 Drools Forward Chaining (Tomado de la referencia [33]).

8. Ontologías en redes sociales.

El crecimiento en el volumen de información generado por medio de redes sociales, ha despertado la atención e interés, tanto de usuarios como de empresas, por extraer, compartir, procesar, generar y usar este conocimiento. Conocimiento que, administrado y aplicado de la forma correcta, puede llegar a convertirse en un bien invaluable. Por lo cual, realizar un uso rentable de la social media se ha convertido hoy en día en un foco de primordial atención al interior de las organizaciones [37]. Sin embargo, gran parte de este conocimiento proveniente de redes sociales no es procesado de la manera correcta debido al desconocimiento en el tratamiento de información de origen no estructurado, ocasionando en algunos casos que sea descartada, sin valorar la ventaja competitiva que esta información le puede suministrar al negocio [35]. Información que una vez procesada, les permitirá mejorar sus estrategias de centralización de clientes, como mecanismo para incrementar la fidelidad de los clientes, reducir las tasas de pérdida de clientes e incrementar las ventas [36].

Con la finalidad de favorecer las estrategias de Multichannel Communication Marketing alrededor de las organizaciones, se propone un enfoque que partiendo del análisis de información

proveniente de comentarios en el muro del usuario conocidos como “*post*”, identifica posibles intencionalidades de consumo que se asocian a un conjunto de características de segmentos previamente descritos.

Existen gran cantidad de casos de aplicación de análisis de información no estructurada para la extracción de conocimiento, que hacen uso de ontologías y redes de ontologías. Uno de estos, es el proyecto SALUS [37] del grupo de Ingeniería Telemática de la Universidad del Cauca que plantea una arquitectura para la anotación semántica automatizada de contenidos en redes sociales, basada en ontologías y vocabularios de dominio. Lo anterior, enfocado principalmente en la definición de métricas para la evaluación de la calidad del contenido en áreas de la salud. Más específicamente, dicho trabajo se concentra en la evaluación de la relevancia de la información, puesto que el objetivo principal de la iniciativa, es poder intercambiar información de forma eficiente entre pacientes y profesionales de la salud. Otro proyecto desarrollado es SoftWiki [38], implementado por Jungmann y Schnefeld en el año 2007, cuya base es Onto-Wiki, lo cual es una iniciativa que integra las capacidades sociales de las personas en un ambiente colaborativo, con el fin de servir como herramienta para la base de conocimiento de las organizaciones, mediante el uso de redes sociales privadas. Para ser más específicos, este proyecto integra:

- ✓ Componentes RDF (término definido por la W3C como un modelo estándar para el intercambio de datos en la Web).
- ✓ RDF Schema (extensión semántica de RDF).
- ✓ OWL (Lenguaje web de ontologías), el cual es un lenguaje diseñado para ser usado por aplicaciones que requieren procesar el contenido de la información, en lugar de solo presentarla de forma humanamente comprensible.

Un tercer proyecto es “Las ontologías son nosotros”. [39] Este nombre del proyecto surge de la pretensión del mismo, de ofrecer un modelo unificado de las redes sociales y la semántica, formulando un esquema genérico y abstracto de las redes semánticas sociales. Específicamente, se basa en una tripleta ontológica compuesta por: actor, concepto e instancia y de igual manera, se incluye el desarrollo de ontologías ligeras expresadas en RDF. Lo anterior, con el fin de desarrollar mecanismos ligeros y de fácil acceso, para la creación de ontologías y metadatos, que se materializan en la construcción de folksonomías. Estas folksonomías se emplean, para designar un sistema de clasificación de objetos web no jerárquicos que nacen de forma natural y democrática por parte de los propios usuarios.

Un cuarto proyecto corresponde a una ontología de extracción usando redes sociales [40]. Esta iniciativa propone ampliar el modelo planteado por Mika (actor-concepto-instancia), para incluir las relaciones entre los actores y los elementos circundantes, relacionando dichos elementos, que pueden corresponder a documentos o datos, a un autor, con el propósito de que estos puedan ser utilizados por colegas o amigos que estén en busca de un mismo elemento. El proyecto tiene componentes más elaborados, como por ejemplo: trabajar con la polisemia de las palabras asociándolas a un contexto adyacente del autor que se ha detectado en su quehacer social en internet.

Por último está el proyecto titulado “Análisis semántico de redes sociales” [41]. Este incluye un proceso de análisis de interacciones en línea, que implementa RDF y SPARQL. Estas extensiones permiten extraer de manera eficiente, parametrizable y de forma semántica, las características de redes SNA (Arquitectura de red utilizada para la conectividad con hosts o mainframes). Lo anterior, haciendo uso de ontologías existentes como FOAF (Ontología que describe personas y sus relaciones dentro de una red social), SIOC (Ontología que describe comunidades en línea), SCOT (Ontología que provee un modelo para expresar conceptos relacionados con actividades a etiquetar) y DOAP (Ontología que describe proyectos de software). La solución planteada en el presente artículo, implementa tripletas ontológicas de la forma: actor, concepto e instancia, con la finalidad de analizar los comentarios provenientes de redes sociales, respecto a la ontología de dominio, y de esta forma establecer una propensión a consumo.

La solución planteada en el presente documento, implementa tripletas ontológicas de la forma: actor, concepto e instancia, con la finalidad de analizar los comentarios provenientes de redes sociales, respecto a la ontología de dominio, y de esta forma, establecer una propensión a consumo.

9. Trabajos relacionados

En el presenta apartado se presentan algunos trabajos relacionados con otros sistemas de recomendación basados en eventos, provenientes de diversas fuentes de datos como la social media.

Con la finalidad de presentar a modo de resumen los principales trabajos relacionados (ver Tabla 1), estos fueron evaluados frente a una serie de principales características entre las que se encuentran: características de respuesta a eventos provenientes de ubicación geográfica, características de social media, características de segmentación, características de inferencia, y características referentes a uso de ontologías para la representación de conocimiento u otro tipo de características semánticas.

Característica	[42]	[43]	[44]	[45]	[46]	[47]	[48]	[49]	[50]
Geoposicionamiento de usuarios						X	x		
Social media	X	X		x	x	X	x	x	X
Segmentación demográfica y psicográfica	X	X	x	x	x	+-			

Inferencia	X	X	x			X			X
Características semánticas	X	X				X	x	x	

Tabla 1. Trabajos relacionados

El símbolo x muestra que la característica es tomada en cuenta, el símbolo +- significa que la característica es abordada de forma parcial.

Respecto a lo descrito en la tabla anterior podemos identificar que gran parte de trabajos relacionados recalcan la importancia de la social media para la extracción de conocimiento valioso y aplicable, que puede ser explotada en diversos campos como el análisis de sentimientos, sistemas de recomendación, perfilamiento y segmentación de usuarios, georreferenciación de usuarios, análisis de tráfico, entre otras aplicaciones.

Entre trabajos relacionados se encuentra el realizado por Chen y Fong [42], el cual propone un Framework de filtrado colaborativo en redes sociales, cuya novedad es la inclusión de medidas de verdad (una recomendación realizada por un familiar tiene mayor valor de verdad que una realizada por otro individuo con el que no se ha mantenido interacción física), por medio de técnicas de minería sobre datos extraídos de la red social Facebook. Valores de verdad que servirán como parámetros de entrada en algoritmos de filtrado colaborativo. En el artículo se destaca lo ya mencionado con relación a la importancia del uso de otras fuentes de datos para incrementar la calidad de las recomendaciones, como la proveniente de redes sociales. Además, de incluir un enfoque de verdad para evaluar el grado de afinidad con la persona o individuo dentro de su red que realiza la recomendación. Los autores consideran, que tiene una mayor relevancia una recomendación realizada por un amigo, familiar o persona cercana con la que se ha tenido contacto físico, frente a otra con la que solo se ha interactuado en un entorno virtual. Dentro del trabajo desarrollado, los autores obtienen y procesan información social de tipo estático, proveniente de datos demográficos, entre la que se encuentran: género, cumpleaños, ciudad de origen, entre otros. Además de información dinámica correspondiente a actividades en redes sociales como: invitaciones, grupos, eventos, blogs, links, frecuencia... Las cuales, se utilizan para medir su relación con otros individuos dentro de su red en seis (6) niveles de verdad.

Por otro lado, está el Framework desarrollado por Kazienko [43], el cual integra varias fuentes de datos para generar recomendaciones personalizadas y relevantes a los miembros de una misma red, con la finalidad de realizar el envío de recomendaciones únicamente a amigos cercanos, compañeros de trabajo o familiares. Igualmente, se utiliza información del perfil del usuario clasificada en estática y dinámica. La primera, la obtiene de forma manual por parte del usuario y la segunda de forma automática por medio del sistema.

Respecto a proyectos relacionados a recomendaciones en tiempo real, encontramos el sistema CP-SMA [47], probado en el evento Milano Design Week. Sistema que realiza recomendaciones de nuevos eventos a usuarios que asisten a la semana de diseño en Milán e interactúan en redes sociales, mediante el análisis semántico de comentarios y la información de los eventos presentados. Sirva de ejemplo, un usuario que indica su ubicación en un evento en redes sociales; seguidamente, el sistema crea un perfil histórico de dicho usuario por medio del análisis de sus comentarios pasados en una red social, hasta este punto el sistema tiene información respecto de los intereses y el evento al que asiste el usuario; posteriormente, basado en las categorías de los eventos, los compara contra la base de intereses extraídos de la red social, para posteriormente recomendar un nuevo evento al que el usuario debería asistir.

Este proyecto combina resultados parciales de las anteriores investigaciones, al utilizar información proveniente de redes sociales y el sector financiero para crear una vista estática del cliente, materializada en la correspondencia a uno o varios segmentos previamente descritos en términos de variables demográficas y psicográficas. Al igual que se incluye una vista dinámica al realizar monitoreo de actividades en redes sociales, específicamente las provenientes de nuevos likes y comentarios realizados en Facebook. De igual manera se incluye dentro del modelo actividades del mundo físico, como las provenientes de transacciones financieras y localización del cliente dentro o cerca de establecimientos de comercio. Lo anterior, con la finalidad de realizar recomendaciones inteligentes y personalizadas mediante la implementación de un motor de reglas, que instancia la vista estática y reacciona frente a eventos producidos desde la perspectiva dinámica del cliente.

10. Conclusiones

Se concluye que al no evidenciar modelos que además de combinar información proveniente de diversas fuentes de datos, permita reaccionar a nuevos intereses y necesidades para el envío de recomendaciones en tiempo cercano al real, de igual modo aprovechar la intencionalidad de consumo manifestadas en el mercado y la generación de aprendizaje acerca del cliente. Se propone la creación de un modelo inteligente de recomendación de campañas basado en perfilamiento de hábitos de consumo, para lo cual se parte de información demográfica del cliente, junto con información psicográfica respecto a intereses manifestados en redes sociales, para así segmentarlos y ofrecer una vista estática respecto a sus hábitos de consumo. Una vista dinámica, que parte del análisis y procesamiento de eventos provenientes de fuentes de datos como: información de actividad en redes sociales (nuevos *likes* y comentarios), información de establecimientos de comercio cercanos respecto de su ubicación geográfica, e información de transacciones financieras. Como resultado, el modelo realiza un análisis exhaustivo de las preferencias y necesidades manifestadas e inferidas por el cliente, y realiza una recomendación inteligente de productos y servicios. En su construcción se aplican herramientas identificadas en el estado del arte como: segmentación, redes sociales, redes de ontologías para el análisis y representación de información y un motor de reglas basado en datos.

IV - CONTRIBUCIONES

1. Modelo inteligente de recomendación de campañas, basado en perfilamiento de hábitos de consumo.

El modelo propuesto incorpora dos componentes: un componente estático de segmentación, que considera aspectos demográficos del cliente (edad, género, estrato, entre otros.) e información psicográfica (intereses). Y un componente dinámico que reacciona frente a nuevas intencionalidades de consumo manifestadas a través de eventos: transaccionales, de actividad social y de ubicación geográfica. Es esto último lo que permite efectuar deducciones de campañas promocionales, en tiempo cercano al real, mediante la instanciación del motor de reglas que realiza dichas deducciones. Estas campañas promocionales son significativas para el individuo, puesto que provienen de los intereses y necesidades expresados por el mismo.

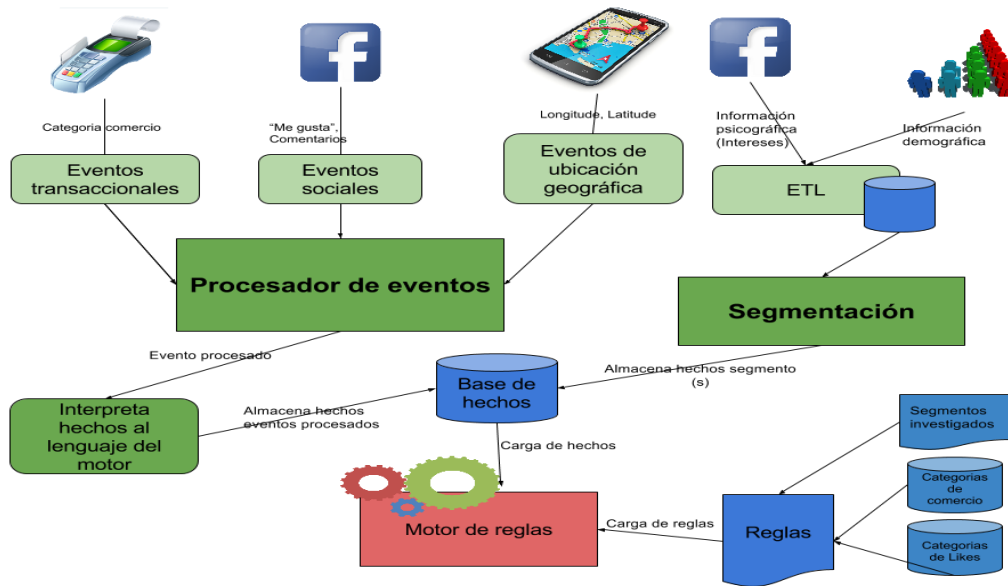


Figura 27. Modelo inteligente de recomendación de campañas.

A continuación, se resume la principal funcionalidad de cada uno de los componentes que conforman la solución propuesta, y como se relacionan y materializan en el motor de reglas.



Figura 12. Componentes de segmentación

En la anterior figura, se aprecia como a partir de los datos demográficos obtenidos del sector financiero y los datos psicográficos correspondientes a intereses obtenidos de social media, se clasifica el usuario dentro de los segmentos de mercado seleccionados, a partir de los estudios realizados por la Cámara de Comercio de Bogotá y la empresa Yanhaas. Lo anterior a fin de, identificar el segmento al que el usuario pertenece, y que será almacenado a manera de hecho en la base de datos de hechos que alimentará el motor de reglas. Sirva de ejemplo, la forma de clasificación y almacenamiento descrita a continuación:

Segmento Solos.

- Edad: entre 18 y 35 años.
- Estado civil: soltero.
- Categorías “me gusta” Facebook: deportes, tecnología y excursiones.

Almacenamiento en base de datos de hechos:

idHecho	idPerson	nombre	descriptor
930	wilisumo@gmail.com	Segmento	Solos
931	wilisumo@gmail.com	Segmento	Open Minded
932	wilisumo@gmail.com	Segmento	Hijos Unicos

Figura 13. Base de hechos, segmentos almacenados.

Componente de eventos sociales.

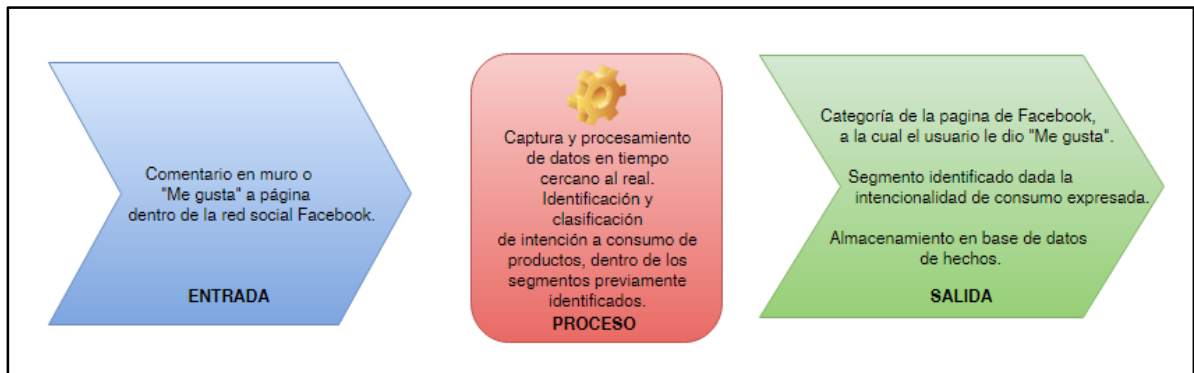


Figura 14. Componente procesador de eventos sociales.

En el presente componente se capturan dos tipos de eventos sociales, los cuales pueden ser descritos en la siguiente lista de actividades:

1. Un nuevo comentario o “me gusta” realizado por el usuario en su “Muro” de Facebook.
2. El sistema detecta si el evento capturado corresponde a un nuevo “me gusta” o un nuevo comentario.

3. En caso de detectar un nuevo “me gusta” el sistema obtiene y almacena su categoría de Facebook, a manera de hecho en la base de datos.
4. En caso de detectar un nuevo comentario, el sistema infiere la categoría que mejor se adhiere a la intención de consumo manifestada por el usuario.

El siguiente ejemplo, sirve para ilustrar mejor el proceso realizado:

Post sin procesar:

“Adoro ir al gimnasio los sábados”

Descripción de la relación en la ontología:

```
<owl:Restriction>
  <owl:onProperty rdf:resource="#adoro"/>
  <owl:someValuesFrom rdf:resource="#Ir_al_gimnasio"/>
</owl:Restriction>
```

Figura 15. Fragmento ontología.

idHecho	idPerson	nombre	descriptor
937	wilisumo@gmail.com	post	Saludable

Figura 16. Base de hechos, almacenamiento hecho comentario.

idHecho	idPerson	nombre	descriptor
933	wilisumo@gmail.com	like	Athlete
934	wilisumo@gmail.com	like	Tours/Sightseeing
947	wilisumo@gmail.com	like	Computers/Technology

Figura 17. Base de hechos, almacenamiento hecho “Me gusta”.

Componente de eventos de ubicación geográfica.



Figura 18. Componente procesador eventos de ubicación geográfica.

En el componente descrito, se captura a partir de una aplicación móvil, la ubicación del usuario en términos de longitud y latitud. Luego, son enviadas en un periodo de tiempo parametrizable de 10 minutos, para ser capturadas en un servidor de aplicaciones mediante un API que recibe la longitud, latitud y correo electrónico del cliente. Que, además consulta en una base de datos de establecimientos de comercio, aquellos que se encuentran en un radio inferior a 500MTS obteniendo su categoría. Finalmente, son ingresados a manera de hecho en el motor de reglas, para su posterior ejecución.

id_commerce	name	CommerceCategory_id_commerce_category	latitute	longitudo
10	CLARO PUNTO DE SUSCRIPCION	23	4.65578	-74.0548
111	DROGAS LA REBAJA	9	4.74417	-74.0569
12	DROGUERIA COLSUBSIDIO	9	4.65698	-74.0544
2	BAR	7	4.74456	-74.0581
2222	RESTAURANTE	5	4.74397	-74.0578
231	DROGUERIA OLIMPICA	9	4.74463	-74.0569

Figura 19. Base de datos de establecimientos de comercio.

id_commerce_category	name
1	AgenciasYAerolineas
10	Educacion
11	ElectrodomesticosYElectronicos
12	General
13	HotelesYClubesSociales
14	Impuestos
15	LibreriasYMediosImpresos
16	RepuestosTalleres
17	Seguros
18	ServiciosClinicosYdeBelleza
19	ServiciosParaElHogarYDecoracion
2	Agropecuario
20	ServiciosPublicos
21	Supermercados
22	Telecomunicaciones
23	TvSuscripcion

Figura 20. Base de datos de categorías de comercio.

idHecho	idPerson	nombre	descriptor	sent
991	wilisumo@gmail.com	estaCerca	TvSuscripcion	0
992	wilisumo@gmail.com	estaCerca	Droguerias	0
993	wilisumo@gmail.com	estaCerca	BaresYRestaurantes	0
994	wilisumo@gmail.com	estaCerca	Droguerias	0

Figura 21. Base de hechos, almacenamiento hecho establecimiento de comercio cercano.

Componente de eventos transaccionales.



Figura 22. Componente procesador de eventos transaccionales.

En lo que respecta al componente de procesamiento de eventos transaccionales, éste extrae la identificación del usuario, y la categoría de comercio en donde se origina la transacción. Además, procesa y almacena su resultado en la base de hechos del motor de reglas (ver figura 24).

idHecho	idPerson	nombre	descriptor
991	wilisumo@gmail.com	Transaccion	TvSuscripcion
992	wilisumo@gmail.com	Transaccion	BaresYRestaurantes

Figura 23. Base de hechos, almacenamiento hechos transaccionales, categoría de comercio.

Una vez almacenados los hechos de segmentación, se activa el motor de reglas con la identificación de cada nuevo evento. Por consiguiente, se ejecuta el algoritmo conocido como “*Pattern matcher*”, que carga la base de datos de hechos a la memoria de trabajo, y realiza instanciación de reglas, para generar deducción basada en *modus ponens*. La cual, en caso de arrojar un resultado afirmativo, realiza envío de mensaje de notificación con la promoción aplicable para el usuario.

idHecho	idPerson	nombre	descriptor
1012	andrea25202015@gmail.com	Segmento	Saludable
1013	andrea25202015@gmail.com	post	Belleza
1014	andrea25202015@gmail.com	like	Health/Beauty

Figura 24. Base de hechos, hechos comentarios y “me gusta”.

```
rule "10% Descuento Servicios clínicos y de Belleza"  
  when  
    $p: Person($hecho : hecho)  
    $e: Email(EnviarEmail ==1)  
    exists (Hecho(person==$p,person.id == $p.id, nombre=="Segmento", descriptor == "Saludable") from $p.hecho)  
    and  
    exists(Hecho(person==$p,person.id == $p.id, nombre=="post", descriptor == "Belleza") from $p.hecho)  
    or  
    exists(Hecho(person==$p,person.id == $p.id, nombre=="like", descriptor == "Health/Beauty") from $p.hecho)  
  then  
    $e.setMessage("10% Descuento en servicios clínicos y de Belleza");  
    $e.setMyTo(""+$p.id);  
    $e.process();  
  end
```

Figura 25. Regla de descuento en servicios clínicos y de belleza.

Como se observa en las gráficas anteriores, los datos almacenados en la base de hechos, concuerdan con los descritos en las reglas del motor, por lo cual se ejecuta la consecuencia de la regla, que para este caso es el envío de un correo electrónico para el usuario al que le aplica la promoción.

2. Componente de segmentación

2.1 Selección de variables

En esta subsección se presenta el resultado de un proceso investigativo, con miras a identificar un conjunto de segmentos en el contexto del mercado colombiano, partiendo inicialmente de la selección de un conjunto de variables demográficas, que según los estudios fueran de mayor relevancia, y que a la vez fuera viable extraer estos datos del sector financiero y de la *social media*. Variables que describen cada uno de los segmentos encontrados.

Entre los criterios de selección utilizados que dieron como resultado final las variables descritas en la siguiente tabla (ver tabla 2), corresponden entre otros factores: cantidad de menciones en publicaciones colombianas, cantidad de menciones en la descripción de segmentos, fuente de datos de fácil disponibilidad o extraíbles del sector financiero o social media.

VARIABLES DEMOGRAFICAS DE MAYOR INFLUENCIA EN LOS HABITOS DE CONSUMO DE LOS COLOMBIANOS.

VARIABLE	DESCRIPCIÓN	DOMINIO
Género.	<p>Esta variable es de relevancia dado que estudios muestran que la posición socioeconómica de la mujer colombiana incide en su disposición y la posibilidad de acceder a diversos servicios. Igualmente, en términos de actitud se observa que la satisfacción de la mujer frente a sus finanzas sigue en aumento, pero sin embargo sigue siendo inferior a la sentida por los hombres [51]. En Colombia el número de mujeres que participan en la responsabilidad del sustento del hogar es creciente [52].</p> <p>Las mujeres se preocupan más por su apariencia física que los hombres [53].</p> <p>Las mujeres colombianas suelen consumir más en vestuario y comida. Sin embargo, esto varía con la edad, una mujer menor a 20 años suele consumir más accesorios, y una ama de casa consume más en el hogar. Por otra parte, los hombres suelen consumir más tecnología y carros, sin embargo, los hombres más jóvenes consumen más en ropa y entretenimiento.</p> <p>Características de producto: las mujeres buscan productos más confiables, durables y de marcas reconocidas [54].</p>	{Masculino, femenino}
Nivel educativo.	Esta variable es de relevancia dado que estudios muestran que la posición socioeconómica de la mujer colombiana incide en su disposición y la posibilidad de acceder a diversos servicios [51]. Estudios muestran que 47% de las mujeres de estratos bajos son amas de casa, la cual se reduce al 20% en los estratos 5 y 6.	{Básico, Técnico, universitario, postgrado}
Edad.	Los hombres jóvenes buscan más emoción y sentimiento a la vida que los hombres de mayor edad. [54] Los colombianos gastan cada vez más su dinero en bienes y servicios que les generen experiencias, este hábito se encuentra principalmente en la población joven. [55]	17-24, 25-39, 40-54, >55.
Localización.	La localización geográfica influye en los hábitos de consumo y en su percepción, como es el caso de un estudio que muestra que en ciudades como Cali, valora más las compañías con compromisos sociales, frente a ciudades como Barranquilla que valora más la calidad.[56]	Longitud, latitud.

Estilo de vida.	<p>Esta variable es relevante en el mercado colombiano dado que estudios muestran que cerca de la mitad de los colombianos son consumidores de redes sociales, y buscan experiencias acordes con su estilo de vida e intereses [57].</p> <p>Las personas solas tienden a tener un estilo de vida saludable, los cuales desarrollan una visión a largo plazo, para lo cual están tomando medidas para asegurar una vejez saludable. Son personas más precavidas frente a la calidad de los productos que consumen [58].</p> <p>En los últimos 13 años en Colombia se ha quintuplicado el número de los hogares unipersonales, lo cual demuestra el potencial para las empresas colombianas.</p> <p>Los acompañados consumen más televisión y hacen menos deporte que los solos [59]. Para los solos su estilo de vida exige mayor autonomía por lo cual consumen más tecnología, en su mayoría computadores portátiles. Al igual, que utilizan redes sociales como Facebook.</p> <p>Personas casadas y de menos de 43 años son clientes potencialmente más rentables [60].</p>	Intereses.
Ocupación.	<p>Esta variable es de relevancia para el consumidor colombiano, dado que estudios demuestran que influye en sus hábitos de consumo, como es el caso del consumo de alimentos fuera de casa [61].</p> <p>Cada vez existe un número mayor de mujeres que trabajan y sustentan el hogar, por lo cual tienden a describirse como responsables, trabajadoras e independientes. Por otro lado, a nivel de consumo tienden a adquirir más productos de belleza que las amas de casa [59].</p>	Empleado, independiente, desempleado.
Nivel de Ingresos.	Estudios demuestran que las mujeres de bajos ingresos ante excesos de liquides, destinan el dinero al cubrimiento de necesidades básicas. Por otro lado, las de estratos medios a productos o servicios y las de estratos altos, al placer y auto gratificación [51].	{Bajo, medio, alto}.
Estrato.	<p>Estudios muestran que 47% de las mujeres de estratos bajos son amas de casa, lo cual se reduce al 20% en los estratos 5 y 6 [52].</p> <p>Las personas de estratos más bajos tienen posibilidades reducidas de invertir y ahorrar, dado que expresan en su mayoría la necesidad de gastar más de lo que ganan, por lo cual requieren mayores facilidades de pago, con cuotas y flujos de dinero pequeños [62].</p>	1, 2, 3, 4, 5 y 6.

Cultura	Variables como el tradicionalismo en la división de roles, que vincula a la mujer con labores y la economía doméstica, la han alineado a la “Macroeconomía del hogar”, siendo un tema masculino [51]. Las recomendaciones de amigos y colegas influyen en la decisión de compra de algunos individuos. [63]	Tradicional, abierta, open mind,
---------	---	----------------------------------

Tabla 2. Descripción de variables seleccionadas.

Las variables demográficas pueden ser combinadas para caracterizar la estructura de grupos de una población [60].

2.2 Definición de segmentos

En esta subsección se presenta los resultados obtenidos, de la realización de un proceso investigativo, cuya finalidad es la identificación de un conjunto de segmentos emergentes en el contexto del mercado colombiano, que cuenten con atributos suficientes que describieran al segmento y permitan el proceso de clasificación de usuarios.

Entre los criterios utilizados para la selección de los segmentos identificados, están la calidad de las fuentes que definen los segmentos en el contexto del mercado colombiano, claridad en su definición, fiabilidad de los datos y la trayectoria de la compañía que identifico dichos segmentos. En este proyecto se seleccionó como fuente de datos, Yanhaas empresa colombiana con más de 20 años de trayectoria en la elaboración de estudios de mercado, y la cámara de comercio de Bogotá. Las cuales identificaron los siguientes grupos de segmentos con sus correspondientes características (ver Tabla 3).

Segmento	Características
Saludables	<p>Consume productos que le motivan a cuidar su figura, y que tengan buen sabor.</p> <p>Buscan llegar a la vejez con buen estado de salud. Precavidos en escoger la calidad de los productos que consumen.</p> <p>Segmento orientado a la salud.</p> <p>Servicios preferenciales en salud (medicina prepagada)</p> <p>Piensan en el bienestar actual y futuro. [58]</p>
Solteros	<p>Interesados en pasar más tiempo al aire libre y realizar actividades deportivas.</p> <p>Estilo de vida de mayor grado de autonomía, que se ve reflejado en un mayor consumo de tecnología, internet y servicios relacionados. [59]</p>
Sensibles	<p>“Mujeres de 25 a 40 años que se aferran a lo estético. Manejan un alto nivel de conocimiento por la moda, la decoración y las tendencias. Buscan vivir en un mundo de estímulos no sólo visuales, sino en todas las experiencias: amistades, alimentación y diversión. Disfrutan de su trabajo y aman salir de compras.” [64]</p>
Creativos totales	<p>"Son mujeres y hombres de 20 a 35 años. Están inmersos en un mundo donde todo es interesante y siempre están buscando novedades para inspirarse. Manejan horarios flexibles y trabajan en áreas diversas con un componente fuerte de tecnología, pero al mismo tiempo son muy creativos. Capaces de atravesar la ciudad para ver algo que pueda darles un estímulo creativo [64]"</p>

Abuelos Activos	“Crecieron en los años 60. Hombres y mujeres de 60 a 70 años. Fueron rebeldes y ahora reinventan lo que significa la tercera edad. No son los típicos abuelitos. Se mantienen activos, estudian y buscan disfrutar de su jubilación [64]”.
Sabelotodo	“Adolescentes y jóvenes de 12 a 20 años que lo saben todo sobre moda y diseño, y utilizan ese conocimiento para estructurar su identidad. Siempre quieren estar a la moda de acuerdo con su propio estilo y viven atentos a la siguiente tendencia. Les encantan la Internet, el cine y las telenovelas. Conocen los productos de su época y de su generación[64]”.
Super Mujeres	<p>Segmento conformado por mujeres que trabajan. Encuentra la mayor fuente de satisfacción tanto en el hogar, como fuera de este.</p> <p>Sobresalen en el consumo mayoritariamente de artículos de belleza. Debido a que permanecen fuera del hogar la mayor parte del tiempo [52].</p>
Hijos únicos	“Mujeres y hombre de 20 a 35 años. Han creado un estilo de vida como el de un niño consentido, inmersos en ellos mismos, narcisistas y los mejores clientes de todo lo último en tecnología. Viven a través de espacios virtuales, como Facebook. Son exigentes y prefieren que sus productos sean únicos. Por eso optan por las nuevas colecciones [64]”.
Independientes	Están entrando a su cuarta década con independencia y gran conocimiento. Son femeninas, pero a la vez fuertes. Han aprendido cómo equilibrar los diferentes componentes de sus vidas y en ellas la familia juega un papel muy importante. Les gustan las artesanías y lo étnico, y se interesan por el detalle en su ropa y sus accesorios [64]”.

Nuevos ricos	“Hombres de 45 a 60 años. En sus carros y en su ropa dejan ver lujo aristocrático y aspiraciones de nuevo rico en un país en desarrollo. Buscan calidad, especialmente en lo que tiene que ver con su vestuario y accesorios. Son miembros de grupos exclusivos. Leen obsesivamente y consultan muchas fuentes de información antes de consumir algún producto o servicio. Prefieren lo importado” [64].
Intelectuales	“Apasionados por la lectura y el estudio, y amigos de la tecnología. Mujeres y hombres de 35 a 50 años. Son jóvenes profesionales que trabajan en lo que aman. Han viajado por el mundo, tienen influencias de muchas disciplinas y culturas, y esto lo demuestran en sus gustos en la cocina y el cine. Están siempre en la búsqueda del nuevo restaurante extranjero para probar cosas nuevas” [64].
Nuevos rebeldes	“Son mujeres y hombres de 45 a 60 años que se identifican como los nuevos rebeldes. Tienen su propia visión del mundo, de la política y de la cultura, y van en contra de la globalización. Esto se nota en los alimentos que consumen, en los lugares que frecuentan y en los empleos que tienen. Les gusta estar en casa y crean espacios para sus pasatiempos y aventuras. Se inclinan por la ecología” [64].

Tabla 3. Descripción segmentos.

A través de las variables demográficas seleccionadas como de mayor relevancia para la definición de segmentos, se procederá con la descripción de los segmentos previamente identificados en términos de estas variables y de atributos psicográficos, que serán extraídos de la actividad del usuario en redes sociales, de forma específica de interés manifestados a partir de “me gusta”. Lo cual, formara parte del componente estático del modelo de perfilamiento, materializado a manera de hechos.

3. Carga de datos demográficos y procesamiento de eventos transaccionales.

En esta sección se presenta el proceso de extracción y procesamiento, de información de carácter estructurado, proveniente del sector financiero transaccional. La cual, fue utilizada para la realización del proceso de segmentación. Entre los datos capturados encontramos: género, edad, localización, estrato, nivel de ingresos, ciclo de vida familiar y estado civil.

En la figura.26 podemos observar el componente de procesamiento de eventos transaccionales y el componente de segmentación. El primero encargado de recibir y procesar los eventos generados por el servicio del sector financiero, que notifica al sistema cada vez que detecta una nueva transacción. Al mismo tiempo, que realiza envío de identificación de usuario y categoría de comercio de donde se origina dicha transacción.

El segundo componente se encarga de procesar los datos demográficos del cliente, y contempla las actividades a continuación descritas:

1. Obtener archivo con información demográfica de clientes del sector financiero.
2. Almacenar información demográfica por usuario, en un data storage temporal.
3. Realizar procesamiento de datos, omitiendo registros incompletos y sintácticamente incorrectos.
4. Cargar datos de clientes.
5. Combinar información demográfica de clientes, con información de intereses proveniente de Facebook.
6. Clasificar clientes, basado en los segmentos de consumo previamente descritos en términos de variables demográficas (CSV sector financiero) y psicográficas (intereses Facebook).

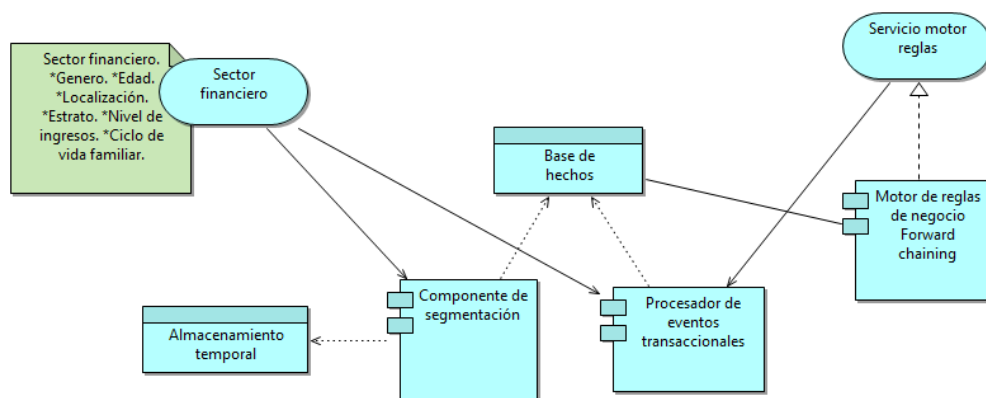


Figura 26. Modelo de componentes procesador eventos transaccionales

4. Componente de procesamiento de eventos.

4.1 Componentes, procesos y técnicas del procesador de eventos de ubicación geográfica.

En esta subsección se presentan los principales componentes, procesos y técnicas implementadas para la construcción del componente de procesamiento de eventos.

Entre los elementos que interactúan con el componente de procesamiento de eventos encontramos:

- Aplicación móvil: se encuentra desarrollada bajo la plataforma Android, cuya única característica y funcionalidad es la captura de la cuenta de usuario o correo electrónico asociadas al teléfono donde se hospeda dicha solución, y la ubicación del usuario en términos de longitud y latitud.
- Base de datos: cuenta con información de establecimientos de comercio, con su correspondiente ubicación y categoría. Ubicación dada en términos de longitud y latitud.
- API: interfaz que recibe y procesa la ubicación del usuario enviada por la aplicación Android hospedada en el celular del cliente, en términos de longitud y latitud, junto con su correo electrónico. Basado en dicha información, esta calcula los establecimientos de comercio en un radio de 500 Metros, posteriormente obtiene la(s) categoría(s) de comercio, cercanas al punto de ubicación e ingresa la información en la base de datos de hechos del usuario.

Una mejor aproximación a como se relacionan e interactúan los elementos anteriormente descritos, se puede apreciar en la siguiente gráfica:

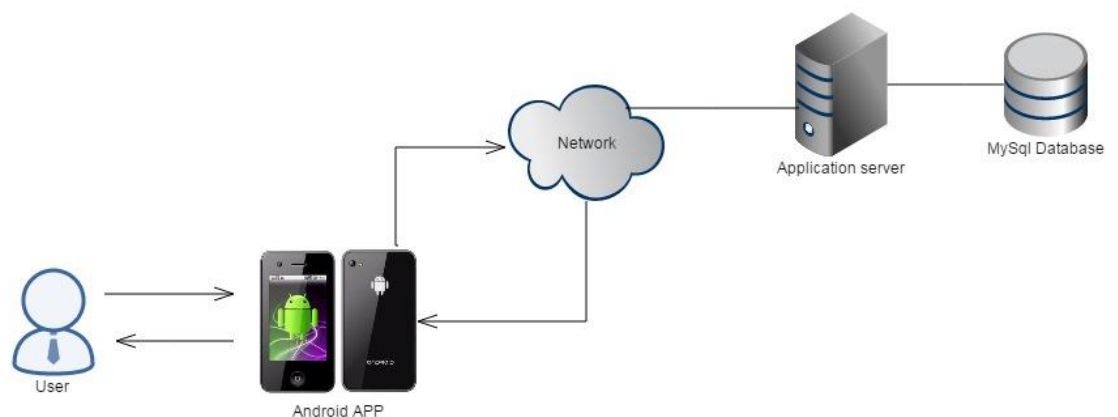


Figura 27. Arquitectura cliente servidor, componente ubicación geográfica.

A continuación, se describe el flujo del proceso ilustrado en la anterior gráfica (ver figura 27):

1. Identificar comunicación con el servidor de destino.
2. En caso de una respuesta exitosa capturar: correo electrónico y ubicación del dispositivo en términos de longitud y latitud.
3. Enviar los datos capturados al servidor de aplicaciones.
4. La aplicación alojada en servidor, recibe los datos enviados.
5. La aplicación consulta las categorías de los comercios en un radio menor o igual a 500 Mts de la longitud y latitud recibidas.

Finalmente, las categorías de comercio obtenidas son almacenadas en la base de hechos del motor de reglas. Se debe subrayar que, con la finalidad de manejar una menor proporción de valores discretos para la construcción de reglas, se optó por seleccionar categorías de comercio, en lugar de establecimientos de manera individual.

4.2 Procesador de eventos Sociales

4.2.1 Procesamiento de comentarios

El primer paso en el componente de procesamiento de comentarios, consiste en su extracción. Para ello, se requirió la construcción de una aplicación que implementa el Api de Facebook en su versión 2.3. El cual, inicialmente realiza la autenticación del usuario, y en caso de no existir registro previo, redirige la solicitud al portal de autenticación (ver figura. 28). Una vez la autenticación se realiza de forma satisfactoria, se solicita acceso a un conjunto de permisos necesarios para la extracción de la información requerida, entre los que se encuentran: “email”, “user_likes”, “user_post”, “user_status”, entre otros (ver Figura.29).

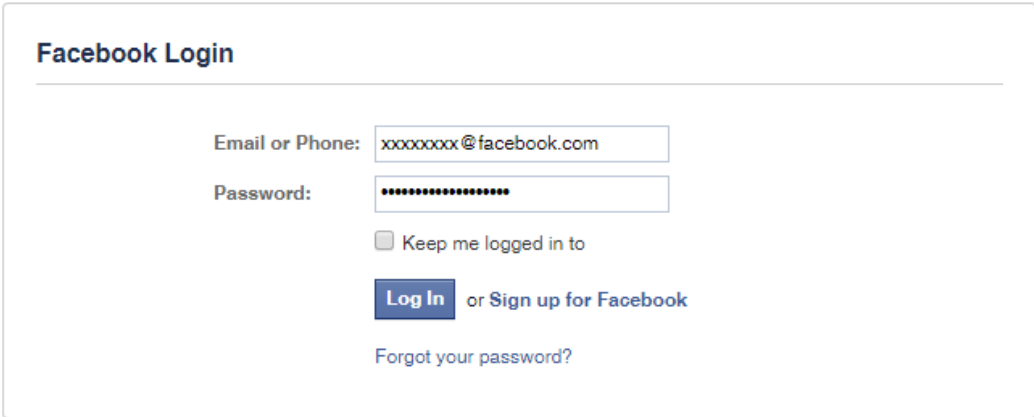


Figura 28. Pantalla de autenticación Facebook

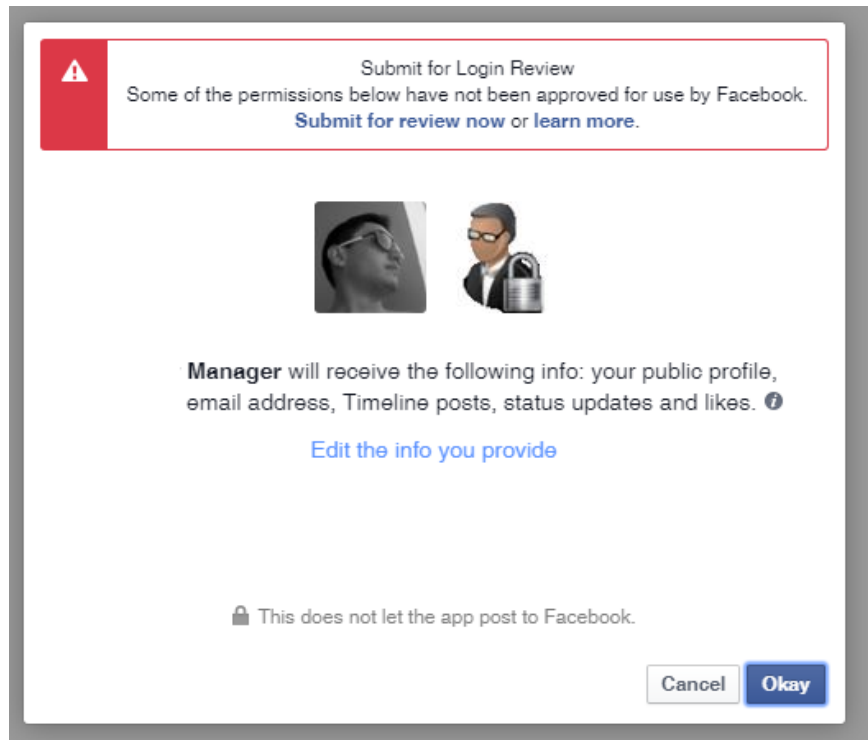


Figura 29. Pantalla de solicitud de permisos

Como segundo paso, se consulta la base de datos de verbos y sinónimos, descritos en la ontología desarrollada que se encuentra especificada en la siguiente subsección, y se comparan con los contenidos en los comentarios. En caso de coincidir y/o hallarse alguno de estos, es extraído el valor de la relación de la ontología, para posteriormente formar parte de la consulta SPARQL (lenguaje de consulta para RDF) [65].

Dentro de la base de datos de verbos y sinónimos podemos encontrar aquellos que demuestran una intención o tendencia positiva a adquirir o consumir ciertos productos o servicios. Al igual que los verbos, las preposiciones son consultadas, buscadas y comparadas, contra la base de datos de preposiciones. Dicha base de datos tiene una estructura conformada por: un primer atributo cuyo valor es sintácticamente igual al descrito en la ontología, al que se le denomina “OntologyValue”. Por otro lado, también cuenta con un atributo cuyo valor contiene un sinónimo del primer atributo, al cual se le denomina “Synonymous”, el cual será comparado contra el texto proveniente del comentario. Para comprender mejor, se pueden apreciar las siguientes tablas:

Base de datos de predicados:

OntologyValue	Synonymous
Asistir_a_concierto	asistir al concierto
bajarPeso	disminuir peso
comida_vegetariana	comida sin calorías
música	escuchar la radio
Deporte	sport
Cine	pelicula

Tabla 4. Base de datos de predicados.

Base de datos de verbos:

OntologyValue	Synonymous
Adoro	amo
Amo	considero
Asiste	llego
Come	engullo
Deseo	ofrezco
Necesito	carezco
Orgullosa	vanidoso
Quiero	deseo

Tabla 5. Base de datos verbos.

Una vez comparado el contenido del comentario, contra el atributo “Synonymous” y en caso de encontrarse sintácticamente igual al descrito en la ontología (figura 30), se procede con la extracción del valor del primer atributo denominado “OntologyValue” (Tabla. 4). De este modo, si obtenemos por parte del usuario el comentario: “deseo asistir al concierto del domingo”, el algoritmo realiza el siguiente proceso:

1. Compara el contenido del comentario con el valor del atributo “Synonymous”.
2. En caso de encontrar valores sintácticamente iguales los extrae. Por lo cual, extrae el valor “deseo” y “asistir al concierto”.

3. Consulta y obtiene el valor del atributo “OntologyValue”, para el atributo con valores en “Synonymous” “deseo” y “asistir a concierto”.
4. Para el caso del ejemplo, extrae el valor “quiero” y “Asistir_a_concierto” correspondientes al atributo “OntologyValue”.

```
<owl:Class rdf:about="http://bisoft.co/suarez-william/ontologies/#Asistir_a_concierto">
  <rdfs:label>Asistir a concierto</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://bisoft.co/suarez-william/ontologies/#Sentimiento"/>
</owl:Class>

<owl:Restriction>
  <owl:onProperty rdf:resource="http://bisoft.co/suarez-william/ontologies/#quiero"/>
  <owl:someValuesFrom rdf:resource="http://bisoft.co/suarez-william/ontologies/#Asistir_a_concierto"/>
</owl:Restriction>
```

Figura 30. Fragmento Owl de la ontología creada

Finalmente, basado en los valores de los atributos obtenidos de “OntologyValue”, el algoritmo procede con la ejecución de una consulta SPARQL (ver figura 31), con la finalidad de inferir la categoría(s) o segmento(s) orientado a consumo al que pertenece la intencionalidad de consumo manifestada en el comentario.

```
select ?screb where {?screb a owl:Class ;
                      rdfs:subClassOf ?restriction
                      ?restriction owl:onProperty screb:quiero
                      owl:someValuesFrom screb:Asistir_a_concierto
}
```

Figura 31. Consulta SPARQL sobre la ontología creada.

Dentro de las categorías o segmentos que conforman la ontología y describen una intencionalidad orientada a consumo se encuentran: Saludables, Soltero, Solos, Sensibles, Super mujeres, entre otros.

4.2.2 Red de ontologías

En el presente apartado se presenta, el resultado del desarrollo de una red de ontologías que permite la exploración y explotación de conocimiento proveniente de social media. Además, de fuente de datos, para la categorización o segmentación de usuarios, identificando intenciones de compra. Como red estará conformada por la implementación de ontologías de varios dominios, entre las que se encuentran:

- a. SIOC (Semantically Interlinked Online Community). Permite la disseminación y transferencia eficiente de información alrededor de comunidades sociales en línea [67]. En particular, se aplicó dado que contiene clases como: `user_account` y `post` que describe cuentas y atributos que pueden ser encontrados en comunidades en línea como Facebook, Twitter, entre otras. Además de contar con propiedades como: `email`, `avatar` y `account_of` para la identificación de usuarios.
- b. GOODRELATIONS. Ontología que describe productos y servicios ofrecidos en la web, cuyo finalidad es la de proveer un espacio por medio del cual clientes, y compañías puedan encontrar productos y proveedores adecuados [68]. En concreto, será implementado para especificar los productos y servicios que describen intencionalidad de consumo dentro de determinadas categorías. Sirva de ejemplo, la categoría “Solos”, la cual puede ser inferida a partir de los usuarios que consumen productos de proveedores de entretenimiento dentro de la categoría de producto y/o servicio “Bares”.
- c. FOOD. Ontología para caracterizar y categorizar el tipo de comidas consumida por una persona, para ello se implementa la Ontología FOOD [69]. Igualmente, fue implementada para categorizar intencionalidades de consumo. Así pues, se puede inferir un usuario “Saludable” a partir de la existencia de una instancia a la clase `Food`, categoría comida saludable.

Ontologías que formarán parte de los recursos ontológicos enmarcados en la metodología de construcción de ontologías Neon, la cual nos permite especificar el contenido de la ontología, identificar objetivos, requerimientos funcionales, fuentes de información ontológicas y no ontológicas, escenarios de uso, usuarios esperados y demás información, que servirá de base para una mejor contextualización y posterior construcción de la ontología [70]. Los resultados obtenidos tras su implementación pueden ser observados en la siguiente tabla.

ONTOLOGY REQUIREMENTS SPECIFICATION DOCUMENT	
1.	Propósito
	El propósito de la ontología es el de proveer un modelo de conocimiento de usuarios, sentimientos e intenciones de adquisición de productos y/o servicios manifestadas, en

	el dominio de redes sociales, que pueda ser utilizado por un modelo inteligente de recomendación de campañas promocionales.
2	Alcance
	La ontología se focalizara en el dominio de redes sociales tales como Facebook y Twitter, basado en identificación de intereses.
3	Lenguaje de implementación
	La ontología será implementada en lenguaje OWL.
4	Usuarios Finales Previstos
	<ul style="list-style-type: none"> • Empresas que requieren realizar recomendaciones promocionales a sus clientes. • Empresas que desean analizar las propensiones de consumo entre sus clientes. • Empresas que desean analizar los intereses de sus clientes expresados por medio de social media. • Usuarios que desean extraer información de preferencias e intereses provenientes de social media, manifestados por medio de lenguaje natural.
5	Usos previstos
	<ul style="list-style-type: none"> • Envío de recomendaciones promocionales basado en identificación de intereses, obtenidos mediante el procesamiento de mensajes provenientes de social media. • Segmentación de usuarios registrados en redes sociales. • Extracción de compradores potenciales de ciertos productos y/o servicios. • Análisis de sentimientos. • Identificación de preferencias y tendencias de consumo en determinados establecimientos de comercio, provenientes de ciertos segmentos poblacionales.
6	Requerimientos Ontológicos
	a. Requerimientos no Funcionales
	1. La ontología soportara el lenguaje español y contendrá notaciones semánticas en inglés para algunos conceptos.

	b. Requerimientos Funcionales: Grupos de preguntas competentes
	<p>USO 1: Realizar Búsquedas</p> <ul style="list-style-type: none"> • ¿Cuál es la identificación del usuario? • ¿Cuáles son los intereses expresados por el usuario? • ¿El usuario practica algún deporte? • ¿El usuario gusta de alguna categoría de comida en especial? • ¿A qué tipo de categorías de comercio, asiste el usuario? • ¿Qué tipo de sentimientos son descriptivos o pertenecientes a cada segmento? • ¿A qué segmento pertenece el usuario?
7	Glosario de Términos descriptivos
	a. Términos para preguntas competentes
	<p>Correo electrónico</p> <p>Interés manifestado</p> <p>Deporte practicado</p> <p>Categoría o tipo de comida</p> <p>Categoría de comercio</p> <p>Sentimiento (Querer, amar, desear)</p> <p>Segmento (Saludable, Hijo Único, Independiente)</p>
	Respuestas
	<p>Correo electrónico</p> <p>Bajar de peso, comprar artículos de lujo, Comprar automóvil, Comprar computador, comprar ropa, comprar muebles, escuchar música, hacer deporte, amo ir de compras, quiero ir cine, adoro ir a spa, entusiasmado por comprar computador, adoro el gimnasio, necesito mejorar mi salud, con ganas de jugar futbol, me gusta el aire libre, me gusta conectarme a internet, me gusta ir al bar, amo tomar fotos, entre otros.</p> <p>Futbol, natación, atletismo, baloncesto, tennis, boxeo, entre otros.</p>

<p>Comida Saludable, comida mexicana, comida colombiana, comida con condimentos, comida sin grasas, comida con grasas...</p> <p>Bares, restaurantes, clubes, joyería, tecnología...</p> <p>Amar, querer, desear, necesitar, orgullo, necesidad, enamorado, feliz...</p> <p>Saludable, mente abierta, Super mujeres, intelectuales , solos, solteros, recién casados, nuevos rebeldes, unique sons, sensibles, abuelos activos, intelectuales, independientes, creativos totales, sabelotodos, nuevos ricos.</p>

Tabla 6. Especificación requerimiento ontología metodología Neon

Entre los recursos ontológicos y no ontológicos, utilizados para la elaboración de la red de ontologías encontramos:

Recursos ontológicos:

- FOAF Ontology: Friend of a friend (<http://xmlns.com/foaf/spec/>)
- GOOD RELATIONS Ontology (<http://purl.org/goodrelations/v1>)
- SIOC ontology (<http://rdfs.org/sioc/spec/>)
- FOOD ontology (<http://data.lirmm.fr/ontologies/food#>)
- Ontology design Patterns (http://ontologydesignpatterns.org/wiki/Main_Page)

Recursos no ontológicos:

- YANHAAS

Descripción de segmentos:

<http://yanhaas.com/index.php/yanhaas-poll/item/234-yanhaas-poll-no-68>

<http://yanhaas.com/index.php/yanhaas-poll/item/243-yanhaas-poll-no-78>

<http://yanhaas.com/index.php/yanhaas-poll/item/232-yanhaas-poll-no-66>

<http://yanhaas.com/index.php/yanhaas-poll/item/231-yanhaas-poll-no-65>

- GRUPO RADAR (http://www.gruporadar.com.uy/01/?page_id=155)

Para el proceso de construcción y desarrollo de la ontología, se procedió inicialmente con la identificación de los objetivos, requerimientos y escenarios de uso. Posteriormente, se construyó una red de ontologías de forma tal, que permitiera describir un conjunto de segmentos a través de relaciones y conceptos. De esta forma, y tomando como referencia el comentario realizado por el usuario, se infiere o deduce la categoría que manifiesta cierta intencionalidad de consumo.

Este proceso se realiza mediante el uso de Tripletas sujeto + verbo + predicado., las cuales describen los tipos de sentimientos e intereses manifestados hacia la adquisición de determinados productos o servicios. Al igual, que proveen en términos de relaciones entre conceptos denominados ObjectProperties y verbos, manifestaciones de interés hacia el consumo de determinados productos y/o servicios. El siguiente ejemplo sirve para ilustrar mejor lo expuesto:

:Person adoro : Cine

:Person Ir_a Cine

:Person amo :Hacer_deporte

Tripletas que a su vez permiten inferir la intencionalidad de compra hacia ciertos tipos de productos y/o servicios que se encuentran enmarcados dentro de segmentos de consumo, al estar los segmentos descritos a manera de relaciones que contienen características compartidas entre las diferentes categorías a las que puede pertenecer un individuo (ver fig. 9).



Figura 32. SubClase Segmento.

Descripción de la intencionalidad de compra hacia ciertos productos y servicios por la categoría o segmento Solos:

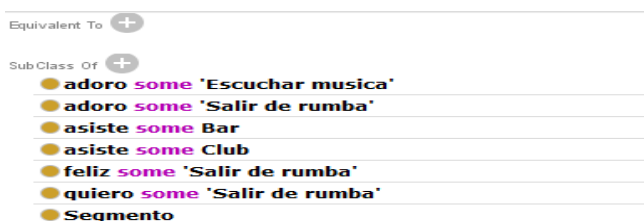


Figura 33. Relaciones subclase solos.

De esta forma y mediante la implementación de “JENA”, *Framework* open source basado en Java para la construcción de web semánticas y aplicaciones de Linken data, y mediante el componente ARQ, procesador de SPARQL (Lenguaje de consulta en RDF desarrollado por la W3C) para Jena [72]. Posteriormente, se realiza una consulta semántica para inferir el segmento de mercado al que muestra intencionalidad de consumo determinado individuo mediante el procesamiento en tiempo real de los comentarios obtenidos de Facebook.

5. Componente de inferencia.

5.1.1 Motor de reglas:

En esta subsección se describen los principales componentes, procesos y técnicas implementadas para el diseño y construcción del motor de reglas, y como este se relaciona e interactúa con los demás componentes que hacen parte de la solución propuesta.

Retomando lo detallado en la sección 3, en lo que respecta a motores de reglas, específicamente los basados en datos (Forward Chaining), los cuales ejecutan acciones basados en inferencia modus ponens entre los hechos almacenados en la memoria de trabajo (Working Memory) y las reglas almacenadas en la memoria de producción (Production Memory) o proceso conocido como “Pattern Matcher”. Por otro lado, dadas las características del proyecto el cual parte de un conjunto de datos preexistentes y/o expresados en tiempo real, como: gustos, intereses, características y necesidades obtenidas e inferidas, en tiempo cercano al real, se desea determinar la campaña aplicable para cada tipo de usuario, llevando el envío de campañas masivas a un mayor grado de individualización. Por lo cual, se determinó seleccionar el motor de reglas de negocio Drools, el cual cumple con todas las funcionalidades requeridas para el desarrollo del proyecto.

El motor de reglas, parte de la información almacenada en la base de datos de hechos, la cual contiene el resultado del proceso de segmentación (componente estático del modelo) y los eventos descritos en el lenguaje del motor de reglas (componente dinámico del modelo). Para posteriormente, inferir utilizando la lógica proposicional modus ponens, la campaña que mejor se adhiere a las necesidades y nuevos intereses manifestados por el usuario. Sirva de ejemplo, la siguiente regla:

```
rule "10% Descuento Servicios clínicos y de Belleza"
when
  $p: Person($hecho : hecho)
  $e: Email(enviarEmail ==1)
  exists (Hecho(person==$p,person.id == $p.id, nombre=="Segmento", descriptor == "Saludable") from $p.hecho)
  and
  exists(Hecho(person==$p,person.id == $p.id, nombre=="post", descriptor == "Belleza") from $p.hecho)
  and
  exists(Hecho(person==$p,person.id == $p.id, nombre=="like", descriptor == "Health/Beauty") from $p.hecho)
then
  $e.setMessage("10% Descuento en servicios clínicos y de Belleza");
  $e.setTo(""+$p.id);
  $e.process();
end
```

Figura 34. Regla de 10% de descuento en servicios clínicos y de belleza.

Conforme a la anterior regla, en esta se identifican dos tipos de hechos: estáticos o de segmentación y dinámicos provenientes de eventos de actividades sociales. Específicamente, asocia en un regla de verdad con conectores lógicos AND: el segmento saludable, el comentario clasificado dentro de la categoría de “Belleza” y la categoría del interés “Health/beauty”. En caso de arrojar una valoración lógica positiva, se procede con él envío de la promoción aplicable al usuario en este caso: “10% de descuento en servicios clínicos y de belleza”.

idHecho	idPerson	nombre	descriptor
1012	andrea25202015@gmail.com	Segmento	Saludable
1013	andrea25202015@gmail.com	post	Belleza
1014	andrea25202015@gmail.com	like	Health/Beauty

Figura 35. Hechos almacenados en la base de hechos.

V - DISEÑO

En el presente apartado se describe como se integran y se relacionan todos los componentes del sistema resultante, como propuesta de solución a la problemática planteada a lo largo del documento, con la explicación de cada una de las entradas y salidas de cada uno de los componentes del sistema.

1. Vista caso de uso arquitectónico

El siguiente diagrama, presenta el caso de uso que guía el diseño de la arquitectura de la aplicación.

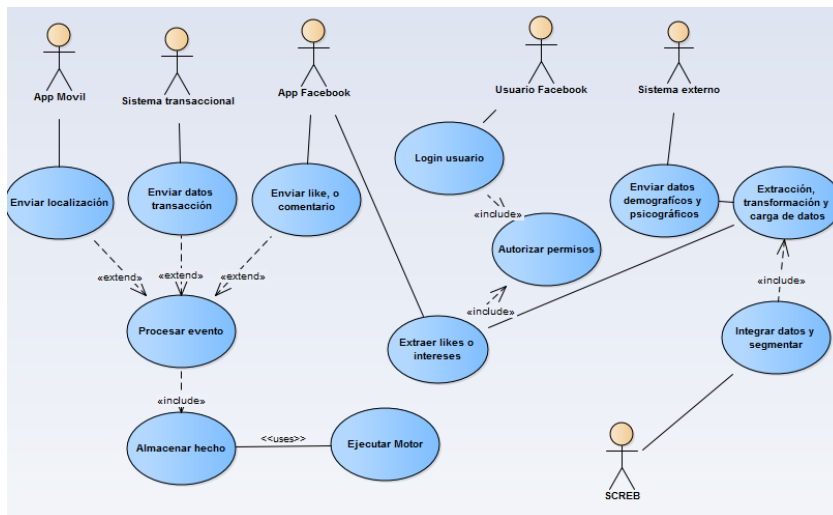


Figura 36. Diagrama caso de uso arquitectónico.

Identificador: CU_ENVIAR_LOCALIZACIÓN		Nombre: Enviar localización
Objetivo: Permite realizar envío de localización de usuario, en términos de longitud y latitud, cada 30 minutos.		
Complejidad: Media	Impacto: Alto	Documento asociado:
Actores: 1. Sistema (App Movil)		
Entrada: Localización de usuario.		Salida: Envío de Longitud, Latitud y email a servidor de aplicaciones.
Descripción: El sistema captura y realiza envío de localización e identificación de usuario, en términos de longitud, latitud y correo electrónico.		
Proceso: 1. Se autoriza acceso de la app a la ubicación del cliente mediante GPS. 2. Se habilita el GPS en el dispositivo móvil. 3. El sistema realiza captura de la ubicación del cliente. 4. El sistema realiza envío de la ubicación e identificación del cliente instalado.		
Precondiciones: 1. El usuario debe haber realizado instalación previa de la app móvil en el cliente.		
Manejo de situaciones anormales: 1. En el punto 2 se debe verificar que el dispositivo cuente con el GPS activado, y para aumentar el grado de confiabilidad respecto a la ubicación, la conexión Wifi.		

Identificador: CU_ENVIAR_TRANSACCIÓN		Nombre: Enviar datos transacción
Objetivo: Permite realizar envío de datos acerca de la transacción realizada por el usuario.		
Complejidad: Media	Impacto: Alto	Documento asociado: Documento de descripción transacciones.

Actores:	
1. Sistema.	
Entrada: Transacción.	Salida: Envío de datos acerca de la transacción
Descripción: El sistema captura y realiza envío de datos acerca de identificación de usuario y categoría de comercio, del sitio de donde se origina la transacción..	
Proceso:	
<ol style="list-style-type: none"> 1. El usuario realiza compra con tarjeta en establecimiento de comercio. 2. El sistema identifica una nueva transacción. 3. El sistema realiza captura de la transacción. 4. El sistema realiza envío de la identificación del usuario y la categoría del establecimiento de comercio de donde proviene la transacción. 	
Precondiciones:	
1. El usuario debe ser poseedor de tarjeta crédito o debito activa, y realizar compra en establecimiento de comercio, por medio de datafono Credibando.	
Manejo de situaciones anormales:	
1. En caso de presentarse un error de comunicación o error en la transacción, se debe iniciar nuevamente el proceso.	

Identificador: CU_ENVIAR_DATOS_SOCIAL_MEDIA		Nombre: Enviar like y/o comentario
Objetivo: Permite realizar envío de datos acerca de un nuevo like y/o comentario realizado por un usuario en la red social Facebook.		
Complejidad: Alta	Impacto: Alto	Documento asociado: API Facebook PHP SDK 4.
Actores:		
1. Sistema.		
Entrada: Like o comentario.		Salida: Envío de datos acerca del contenido del like y/o del comentario.
Descripción:		

El sistema captura y realiza envío de datos acerca de la categoría de la página en la cual el usuario ha realizado un nuevo <i>like</i> o el contenido de un nuevo comentario (estado) realizado en su perfil..
Proceso: <ol style="list-style-type: none"> 1. El usuario realiza un nuevo like o comentario en la red social Facebook. 2. El sistema identifica la existencia de un nuevo like o comentario. 3. El sistema realiza captura de la información referente al like o comentario. 4. El sistema realiza envío de la categoría de la página de donde proviene el like, y/o el contenido del texto del nuevo comentario realizado.
Precondiciones: <ol style="list-style-type: none"> 1. El usuario debe haber aceptado a los permisos acerca de 'email','user_likes','user_posts','user_status' y 'read_stream'.
Manejo de situaciones anormales: <ol style="list-style-type: none"> 1. En caso de que el usuario revoque los permisos en el sistema, estos serán solicitados nuevamente en un próximo ingreso.

Identificador: CU_PROCESAR_EVENTO_LOCALIZACION		Nombre: Procesar localización de usuario
Objetivo: Permite obtener las categorías de comercio cercanas a menos de 500mts, de la ubicación del usuario.		
Complejidad: Media	Impacto: Alto	Documento asociado: Base de datos de categorías y establecimientos de comercio.
Actores: 1. Sistema.		
Entrada: Longitud, Latitud, correo electrónico.		Salida: Categoría(s) de comercio de los establecimientos a menos de 500Mts de la ubicación dada.
Descripción: El sistema procesa información correspondiente a la ubicación del usuario en términos de longitud y latitud, y consulta categorías de comercio de establecimiento a menos de 500Mts de la ubicación dada.		
Proceso: 1. El sistema obtiene longitud y latitud del usuario.		

<ol style="list-style-type: none"> 2. El sistema consulta en la base de datos, los establecimientos de comercio a menos de 500MTs de la longitud y latitud dadas. 3. El sistema retorna la categoría(s) de comercio, de establecimientos cercanos. 4. Continúa el caso de uso Almacenar Hecho. 		
Precondiciones: <ol style="list-style-type: none"> 1. Se debe haber capturado la localización del usuario en términos de longitud y latitud. 2. Debe existir una base de datos de establecimientos y categorías de comercio. 		
Manejo de situaciones anormales: <ol style="list-style-type: none"> 1- En caso de no encontrar establecimientos de comercio cercanos, el sistema no continúa con el caso de uso Almacenar hecho. 		
Identificador: CU_PROCESAR_EVENTO_TRANSACCION		Nombre: Procesar transacción usuario
Objetivo: Permite obtener las categorías de comercio de donde se origina la transacción		
Complejidad: Media	Impacto: Alto	Documento asociado: Base de datos de categorías y establecimientos de comercio.
Actores: <ol style="list-style-type: none"> 1. Sistema. 		
Entrada: Transacción.		Salida: Categoría(s) de comercio de establecimientos de donde se origina la transacción e identificación de usuario que realiza la transacción.
Descripción: El sistema procesa información correspondiente a la transacción y obtiene la categoría de comercio de donde se origina la transacción como la identificación del usuario que la realiza.		
Proceso: <ol style="list-style-type: none"> 1. El sistema identifica una nueva transacción. 2. El sistema extrae la identificación de la transacción. 3. El sistema cruza la información extraída contra una tabla de propietarios de tarjeta. 4. El sistema obtiene la categoría de comercio de donde se origina la transacción a partir de los datos previamente extraídos. 		
Precondiciones: <ol style="list-style-type: none"> 1. Se debe realizar la transacción por medio de Datafono asociado a la empresa Credibanco. 		

Manejo de situaciones anormales:

1. En caso de error en la conexión, se debe finalizar la comunicación.

Identificador: CU_PROCESAR_EVENTO_SOCIAL_MEDIA		Nombre: Procesar like y/o comentario
Objetivo: Permite procesar un nuevo like y/o comentario realizado por un usuario en la red social Facebook, con la finalidad de identificar una intencionalidad de consumo en tiempo cercano al real.		
Complejidad: Alta	Impacto: Alto	Documento asociado: API Facebook PHP SDK 4 Base de datos de verbos y predicados.
Actores: 1. Sistema.		
Entrada: Like o comentario.		Salida: Intencionalidad de consumo manifestado en el like y/o comentario proveniente de la red social.
Descripción: El sistema realiza procesamiento de un nuevo like o comentario, proveniente de la red social Facebook, en caso de corresponder a un nuevo like, el sistema realiza extracción de la categoría de la página asociada a este, y en caso de provenir de un nuevo comentario realiza análisis de texto sobre su contenido.		
Proceso: <ol style="list-style-type: none"> 1. El sistema detecta un nuevo evento social proveniente del usuario en la red social Facebook. 2. El sistema identifica si el evento corresponde a un nuevo like o a un nuevo comentario. 3. En caso de provenir de un nuevo like, el sistema extrae la categoría de la página asociada a este. 4. En caso de provenir de un nuevo comentario el sistema analiza su contenido, identificando verbos y predicados. 5. El sistema consulta la base de datos de verbos y predicados, contra los incluidos en el comentario. 6. En caso de encontrarse el verbo y predicado en la base de datos, el sistema extrae el valor de cómo se encuentra estas en la definición en términos de la ontología. 7. El sistema realiza ejecución de consulta SPARQL sobre la ontología de dominio, tomando como parámetros los valores obtenidos del punto anterior. 		

8. El sistema obtiene, el nombre del segmento que clasifica al comentario, basado en la descripción de intereses y aptitudes de consumo descritos en la ontología de dominio para cada segmento de manera individual.		
Precondiciones:		
1. Se debe haber ejecutado de forma exitosa el caso de uso CU_ENVIAR_DATOS_SOCIAL_MEDIA .		
Manejo de situaciones anormales:		
<ol style="list-style-type: none"> 1. En caso de presentarse un comentario vacío, este será desestimado. 2. En caso de presentarse un comentario con igual contenido al realizado en un periodo menor a 10 minutos, igualmente será desestimado. 3. El sistema no toma en cuenta los likes y comentarios, anteriores a la fecha de suscripción del usuario. 		
Identificador: CU_ALMACENAR_HECHO	Nombre: Almacenar evento	
Objetivo: Permite almacenar el resultado del procesamiento de un evento en la base de hechos		
Complejidad: Media	Impacto: Alto	Documento asociado: Base de datos de hechos.
Actores:		
1. Sistema.		
Entrada: Resultado del evento procesado.	Salida: Evento almacenado a manera de hecho.	
Descripción:		
El sistema almacena el evento procesado, a manera de hecho, en la base de datos dispuesta para tal fin, que servirá como elemento de entrada al motor de reglas.		
Proceso:		
<ol style="list-style-type: none"> 1. El sistema obtiene un nuevo evento procesado. 2. El sistema verifica el tipo de evento procesado. 3. El sistema almacena el evento procesado en forma de hecho, tomando en cuenta si corresponde a un evento transaccional, de localización o social. 		
Precondiciones:		
1. Se debe haber ejecutado de forma exitosa un caso de uso de procesamiento de eventos..		

Identificador: CU_EJECUTAR_MOTOR		Nombre: Ejecutar motor de reglas
Objetivo: Ejecuta las reglas almacenadas en la base de datos de hechos.		
Complejidad: Media	Impacto: Alto	Documento asociado: Base de datos de hechos.
Actores: 1. Sistema.		
Entrada: Base de datos de hechos. Reglas		Salida: Envío de correo electrónico con promoción.
Descripción: El sistema ejecuta el proceso conocido como “ <i>Pattern matcher</i> ” en el cual cruza la información proveniente de la base de datos de hechos, contra la descrita en las reglas, y en caso de cumplirse alguna regla para los datos dados ejecuta su consecuencia, que se materializa en él envío de un correo electrónico al usuario final.		
Proceso: 1. El sistema consulta detecta un nuevo evento procesado. 2. El sistema carga la base de datos de hechos. 3. El sistema carga las reglas ingresadas al motor. 4. El sistema ejecuta el algoritmo “ <i>Pattern matcher</i> ” 5. El sistema ejecuta las consecuencias de cada regla, cumplida. 6. El sistema queda en ejecución y a la espera de un nuevo evento.		
Precondiciones: 1. Debe existir un procesamiento previo de eventos tipo: transacción, ubicación o social. 2. Se deben contar con hechos almacenados previamente en la base de datos.		
Manejo de situaciones anormales: 1. En caso de presentarse un error de comunicación, este será procesado posteriormente. 2. En caso de presentarse eventos de manera simultánea serán almacenados en una cola de mensajes, para su posterior ejecución. 3. En caso de que no se cumpla ninguna regla, para la base de hechos cargada, no se realizara envío de correo electrónico.		

2. Diagrama de actividades

En los siguientes diagramas se describe el flujo de trabajo entre el usuario y el sistema, el proceso de software y las diferentes acciones realizadas para la ejecución de los principales casos de uso del sistema.

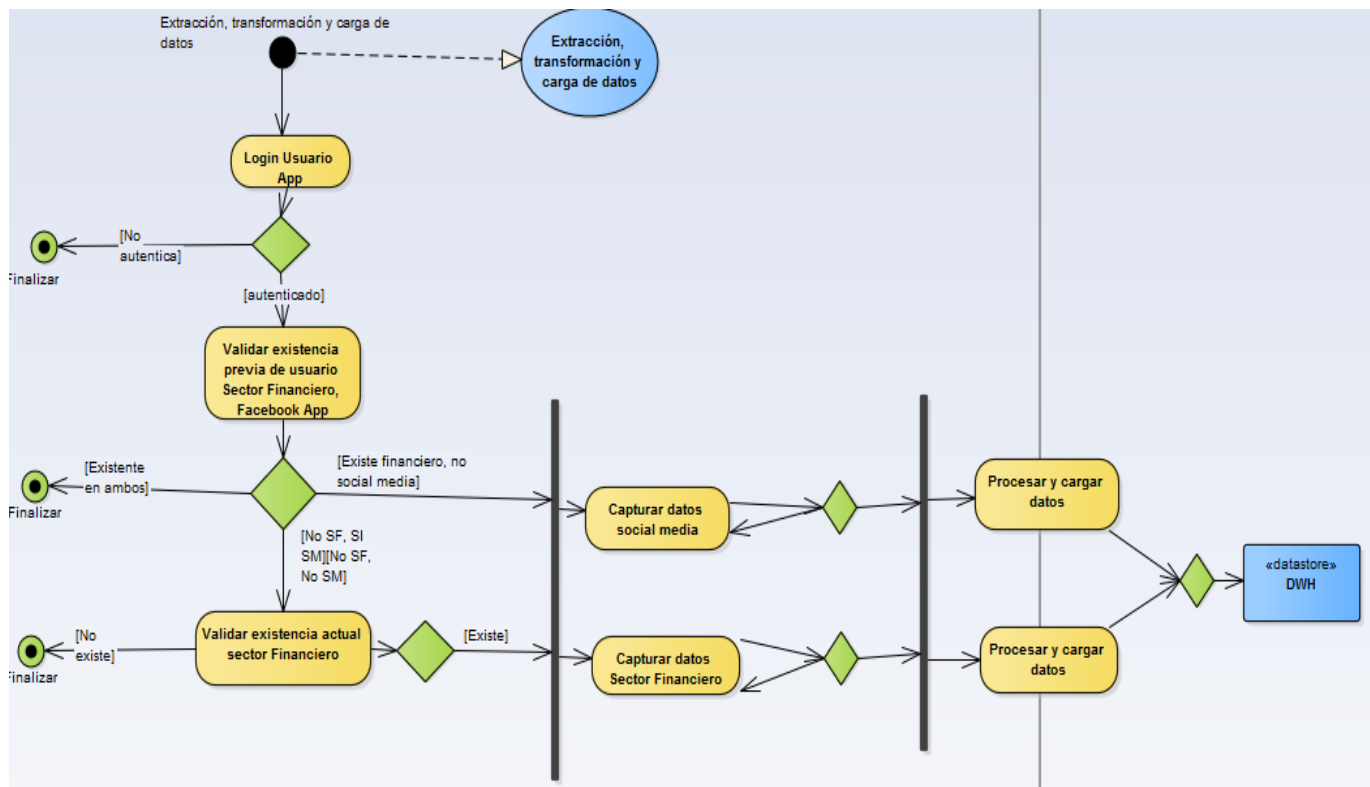


Figura 37. Diagrama de actividades Extracción transformación y carga de datos.

En el presente diagrama se describe el flujo de actividades desarrolladas para cumplir con lo dispuesto en el caso de uso de extracción, transformación y carga de datos **CU_ETL**.

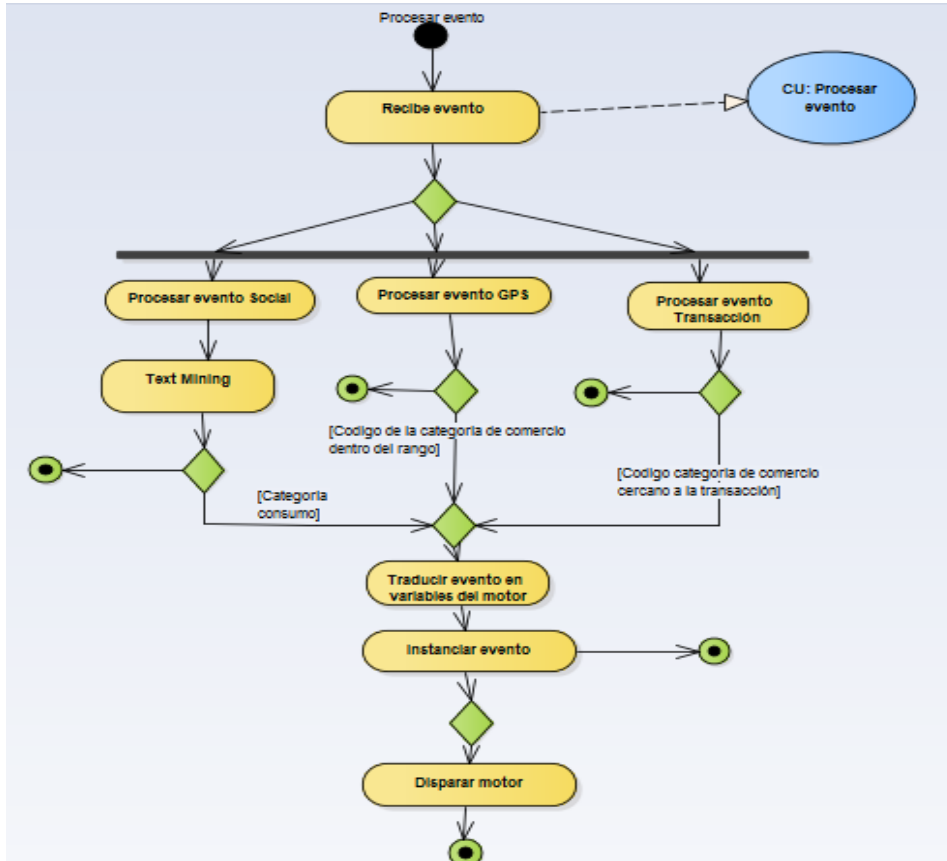


Figura 38. Diagrama de actividades CU procesar evento.

En el diagrama anterior se describen cada una de las actividades realizadas, durante la ejecución del caso de uso de procesamiento de eventos. En particular, aquellas de tipo transaccional, de ubicación y sociales, cuyo resultado es el almacenamiento del evento procesado y traducido en variables del motor de reglas.

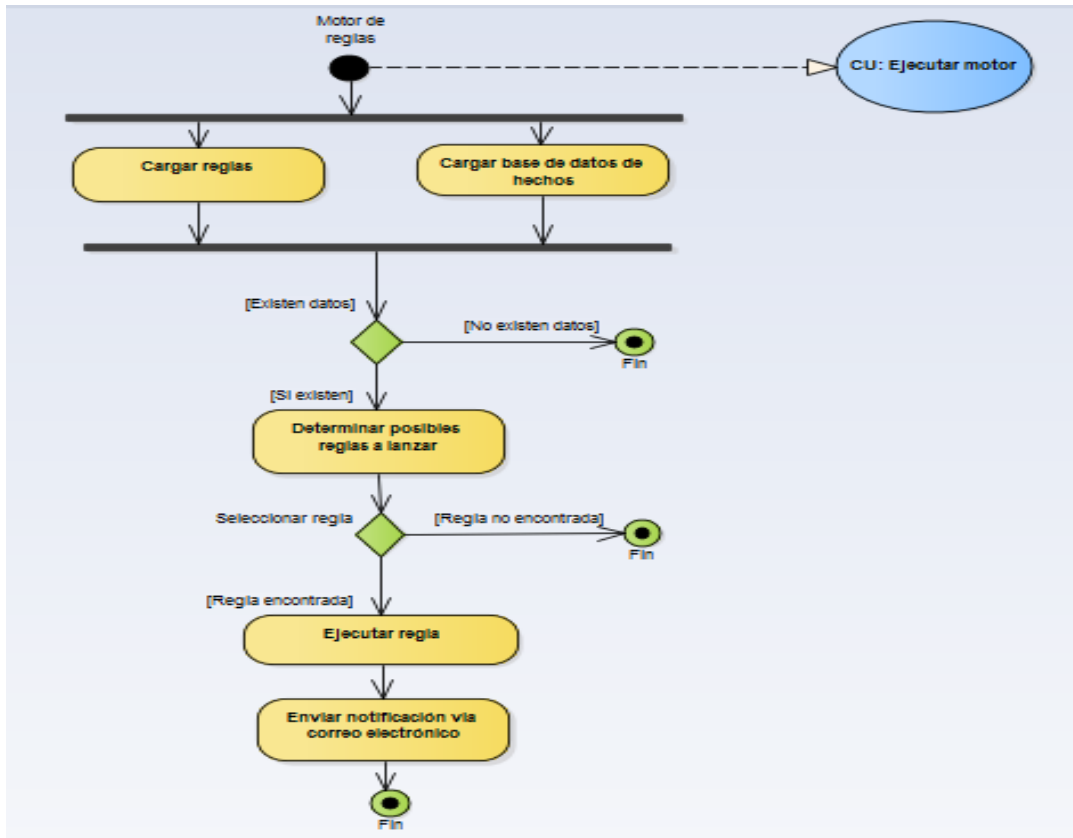


Figura 39. Diagrama de actividades CU ejecutar motor.

En el presente diagrama se describe la secuencia de actividades realizadas en el caso de uso de ejecución del motor de reglas. Cuyo resultado final, es el envío de una notificación vía correo electrónico al usuario final.

3. Diagrama de secuencia Facebook

El siguiente diagrama describe la interacción existente entre: el sistema de extracción de información social, el usuario, el server huésped de la aplicación y finalmente, el servidor de Facebook que procesa las peticiones enviadas por el cliente.

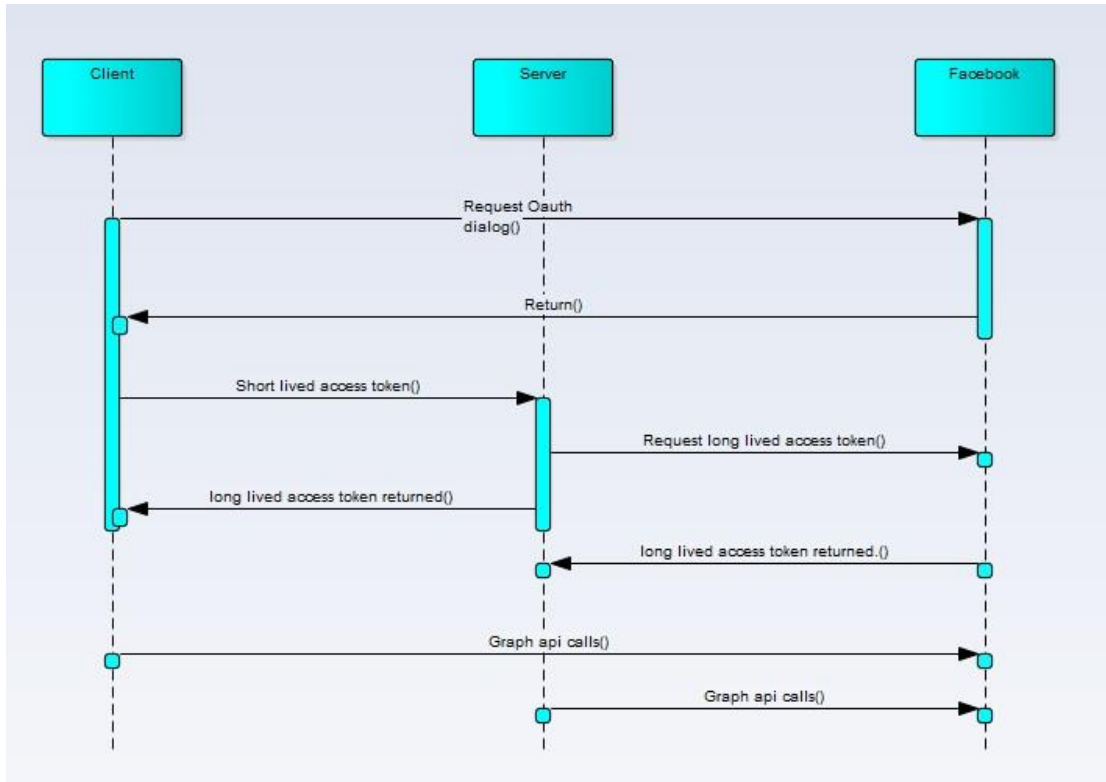


Figura. 42. Diagrama de secuencia app Facebook.

4. Diagrama de despliegue

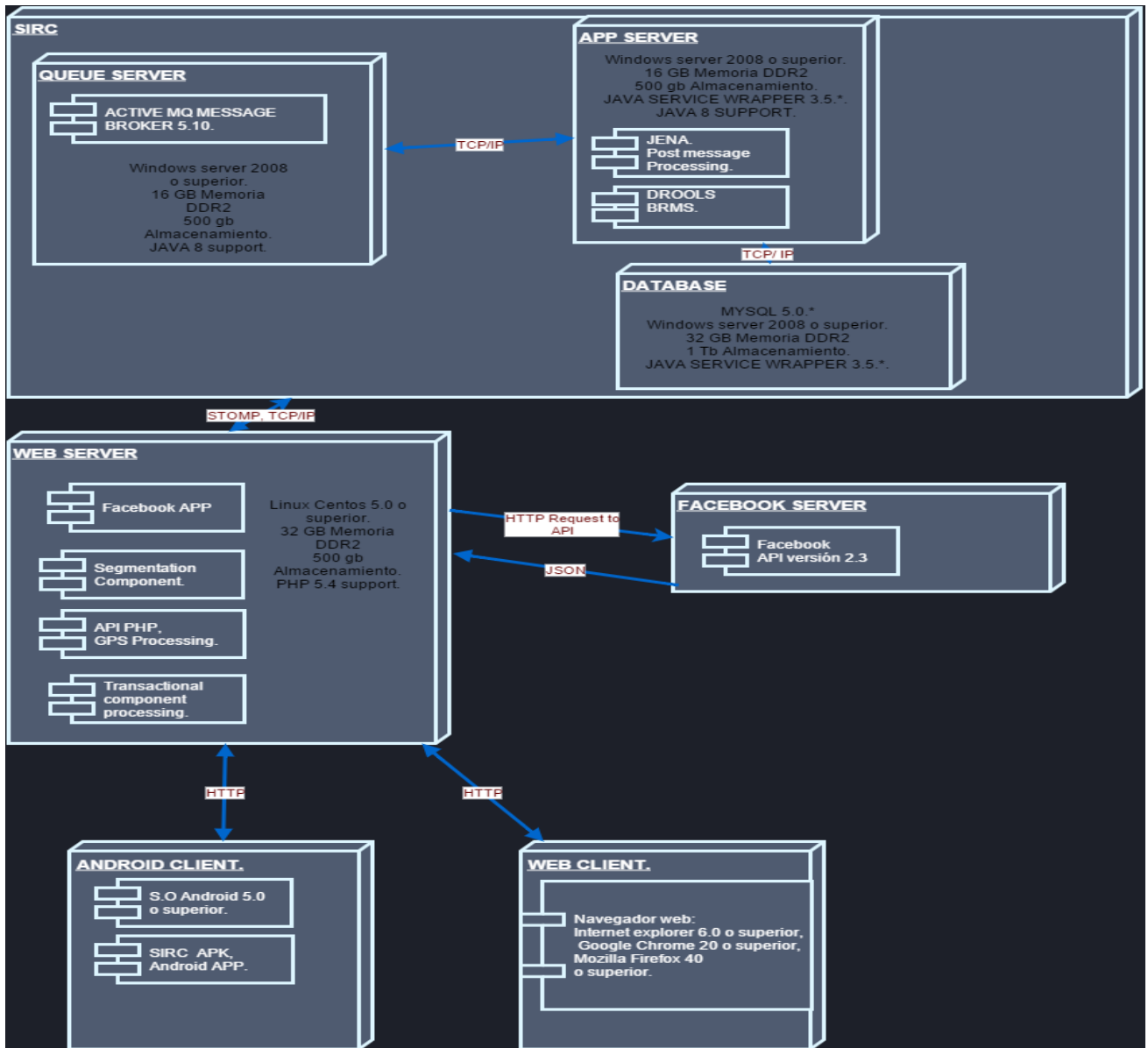


Figura. 43. Diagrama de despliegue del sistema.

5. Diagrama de componentes

El siguiente diagrama se muestra los componentes de la arquitectura propuesta para la solución desarrollada.

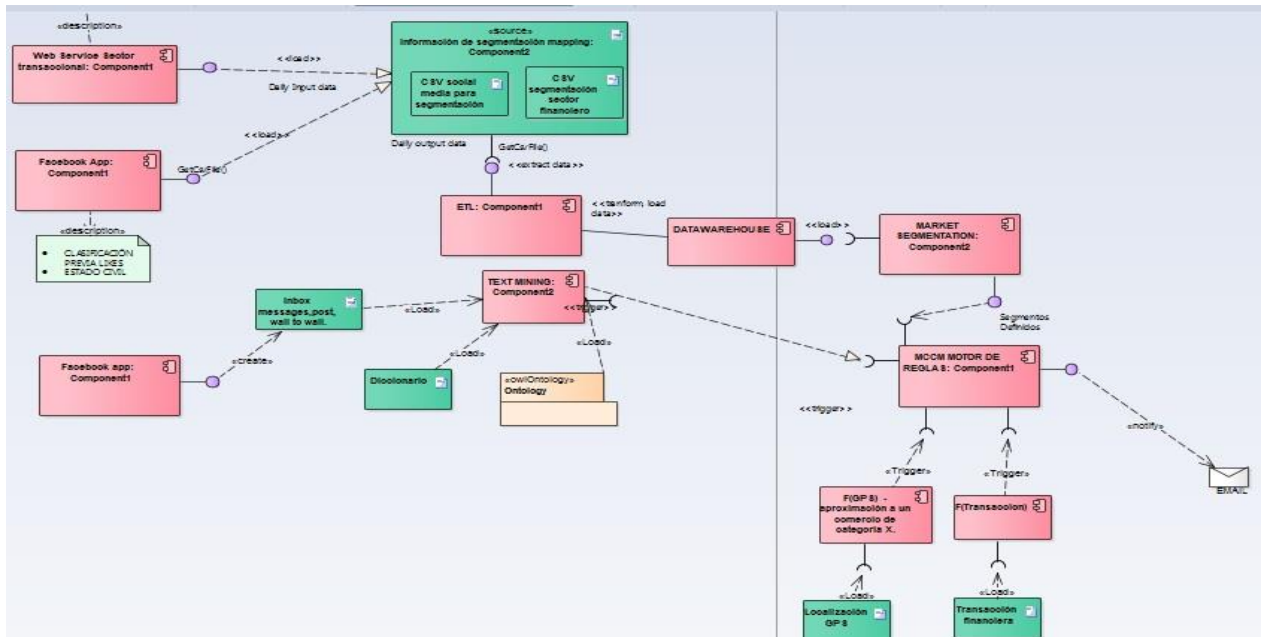


Figura 43. Diagrama de componentes del sistema.

El diagrama representa cada uno de los elementos que hacen parte del sistema, junto con sus componentes e integraciones.

Entre los principales componentes se encuentran:

Un componente destinado al proceso de segmentación, que recibe datos demográficos provenientes del sector financiero y datos psicográficos correspondientes a intereses extraídos de social media. Así mismo, realiza el proceso de segmentación del usuario basado en los segmentos identificados en el [capítulo IV \(subsección 1.2\)](#) del presente documento.

Un componente destinado al proceso de análisis de texto a partir de los comentarios provenientes de la red social Facebook. Explicado a profundidad en el [capítulo IV \(subsección 2.2\)](#).

Un componente de procesamiento de eventos de localización, que basado en datos provenientes del celular del usuario en términos de latitud y longitud, obtiene los establecimientos de comercio en un rango de 500Mts, de la ubicación dada, el cual fue explicado a profundidad en el [capítulo IV, subsección 2.1](#).

Un componente de procesamiento de eventos de transaccionales, que extrae la categoría de comercio de donde proviene dicha transacción, el cual fue explicado a profundidad en el [capítulo IV, subsección 2.3](#).

Por último, un componente central encargado de ejecutar las reglas de negocio. De modo que, a partir de cada uno de los hechos almacenados, infiere la promoción aplicable para cada usuario. Finalmente plasmada a través de un correo electrónico ([ver capítulo IV, sección 3](#)).

VI – CONCLUSIONES Y TRABAJO FUTURO

Concluyó a partir de la consolidación de los estudios revisados, que gran parte de las deficiencias encontradas en los sistemas de envío de campañas masivas, corresponden a no ser llevadas a un mayor grado de individualización, lo que ocasiona entre otras cosas la generación de campañas que no se adaptan a los intereses, características y necesidades de los clientes. Por lo cual, se creó un modelo inteligente de recomendación de campañas basado en perfilamiento de hábitos de consumo que toma en cuenta aspectos estáticos y dinámicos del cliente, para así mejorar los niveles de personalización y de generación de aprendizaje.

A fin de construir el modelo de perfilamiento se inició con una etapa previa de investigación, en la cual se identificaron segmentos de mercado de consumo en Colombia, descritos en términos de variables demográficas (edad, género, estrato, nivel de ingresos) y psicográficas (intereses). Con el objetivo de ser implementados como punto de partida para el diseño y construcción del componente estático del modelo, correspondiente a clasificación de usuarios. Además de interpretar su propensión a consumo en términos de reglas, construidas junto con los resultados de los demás componentes del modelo. Con esto se lograron los dos primeros objetivos del proyecto.

En segunda instancia se diseñó, el componente dinámico del modelo compuesto por eventos provenientes del mundo físico, y virtual. Entre los eventos del mundo físico encontramos: eventos de ubicación geográfica, y de transacciones financieras. Por otra parte, entre los eventos del mundo virtual tenemos aquellos que provienen de actividad en redes sociales como: “me gusta” o “likes” y comentarios. Además, se seleccionó un motor de reglas como componente central del modelo, el cual requería a partir de las salidas dadas por el componente estático del modelo asociado a los segmentos establecidos en las investigaciones previas, el componente dinámico del modelo relacionado a la intencionalidad de compra expresada y a la propensión de consumo de dichos segmentos, deducir la campaña mediante la evaluación lógica de los anteriores resultados en tiempo cercano al real. Con lo cual se logró el tercer objetivo del proyecto.

Posteriormente, y con el objetivo de validar el modelo propuesto, se construyó un prototipo funcional. Dicho prototipo desarrolla el componente estático del modelo a partir de la segmentación de usuarios, por medio de datos ficticios provenientes del sector financiero. Luego, se

procedió con la construcción de los componentes dinámicos del modelo, a partir de los siguientes componentes: componente de procesamiento de eventos transaccionales, componente de procesamiento de eventos de ubicación geográfica, y por último un componente de procesamiento de actividad en redes sociales. Finalmente, los resultados del procesamiento de datos por parte de los diferentes componentes son almacenados en la base de datos de hechos, las cuales son instanciadas por el componente central del sistema, desarrollado a partir del motor de reglas de negocio Drools. Con esto, se logró el quinto objetivo del proyecto.

La principal contribución del proyecto, es el modelo inteligente de recomendación de campañas, el cual incluye aspectos estáticos y dinámicos, reaccionando en tiempo real frente a intencionalidades de consumo manifestada por los clientes, desde diversas fuentes de datos como la encontrada en redes sociales, ubicación geográfica y transacciones financieras. Logrando obtener recomendación de campañas masivas de manera individualizada acorde con las nuevos intereses y necesidades expresadas por los usuarios.

En cuanto a las lecciones aprendidas, se identificó la importancia de reaccionar en tiempo real a la intencionalidad de consumo manifestadas desde diferentes fuentes, de tal forma que puedan ser materializadas en forma de campañas promocionales, las cuales pueden alcanzar a tener un efecto positivo mayor a las enviadas tiempo después de que la intención de consumo fue manifestada. De igual forma, vale la pena recalcar la tendencia global entorno al posicionamiento de productos y servicios de acuerdo a las preferencias actuales del cliente, en un lugar y tiempo específico, lo que se resume en campañas instantáneas y personalizadas. El siguiente ejemplo ilustra mejor, lo anteriormente expuesto: Durante el Superbowl 2013, al presentarse una falla en el suministro de energía, la compañía Oreo aprovechó para realizar una microcampaña, basada en un tweet “¿Sin energía?, No hay problema”. Mostrando una galleta solitaria con el slogan “aún la puedes mojar en la oscuridad”, haciendo referencia al hábito de comer estas galletas con leche, campaña que logró volverse viral en redes sociales. Ahora bien, al diseñar campañas significativas para el consumidor, en cuanto interpretan sus propios gustos y preferencias, mediante la reacción a eventos que suceden en su vida cotidiana, ya sea en su mundo físico o virtual, los resultados en la efectividad de campañas de marketing seguramente se incrementarían de manera destacable. Sin embargo, este tipo de campañas pueden verse limitadas por las restricciones a nivel de privacidad tanto de plataformas como de usuarios. En el caso de nuestro proyecto, nos vimos limitados a las condiciones de acceso a la información dispuestos por Facebook, el cual no permite acceder a información de comentarios en tiempo real para realizar tareas relacionadas a análisis de texto, por lo cual el prototipo fue validado solo en las cuentas previamente inscritas al interior de la aplicación en esta red social.

Como trabajo futuro, se plantea ampliar el modelo con la inclusión de otros tipos de eventos tanto del mundo físico, como virtual y la alimentación con otras fuentes de datos, que permitan realizar un análisis más exhaustivo de intereses y preferencias de los usuarios. Igualmente la inserción dentro del modelo de un componente de generación de reglas automáticas, que tomen en cuenta eventos y aspectos globales del entorno para su generación, como los ilustrados en la campaña de galletas Oreo. También, el desarrollo de una red de ontologías propias del consumo y darle capacidad de aprendizaje al motor, reteniendo hechos, deducidos a partir de las mismas reglas. Por último se plantea, la inclusión futura de un componente de análisis de sentimientos, para mejorar la eficacia en la identificación de intencionalidades de consumo manifestadas por medio de comentarios en redes sociales.

REFERENCIAS

- [1] Loveman, G. W., «Employee satisfaction, customer loyalty, and financial performance an empirical examination of the service profit chain in retail banking». *Journal of Service Research*, 1998, 1(1), 18-31.
- [2] D. Loshin, “Using Information to Develop a Culture of Customer Centricity”, Chapter 8 *Customer Profiling*, 2013.
- [3] P. A. Phillips, Davies, F. M., & Moutinho, L., “The interactive effects of strategic marketing planning and performance: A neural network analysis. *Journal of Marketing Management*”, 2001, 17(1-2), 159-182.
- [4] Blog, [en línea] <http://blogs.gartner.com/adam-sarner/2015/04/07/multichannel-campaign-management-in-2015/>.
- [5] H. Interactive, “Online survey of 4,271 checking account customers”, 2012.
- [6] Primary Bank, “Definition. Online”, [en línea] “<http://www.investopedia.com/>”.
- [7] R. Hallowell, “The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International journal of service industry management*”, 1996, 7(4), 27-42.
- [8] D. Loshin, “Using Information to Develop a Culture of Customer Centricity - Chapter 8 *Customer Profiling*. “, 2013.
- [9] K. Maru, R. Prince, “A psychographic segmentation of industrial family businesses Original Research Article”, *Industrial Marketing Management*, 1996, Volume 25, Issue 3, Pg 223-234.
- [10] D. Loshin, “Using Information to Develop a Culture of Customer Centricity” ,2013.
- [11] Lisa Artur, “Are You Using Multi-Channel Campaign Management to Lead? “, *Forbes*, 2011.
- [12] Deloitte, “Taking retail bank cross-selling to the next level”, 2013, Pg. 1 -3.
- [13] J. Facchin , “Las Redes Sociales como motor para el Up-Selling y Cross-Selling”, 2013.
- [14] W.K. Chan, Yuen Yau Chiu, Yuen Tak Yu, “A web search-centric approach to recommender systems with URLs as minimal user contexts”, Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong, China.
- [15] Finsights, “Analytics cross-selling-retail-banking”, 2010.
- [16] P. Russom, “Big Data Analytics Introduction”, TWDI RESEARCH. 2011.
- [17] Y. Park, K. Nyeong , “Individual and group behavior-based customer profile model”.
- [18] J. Mullich, “4 Approaches for Integrating Social Media Data with Enterprise Systems”, [en línea] <http://data-informed.com/4-approaches-for-integrating-social-media-data-with-enterprise-systems/#sthash.ZV6BqfP5.dpuf>, 2013.
- [19] K. Trainora, J. Andzulisb, A. Rapp, “Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM”.
- [20] M. Cross, “Chapter 2: Opportunities of Social Media, Social Media Security Leveraging Social Networking While Mitigating Risk”, 2013.
- [21] D Loshin, “Using Information to Develop a Culture of Customer Centricity - Chapter 1 What is Customer centricity”, 2013.
- [22] M. Xu, J. Walton. "Gaining customer knowledge through analytical CRM." *Industrial management & data systems*, 105.7 (2005): 955-971.

- [23] S.N.H Jansen, “Customer Segmentation and Customer Profiling for a Mobile Tele-communications Company Based on Usage Behavior”, 2007.
- [24] F. Hamka, H. Bouwman, M. de Reuver a, M. Kroesen, “Mobile customer segmentation based on smartphone measurement”, 2014.
- [25] J. van Dam, M. van de Velden, “Online profiling and clustering of Facebook users”, 2014.
- [26] Yanhaas, “Redes Sociales: Una nueva forma de vincularse con el consumidor”, Yanhaas Poll. 2011, [en línea] <http://www.yanhaas.com/index.php/yanhaas-poll/item/241-yanhaas-poll-no-76>.
- [27] N. Robehmed, “What is a Startup?”, Forbes, 2013.
- [28] D. Loshin , “ Using Information to Develop a Culture of Customer Centricity - Chapter 9 Customer Data Analytics”, 2013.
- [29] Y. Park, K. Chang, “Individual and group behavior-based customer profile model”.
- [30] W. Ribarsky, D. Xiaoyu, W. Dou, “Social media analytics for competitive advantage Original Research Article Computers & Graphics”, Volume 38, February 2014, Pages 328-331
- [31] T. Morgan. “Business Rules and Information Systems. Boston: Addison-Wesley Publishing”, 2002.
- [32] R. G. Ross, “Principles of the Business Rule Approach”, Addison-Wesley, 2003.
- [33] Drools, [en línea] https://docs.jboss.org/drools/release/6.1.0.Final/drools-docs/html_single/#HybridReasoningChapter.
- [34] A. Y.K Chua, S. Banerjee, “Customer knowledge management via social media: the case of Starbucks. Journal of Knowledge Management”, 2013 17:2, 237-249.
- [35] Wu He, Z. Shenghua, L. Ling, “Social media competitive analysis and text mining: A case study in the pizza industry”.
- [36] D. Loshin, “Using Information to Develop a Culture of Customer Centricity”, 2013.
- [37] D. Lopez, E. Caldón, “Enriquecimiento Semántico de Contenidos en Redes Sociales basado en Ontologías y Vocabularios de Dominio”, [en Línea] <http://seer.ufrgs.br/index.php/cadernosdeinformatica/article/view/v4n2p19-26/8728>.
- [38] J. Tramullas Saz, P. Garrido, “Organización de información en herramientas WIKI: aplicación de ontologías en wikis semánticos”.
- [39] P. Mika, “Ontologies are us: A unified model of social networks and semantics”, Department of Computer Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands, [en línea], <http://www.websemanticsjournal.org/index.php/ps/article/view-File/106/104>.
- [40] M. Hamasaki, Y. Matsuo, T. Nishimura, H. Takeda, “Ontology Extraction using Social Network”, [en línea] <https://staff.aist.go.jp/masahiro.hamasaki/pub/swecka07.pdf>.
- [41] G. Erétéo, M Buffa, F Gandon, MLeitzelman, FLimpens, “Leveraging Social data with Semantics”, [en línea] http://www.w3.org/2008/09/msnws/papers/ereteo_et_al_2008_leveraging.html.
- [42] W. Chen and S. Fong, “Social Network Collaborative Filtering Framework and Online Trust Factors: a Case on Facebook”. IEEE, 2010, 266-273.
- [43] P. Kazienko and K. Musial, “Recommendation Framework for Online Social Networks”, Advances in Web Intelligence and Data Mining, Springer, Vol. 23, 2006, 111-120.
- [44] A. Ortigosa, J. Martin, R. Carro, “SENTIMENT ANALYSIS IN FACEBOOK AND ITS APPLICATION TO E-LEARNING”.

- [45] A. Balahur, J. M Perea, Ortega, “Sentiment analysis system adaption for multilingual processing: The case of tweets”.
- [46] P. D. Turney, “Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews”.
- [47] “Recommending Venues Using Continous Predictive Social Media Analytics”, 2013.
- [48] “Tracking Movemets attention of crowds in Real Time Analysing Social Streams”, 2012.
- [49] Z. Suna, L.Hana, W. Huanga, X. Wanga, X. Zenga, M. Wanga, H. Yanb, “Recommender systems based on social networks”, 2014.
- [50] M. Mostafa. “More than words: Social networks’ text mining for consumer brand sentiments”.
- [51] Yanhass, “Mujeres y finanzas: Una combinación interesante”, Yanhass Poll, 2010, [en línea] <http://www.yanhaas.com/index.php/yanhaas-poll/item/236-yanhaas-poll-no-71>.
- [52] Yanhass, “La Super Mujer”, Yanhass Poll, 2009, [en línea] <http://www.yanhaas.com/index.php/yanhaas-poll/item/231-yanhaas-poll-no-65>.
- [53] Yanhass, “¿Cómo innovar en el mercado colombiano?”, Yanhass Poll, 2006, [en línea] http://www.yanhaas.com/index.php/yanhaas-poll/item/download/53_73b5ba581272fe8f1977339bf17e6a6b.
- [54] Redacción Portafolio, “Consumidor colombiano”, Portafolio.co.
- [55] Dane, “IPC DANE”, 2013.
- [56] Yanhass, “¿Por qué pagaría más el colombiano?”, Yanhass Poll, [en línea] <http://www.yanhaas.com/index.php/yanhaas-poll/item/211-yanhaas-poll-no-43>.
- [57] Yanhass, “Redes Sociales: Una nueva forma de vincularse con el consumidor”, Yanhass Poll, 201, [en línea] http://www.yanhaas.com/index.php/yanhaas-poll/item/download/93_6f61f749bd1eee89738e0fed7ab1a511.
- [58] Yanhass, “Saludable”, Yanhass Poll, 2009, [en línea] http://www.yanhaas.com/index.php/yanhaas-poll/item/download/85_d42e81ff5162c18ffe2a29198e3d0424.
- [59] Yanhass, “Solos”, Yanhass Poll, 2009, [en línea] http://www.yanhaas.com/index.php/yanhaas-poll/item/download/85_d42e81ff5162c18ffe2a29198e3d0424.
- [60] J. Hee, S. Chang, “Intelligent profitable customers segmentation system based on business intelli-gence tools”, 2005.
- [61] Yanhass, “Hoy más colombianos comen fuera de casa”, Yanhaas Poll, [en línea] <http://www.yanhaas.com/index.php/yanhaas-poll/item/242-yanhaas-poll-no-77>.
- [62] Yanhass , Finanzas del consumidor de bajos ingresos. Yanhass Poll, 2005, [en línea] <http://www.yanhaas.com/index.php/yanhaas-poll/item/201-yanhaas-poll-no-33>.
- [63] Accenture, “Impact of Online Digital Communication on Customer Buying Decision”, Accen-ture, 2012.
- [64] Redacción el tiempo, “Los 10 perfiles de los consumidores Bogotanos”, El tiempo. 2009.
- [65] W3C, “SPARQL Query Language for RDF”, [en línea] <http://www.w3.org/TR/rdf-sparql-query/>.
- [66] F. Tim, et al. "Social networking on the semantic web." *The Learning Organization* 12.5 (2005): 418-435.
- [67] Breslin, John G., et al. "SIOC: an approach to connect web-based communities." *International Journal of Web Based Communities* 2.2 (2006): 133-142.
- [68] H. Martin. "Goodrelations: An ontology for describing products and services offers on the web." *Knowledge Engineering: Practice and Patterns*. Springer Berlin Heidelberg, 2008. 329-346.

- [69] LIRMM ontologies publishing platform, Food Ontology, Web: [http:// data.lirmm.fr/ontologies/food#](http://data.lirmm.fr/ontologies/food#)
- [70] M. Suárez, “NeOn Methodology for building ontology networks: specification, scheduling and reuse”, Diss. Informatica, 2010.
- [71] Z. Anneke, M. Janssen, and K. Jeffery. "Towards an e-infrastructure to support the provision and use of open data." Conference for E-Democracy and Open Government. 2013.
- [72] Jena Apache. [en línea], <https://jena.apache.org/>.

ANEXO 2

CARTA DE AUTORIZACIÓN DE LOS AUTORES (Licencia de uso)

Bogotá, D.C., 15 de Diciembre del 2015

Señores
Biblioteca Alfonso Borrero Cabal S.J.
Pontificia Universidad Javeriana
Ciudad

Los suscritos:

William Hernán Suárez Molano, con C.C. No 1049614414
_____, con C.C. No _____
_____, con C.C. No _____

En mi (nuestra) calidad de autor (es) exclusivo (s) de la obra titulada:

Modelo Inteligente de Recomendación de Productos, Basado en Perfilamiento de Hábitos de Consumo.

(por favor señale con una "x" las opciones que apliquen)
Tesis doctoral Trabajo de grado Premio o distinción: **Si** **No**
cuál: _____

presentado y aprobado en el año 2015, por medio del presente escrito autorizo (autorizamos) a la Pontificia Universidad Javeriana para que, en desarrollo de la presente licencia de uso parcial, pueda ejercer sobre mi (nuestra) obra las atribuciones que se indican a continuación, teniendo en cuenta que en cualquier caso, la finalidad perseguida será facilitar, difundir y promover el aprendizaje, la enseñanza y la investigación.

En consecuencia, las atribuciones de usos temporales y parciales que por virtud de la presente licencia se autorizan a la Pontificia Universidad Javeriana, a los usuarios de la Biblioteca Alfonso Borrero Cabal S.J., así como a los usuarios de las redes, bases de datos y demás sitios web con los que la Universidad tenga perfeccionado un convenio, son:

AUTORIZO (AUTORIZAMOS)	SI	NO
1. La conservación de los ejemplares necesarios en la sala de tesis y trabajos de grado de la Biblioteca.	x	
2. La consulta física (sólo en las instalaciones de la Biblioteca)	x	
3. La consulta electrónica – online (a través del catálogo Biblos y el Repositorio Institucional)	X	
4. La reproducción por cualquier formato conocido o por conocer	x	
5. La comunicación pública por cualquier procedimiento o medio físico o electrónico, así como su puesta a disposición en Internet	X	
6. La inclusión en bases de datos y en sitios web sean éstos onerosos o gratuitos, existiendo con ellos previo convenio perfeccionado con la Pontificia Universidad Javeriana para efectos de satisfacer los fines previstos. En este evento, tales sitios y sus usuarios tendrán las mismas facultades que las aquí concedidas con las mismas limitaciones y condiciones	X	

De acuerdo con la naturaleza del uso concedido, la presente licencia parcial se otorga a título gratuito por el máximo tiempo legal colombiano, con el propósito de que en dicho lapso mi (nuestra) obra sea explotada en las condiciones aquí estipuladas y para los fines indicados, respetando siempre la titularidad de los derechos patrimoniales y morales correspondientes, de acuerdo con los usos honrados, de manera proporcional y justificada a la finalidad perseguida, sin ánimo de lucro ni de comercialización.

De manera complementaria, garantizo (garantizamos) en mi (nuestra) calidad de estudiante (s) y por ende autor (es) exclusivo (s), que la Tesis o Trabajo de Grado en cuestión, es producto de mi (nuestra) plena autoría, de mi (nuestro) esfuerzo personal intelectual, como consecuencia de mi (nuestra) creación original particular y, por tanto, soy (somos) el (los) único (s) titular (es) de la misma. Además, aseguro (aseguramos) que no contiene citas, ni transcripciones de otras obras protegidas, por fuera de los límites autorizados por la ley, según los usos honrados, y en proporción a los fines previstos; ni tampoco contempla declaraciones difamatorias contra terceros; respetando el derecho a la imagen, intimidad, buen nombre y demás derechos constitucionales. Adicionalmente, manifiesto (manifestamos) que no se incluyeron expresiones contrarias al orden público ni a las buenas costumbres. En consecuencia, la responsabilidad directa en la elaboración, presentación, investigación y, en general, contenidos de la Tesis o Trabajo de Grado es de mí (nuestro) competencia exclusiva, eximiendo de toda responsabilidad a la Pontificia Universidad Javeriana por tales aspectos.


Sin perjuicio de los usos y atribuciones otorgadas en virtud de este documento, continuaré (continuaremos) conservando los correspondientes derechos patrimoniales sin modificación o restricción alguna, puesto que de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación de los derechos patrimoniales derivados del régimen del Derecho de Autor.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables. En consecuencia, la Pontificia Universidad Javeriana está en la obligación de RESPETARLOS Y HACERLOS RESPETAR, para lo cual tomará las medidas correspondientes para garantizar su observancia.

NOTA: Información Confidencial:

Esta Tesis o Trabajo de Grado contiene información privilegiada, estratégica, secreta, confidencial y demás similar, o hace parte de una investigación que se adelanta y cuyos resultados finales no se han publicado. Si No

En caso afirmativo expresamente indicaré (indicaremos), en carta adjunta, tal situación con el fin de que se mantenga la restricción de acceso.

NOMBRE COMPLETO	No. del documento de identidad	FIRMA
William Hernán Suárez Yolaro	1049614414	

FACULTAD: INGENIERÍA DE SISTEMAS.

PROGRAMA ACADÉMICO: MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN.

ANEXO 3
BIBLIOTECA ALFONSO BORRERO CABAL, S.J.
DESCRIPCIÓN DE LA TESIS O DEL TRABAJO DE GRADO
FORMULARIO

TÍTULO COMPLETO DE LA TESIS DOCTORAL O TRABAJO DE GRADO						
Modelo Inteligente de Recomendación de Productos, Basado en Perfilamiento de Hábitos de Consumo.						
SUBTÍTULO, SI LO TIENE						
AUTOR O AUTORES						
Apellidos Completos			Nombres Completos			
SUÁREZ MOLANO			WILLIAM HERNÁN			
DIRECTOR (ES) TESIS O DEL TRABAJO DE GRADO						
Apellidos Completos			Nombres Completos			
SUÁREZ SANTAMARIA			LUIS IGNACIO			
FACULTAD						
INGENIERÍA DE SISTEMAS Y COMPUTACIÓN						
PROGRAMA ACADÉMICO						
Tipo de programa (seleccione con "x")						
Pregrado	Especialización	Maestría	Doctorado			
		X				
Nombre del programa académico						
MAESTRÍA EN INGENIERIA DE SISTEMAS Y COMPUTACIÓN						
Nombres y apellidos del director del programa académico						
ANGELA CRISTINA CARRILLO RAMOS						
TRABAJO PARA OPTAR AL TÍTULO DE:						
Magister en Ingeniería de Sistemas y Computación.						
PREMIO O DISTINCIÓN <i>(En caso de ser LAUREADAS o tener una mención especial):</i>						
CIUDAD		AÑO DE PRESENTACIÓN DE LA TESIS O DEL TRABAJO DE GRADO			NÚMERO DE PÁGINAS	
BOGOTÁ		2015			60	
TIPO DE ILUSTRACIONES (seleccione con "x")						
Dibujos	Pinturas	Tablas, gráficos y diagramas	Planos	Mapas	Fotografías	Partituras
		X			X	
SOFTWARE REQUERIDO O ESPECIALIZADO PARA LA LECTURA DEL DOCUMENTO						
Nota: En caso de que el software (programa especializado requerido) no se encuentre licenciado por la Universidad a través de la Biblioteca (previa consulta al estudiante), el texto de la Tesis o Trabajo de Grado quedará solamente en formato PDF.						

MATERIAL ACOMPAÑANTE					
TIPO	DURACIÓN (minutos)	CANTIDAD	FORMATO		
			CD	DVD	Otro: ¿Cuál?
Vídeo					
Audio					
Multimedia					
Producción electrónica					
Otro: ¿Cuál?					
<p align="center">DESCRIPTORES O PALABRAS CLAVE EN ESPAÑOL E INGLÉS</p> <p>Son los términos que definen los temas que identifican el contenido. <i>(En caso de duda para designar estos descriptores, se recomienda consultar con la Sección de Desarrollo de Colecciones de la Biblioteca Alfonso Borrero Cabal S.J en el correo biblioteca@javeriana.edu.co, donde se les orientará).</i></p>					
ESPAÑOL			INGLÉS		
sistemas de recomendación			recommender systems		
perfilamiento			profiling		
campañas masivas			massive campaigns		
social media analisis			social media analysis		
<p align="center">RESUMEN DEL CONTENIDO EN ESPAÑOL E INGLÉS (Máximo 250 palabras - 1530 caracteres)</p>					
<p>Segmentation and sent to mass system campaigns, aren't based on a thorough analysis of individual preferences and interests expressed and inferred. Which can be performed throw data analyzing on different sources. Allowing to improve don't only knowledge about customer, also to turn react to new likes and improve effectiveness level on recommendation. Hence we arise creation of a profiling and consumption habits intelligent system based, which react providing personalized offers, against individual preferences and needs, through events processing on physical and real world, like events on networks, geographical location and financial transaction.</p> <p>Sistemas de segmentación y envió de campañas masivas, desean mejorar el nivel de individualización en las recomendaciones. Lo anterior, se puede lograr mediante el análisis exhaustivo a nivel individual de preferencias e intereses expresados e inferidos. Intereses que pueden ser establecidos mediante el análisis de varias fuentes de datos, que permiten mejorar no solo el conocimiento acerca del cliente, si no a su vez reaccionar frente a nuevos eventos en su vida cotidiana, mejorando los niveles de efectividad en la recomendación. Es por esto que se plantea la creación de un modelo inteligente de recomendación de campañas basado en perfilamiento de hábitos de consumo, que reacciona proveyendo oferta adaptada a preferencias y necesidades del individuo. Esto mediante el análisis y procesamiento de eventos del mundo físico y virtual como: eventos en redes sociales, eventos de ubicación geográfica y eventos provenientes de transacciones financieras</p>					