

Universidad Internacional de La Rioja (UNIR)

Escuela de Ingeniería

**Máster en Análisis y Visualización de Datos
Masivos**

**Sistema Recomendador
para áreas de:
Ventas/Mercadotecnia
de Yamaha Motor de
México.**

Trabajo Fin de Máster

presentado por: Cueva Morales, Gabriela

Director: Ruiz Moreno, Carlos

Resumen

Este es un trabajo que documenta el piloto experimental: Sistema de Recomendación para áreas de Ventas o Mercadotecnia de Yamaha Motor de México, que recomienda productos a ofrecer a los distintos distribuidores de la empresa, en éste se detalla el procedimiento para resolver la problemática de la empresa a través de Inteligencia Artificial y muestra también el marco teórico general en que se engloba este desarrollo, describe brevemente el filtrado colaborativo basado en Vecindad y analiza los pasos para la construcción de un Sistema Recomendador.

Debido a la relación que tiene con Business Intelligence y Business Analytics en este trabajo se describen también los conceptos básicos y se enuncian algunas herramientas de BI y sus diferencias con los Sistemas Recomendadores.

Palabras Clave: Sistemas Recomendadores, Filtrado Colaborativo, Información Contextual, Análisis y Visualización de Datos.

Abstract

This work is the documentation of the experiment: Recommender System for Sales or Marketing departments of Yamaha Motor de México, which recommends products to offer to the company's dealers, in this work is the detailed procedure to solve the company's problem using Artificial Intelligence, and it shows the theoretical framework in which this development is encompassed, describes briefly the Neighborhood-Based Collaborative Filtering and analyze the steps to the construction of a Recommender System.

Due to the relation that there is between Business Intelligence, Business Analytics and Recommender Systems, this work describes the basic concepts and enunciates some BI tools and its differences with Recommender Systems.

Keywords: Recommender systems, Collaborative Filtering, Context-awareness, Visual Analytics.

A. Índices

Por página

Resumen.....	2
Abstract.....	2
A. Índices	3
Por página	3
De figuras	4
De tablas.....	4
B. Introducción	5
1. Marco Teórico	8
1.1 Sistemas Recomendadores Historia y Desarrollo	8
1.2 Sistemas Recomendadores hoy	9
1.3 Modelos básicos de Sistemas de Recomendación	12
1.4 Sistemas Recomendadores basados en Vecindad	14
1.5 Business Intelligence	17
1.6 Herramientas de BI.....	19
1.7 Herramientas de BI Vs Sistemas Recomendadores	22
2. Piloto Experimental Sistema Recomendador para el área de Ventas de Yamaha Motor de México.....	23
2.1 Problemática.....	23
2.2 Objetivos	24
2.3 Resultados esperados	25
2.4 Metodología.....	26
2.5 Desarrollo	29
2.5.1. Captura de Datos.....	31
2.5.2. Filtrado y Pre-procesamiento de Datos	32
2.5.3. Desarrollo de Sistema Recomendador con R.....	33
2.5.4. Desarrollo de visualización de refacciones necesarias con R	40
2.5.5. Desarrollo de interfaz de Sistema Recomendador con Shiny.....	42
2.5.6. Visualización de compras por distribuidor y modelo	45
2.5.7. Publicación de resultados en sitios públicos.....	51
2.6 Resultados	51
2.7 Discusión.....	54
2.8 Conclusiones	55
2.9 Líneas Futuras de Investigación	56
C. Anexos.....	58

D. Bibliografía	63
Libros	63
Noticias	63
Artículos	63
Enlaces	64
Bibliografía para el Desarrollo	64

De figuras

Figura 1-1 Definición de Sistemas Recomendadores	9
Figura. 1-2 Arquitectura de alto nivel de un Sistema Recomendador basado en contenido. Pasquale Lops, Marco de Gemmis y Giovanni Semeraro (2011)	16
Figura 1-3 Diferentes Tecnologías y aspectos que forman parte de la Inteligencia de Negocio creado por Forrester Research.....	18
Figura 1-4 Pentaho integración (pentaho.com).....	19
Figura 1-5 Dashboard WebFOCUS	20
Figura 2-1 Metodología de Piloto experimental.....	26
Figura 2-2 App Template - Shiny Articles	30
Figura 2-3 Filtrado de caracteres.....	32
Figura 2-4 Tabla tipo con ratings	35
Figura 2-5 Resultado de la aplicación de la similitud coseno	37
Figura 2-6 Resultado de la recomendación graficado	39
Figura 2-7 Resultado graficado de las partes requeridas.....	42
Figura 2-8 Código de Aplicación de Shiny	43
Figura 2-9 Resultado de Aplicación Shiny	44
Figura 2-10 Resultado graficado modelos comprados por distribuidor.....	48
Figura 2-11 Resultado final de Aplicación Shiny.....	50
Figura 2-12 Resumen Resultados Encuesta Preguntas 1 a 8.....	53
Figura 2-13 Resumen Resultados Encuesta Pregunta 9	54
Figura C-1 Figuras a detalle de resultados de evaluación	62

De tablas

Tabla 1 Objetivos conceptuales de Sistemas Recomendadores Básicos. Aggarwal (2016)	14
Tabla 2 Pros y Contras Filtrado Colaborativo contra Basado en Contenido.....	14
Tabla 3 Técnicas actuales de BI - JAYANTHI RANJAN (2009)	21
Tabla 4 Tabla tipo para construir matriz de ratings	33
Tabla 5 Tabla tipo con formato para tabla de ratings	35
Tabla 6 Preguntas realizada en la encuesta de satisfacción.....	52

B. Introducción

Inicié trabajando en Yamaha Motor de México en Febrero de 2003, trabajé ahí por casi 5 años y en este tiempo vi el desarrollo de esta empresa que mundialmente es una reconocida marca pero en México iba iniciando para aquel tiempo.

Trabajaban prácticamente llevando todos sus procesos manualmente, desde los procesos de Cobranza, los de Venta a distribuidores, Servicio Técnico, Campañas de mercadotecnia, etc.

Durante el tiempo que trabajé ahí se desarrollaron los principales sistemas, con miras a tener un control informático de las principales áreas de la empresa. Se adquirió la infraestructura e iniciamos a programar sistemas que estarían en la web para que los distribuidores alimentaran también los sistemas de Ventas, Refacciones y Servicio que serían los primeros en realizarse.

Hoy en día tienen este desarrollo mucho más maduro que el que iniciamos hace 15 años pero toda esta información permanece ahí y crece sin ser analizada, revisada, optimizada, etc. Es por eso que se me ha ocurrido implementar con ellos este piloto para recomendar ventas inteligentes dirigidas a los distribuidores, en base a las ventas que hacen a los distintos distribuidores.

Actualmente en el sector de motocicletas en México el mercado está dividido llevándose la marca Italica un 60% del mercado y el 40% restante se divide entre el resto de marcas incluyendo Yamaha. El financiero (2017). Naturalmente la empresa busca crecer motivando a sus distribuidores a la compra, pero al no analizar sus datos como hasta ahora, la recomendación de productos se ha venido haciendo al "feeling" de los vendedores, sin tener la información clave para hacer recomendaciones de valor que podrían hacer crecer a sus distribuidores y con ellos las ventas de la empresa.

La empresa cuenta con una red de distribuidores a todo lo largo de la República Mexicana en distintos puntos de diferente nivel socioeconómico. Cada distribuidor tiene su historia, su experiencia y sus estrategias de ventas dirigidas a su tipo de mercado específico. Algunos distribuidores son grandes otros son pequeños, la mayoría de ellos inició prácticamente con la empresa.

Además de la problemática de su área de ventas, la falta del análisis de su información impacta todas las áreas del negocio, como la de servicio y garantías, ellos desconocen el histórico de partes que más se requieren cambiarse, o los motivos por los que se requieren servicios más continuamente, o las razones por las que se reclaman garantías, etc.

Naturalmente la experiencia de sus empleados de confianza en áreas clave del negocio aportan el conocimiento y experiencia que los ha ayudado a crecer, junto con el mercado que se encuentra en constante crecimiento, pero a pesar de esto tienen una gran área de oportunidad en este sentido.

Este piloto tiene como objetivo desarrollar este sistema que recomiende a los distribuidores a través del área de ventas de la empresa, las compras adecuadas para crecer en el negocio en base a lo que distribuidores de similares características a las suyas han comprado permitiendo que el área comercial pueda analizar a mediano plazo el crecimiento del negocio en función del crecimiento de sus distribuidores.

También se pretende cruzar la información de modelo de motocicletas con la base de datos de servicio y garantías para obtener un listado de las partes que se le recomendará comprar al distribuidor pues tienen gran probabilidad de requerir un cambio a corto plazo y es de vital importancia que formen parte de su inventario. Esto último con la finalidad mejorar los plazos de respuesta del área de servicio y abonar con esto a la satisfacción del cliente.

El análisis de información hoy en día es de vital importancia para las empresas, y más aún la inteligencia artificial aplicada al negocio ayuda a las empresas a automatizar procesos de análisis, segmentación, o como en este caso de recomendación, que manualmente no son posibles de efectuar o que conllevan errores.

Por esto, estoy segura de que incorporar Inteligencia Artificial a los sistemas de Yamaha Motor de México será un parte aguas que le ayudará a la empresa a incrementar sus ventas, a mejorar su servicio al cliente aumentando su satisfacción y por supuesto les dará información clave para la toma de decisiones del negocio.

A lo largo de este trabajo se describirán los Sistemas Recomendadores, los tipos que hay, haciendo especial énfasis en el filtrado colaborativo basado en vecindad tratado en el tema 1.4, que es el utilizado en este trabajo, también se hablará de su historia, cómo han ido evolucionando y hacia dónde se dirigen.

Durante el planteamiento de este trabajo recibí la retroalimentación de un revisor que me cuestionó porqué utilizar un Sistema Recomendador en vez de una herramienta de Business Intelligence, así que me di a la tarea de analizar distintas herramientas y sus diferencias en relación a los sistemas Recomendadores , dando como resultado los temas 1.5 a 1.7 de este trabajo donde se muestran los resultados de esta investigación, se compara a los Sistemas Recomendadores con las Herramientas de Business Intelligence y se habla de los resultados obtenidos.

Por último todo el Capítulo 2 se dedica al Piloto Experimental objeto de este trabajo, se explica la forma que se encontró para resolver la problemática de Yamaha Motor de México y se describe desde su desarrollo hasta sus resultados en el marco de los Sistemas Recomendadores , su historia y desarrollo.

1. Marco Teórico

1.1 Sistemas Recomendadores Historia y Desarrollo

Primeramente me gustaría definir qué son los Sistemas Recomendadores, para lo cual consideraremos la definición dada por Francesco Ricci:

"Los Sistemas Recomendadores son herramientas de software y técnicas que proveen sugerencias de ítems para ser usados por un usuario. Las sugerencias proporcionadas van dirigidas a ayudar a sus usuarios a decidir, qué ítems comprar, qué música oír, o qué noticias leer. Está probado que los Sistemas Recomendadores son medios valiosos para los usuarios en línea para hacer frente a las grandes cantidades de información que hay y se han convertido en una de las más poderosas y populares herramientas para el comercio en línea.

'Ítem' es un término general usado para denotar lo que los sistemas recomiendan a los usuarios. Un RS normalmente se enfoca en un tipo específico de ítem y acorde a su diseño, interfaz gráfica, y a la técnica de recomendación usada para generar las recomendaciones, está personalizado para proveer útiles y efectivas sugerencias de ese tipo específico de ítem.

Los RS están dirigidos principalmente a individuos que no tienen suficiente experiencia para evaluar abrumadoras cantidades de alternativas de ítems disponibles en la Web."

Francesco Ricci (2011)

O como lo define más brevemente Ricardo Moya (2013): "Un sistema de recomendación es un sistema inteligente que proporciona a los usuarios una serie de sugerencias personalizadas (recomendaciones) sobre un determinado tipo de elementos (ítems). Los sistemas de recomendación estudian las características de cada usuario y mediante un procesamiento de los datos, encuentra un subconjunto de ítems que pueden resultar de interés para el usuario."

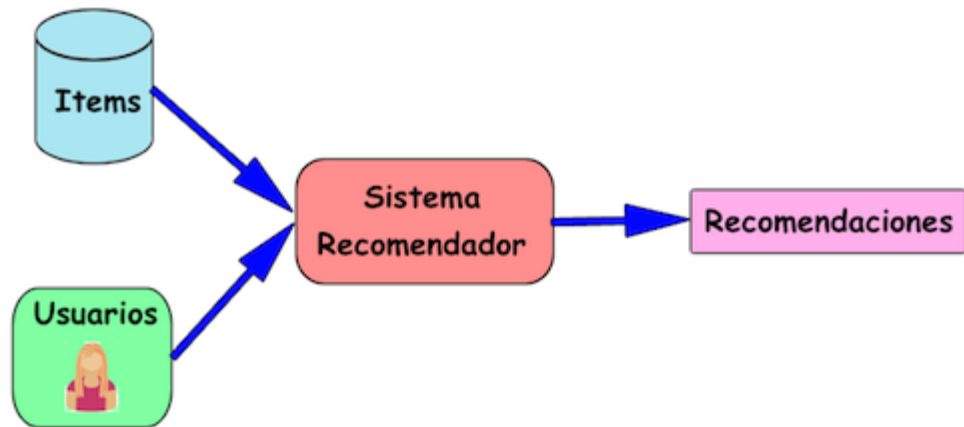


Figura 1-1 Definición de Sistemas Recomendadores

Los Sistemas Recomendadores nacen por la necesidad de ofrecer al usuario productos que le puedan resultar de interés y evidentemente para motivar a la compra de éstos.

Usualmente los Sistemas Recomendadores se han implementado en Internet para mostrar al cliente las recomendaciones, debido a la creciente venta de productos en línea y facilitando la promoción de los productos así como la retroalimentación del cliente que indica lo que le gusta y no le gusta a través de los llamados "ratings" de manera directa y explícita pero, también de manera indirecta registrando las visitas del cliente a la página del producto o cuando el cliente lo compra.

1.2 Sistemas Recomendadores hoy

El desarrollo de los Sistemas Recomendadores se ha ido enfrentando a varios retos, desde saber qué productos recomendar y saber evaluar qué tan efectiva es la recomendación, hasta temas de privacidad de la información, dado que en algunos casos se recoge la trayectoria del usuario por la página registrando sus intereses o necesidades. Pero a pesar de los retos ha permanecido en estudio durante muchos años dado a la importancia que ha significado para el usuario este valor añadido de los sitios web y, para la industria que ha visto la utilidad reflejada en sus ventas y en la lealtad de sus clientes.

Ya es una práctica común para nosotros que cuando queremos hacer una compra o asistir a algún lugar, buscamos la recomendación de otras personas que han ido, saber su opinión sobre el producto o servicio antes de hacer la compra o de adquirir el servicio es fundamental para la compra, asimismo saber que a las personas les ha gustado facilita nuestra decisión de compra e incluso la induce.

Cuando una empresa cuenta con un sistema de recomendación de productos, éste representa comodidad para el usuario, ya que en vez de navegar por miles de productos nos muestra sólo los que "sabe" que nos gustarán, nos evita perder tiempo y sobretodo nos muestra productos que tal vez de otro modo no habríamos podido encontrar.

La diversidad de servicios que podemos encontrar de este modo son ilimitados, desde ropa, películas, un restaurante que además esté cercano a nosotros, el automóvil ideal para nuestras necesidades, las vacaciones en el lugar idóneo etc. son sólo algunos ejemplos de aplicación de estos modelos.

También como señala Estela Gil en su artículo también existe el: "factor cotilla", aquel que hace que por un recomendador cambies de opinión o de compra. Aquel que te indica lo que más han comprado los usuarios."

Por otro lado, visto desde la perspectiva del negocio los Sistemas Recomendadores han significado incrementos de ventas que ni los mismos empresarios creyeron posibles, tal vez el caso de éxito más conocido de los Sistemas Recomendadores es Netflix.

Netflix originalmente era una empresa de renta de películas en formato DVD que venía redefiniéndose por varios años buscando siempre dar un servicio innovador al público, en sus inicios competía con Blockbuster y no parecía despuntar como lo hizo tras el desarrollo de los sistemas Web, el Internet y por supuesto los Sistemas Recomendadores. Su constante interés en desarrollar su empresa de una forma más moderna y funcional lo llevó a evolucionar a la renta de películas por internet y cambió de tal forma la historia de la renta de películas que la empresa fuerte que era Blockbuster y que incluso se rehusó a asociarse con Netflix terminó por quebrar y desaparecer definitivamente del mercado, dejando a Netflix como líder en el ramo. Netflix implementó un Sistema Recomendador que a partir del género que se elige,

de las puntuaciones o "raitings" de otros usuarios, y del propio usuario que de forma histórica ha generado, recomienda películas al usuario y además le explica la razón de la recomendación, "porque viste...." facilitando la tarea de la elección de películas y obteniendo la preferencia del mercado. Esta historia de éxito nos ilustra que quedarse en la zona de confort sólo puede llevar a la bancarrota y la extinción y que aplicar Inteligencia Artificial al negocio puede llevar de no figurar en el mercado a ser la empresa número uno en el sector.

Para Netflix ha sido tan importante su Sistema Recomendador que incluso realizó el GrandPrix, un concurso en el que estableció un premio millonario a quién pudiera mejorar las recomendaciones de su Sistema Recomendador. Xavier Amatriain (2012).

También Amazon utiliza los Sistemas Recomendadores para recomendar discos, libros, ropa, etc. Otra empresa de las líderes en venta de productos en línea que siempre está innovando en la aplicación de tecnología en su negocio y que es otro caso de éxito indiscutible para los Sistemas Recomendadores.

Pero tal vez el caso que más se relaciona con el presente trabajo sea el de Harley-Davidson una empresa del ramo de Yamaha Motor de México que se dedica a vender motocicletas.

Ellos al principio vendían tan sólo un par de motocicletas en su distribuidor de NY, hasta que Asaf Jacobi responsable del distribuidor de la marca en Nueva York sugirió implementar Inteligencia Artificial como parte de su estrategia de mercadotecnia y de un solo golpe logró crecer de 2 a 15 motocicletas vendidas en un fin de semana.

Ahí inició su caso de éxito con la Inteligencia Artificial, de estimar un crecimiento del 2% logró crecer sus ventas al 2,930% Brad Power (2017) dando muestra de que de establecerse en el área de confort habría efectivamente logrado ese 2% pero al adoptar la Inteligencia Artificial en el negocio pudo lograr lo que la compañía no podría ni haber imaginado tan sólo unos meses atrás.

A razón de esto muchas industrias como Harley-Davidson, Amazon, Google, Netflix, Facebook entre muchas otras han incorporado estos sistemas a sus comercios en línea obteniendo notables resultados.

Aún hasta el día de hoy se sigue estudiando la manera de hacer mejores recomendaciones al usuario, que no sean obvias, que sean apegadas a las cambiantes necesidades o gustos del usuario, mostrar productos que sorprendan al usuario, recomendar con mayor diversidad de productos, etc. también se estudian formas de explicar al usuario el porqué de las recomendaciones hechas como lo hace Netflix, para transparentar el proceso de recomendación hacia el usuario y aumentar la confianza en el sistema.

Actualmente existen muchos estudios y desarrollos en este tema usados para la investigación o en sitios comerciales y se siguen desarrollando modelos, explicaciones y algoritmos cada vez más completos y efectivos.

1.3 Modelos básicos de Sistemas de Recomendación

Los modelos básicos de Sistemas Recomendadores son los modelos basados en contenido y los modelos colaborativos, cada uno de estos tomará datos de distintas tipos de datos, el primero utilizará información de los atributos de los usuarios o de los ítems, mientras que el segundo utilizará interacciones del tipo "user-ítem" como ratings o comportamientos de compra. Otros modelos básicos son el modelo basado en conocimiento y los híbridos.

En este punto me gustaría definir también el filtrado colaborativo, con la definición que le da J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen (2007):

"El Filtrado colaborativo es el proceso de filtrar o evaluar ítems usando la opinion de otras personas. Mientras que el término filtrado colaborativo (CF) tiene sólo un poco más de una decada, CF tiene sus raíces en algo que los humanos vienen haciendo por siglos - compartiendo opiniones con otros."

Mientras que el filtrado basado en contenidos se puede definir de la siguiente manera: Ricardo Moya (2013) "Las recomendaciones se basan en el conocimiento

que se tiene sobre los ítems que el usuario ha valorado (ya sea de forma implícita o explícita), y se le recomendarán ítems similares que le puedan gustar o interesar."

Charu Aggarwal (2016) habla de seis tipos de modelos:

- » **Filtrado colaborativo basado en vecindad:** También conocido como algoritmos basados en memoria, está basado en el hecho de que usuarios similares despliegan patrones similares de comportamiento de ratings y similares ítems reciben ratings similares. Este filtrado es un modelo sencillo del tipo colaborativo.
- » **Modelo de Filtrado colaborativo:** Es un modelo que se nutre por el comportamiento de los ítems o de los usuarios. Usualmente se utilizan los ratings que dan los usuarios a los ítems para evaluar las opciones de recomendación.
- » **Sistema Recomendador basado en contenidos:** Es un tipo de SR en donde a los comportamientos de compra y ratings se les suma las características del ítem. Es recomendado para nuevos ítems que no tienen todavía información de rating.
- » **Sistema Recomendador basado en conocimiento:** Es un SR que está basado en los requerimientos o restricciones que especifican los usuarios en algunos campos de entrada explícita. Este tipo de filtrado es recomendable cuando los ítems nos son comprados muy frecuentemente.
- » **Sistemas Recomendadores demográficos:** Este SR depende de la ubicación geográfica del usuario y pretende hacerle recomendaciones en base a ésta, considerando que para el usuario es preferible encontrar un producto cercano o pensando en que las preferencias de los usuarios están determinadas por su localidad.
- » **Sistema Recomendador basado en conjunto híbrido:** Es un modelo que se basa en el hecho de que en muchos casos reales un sólo tipo de filtrado no es suficiente y busca hacer uso de combinaciones de ellos para el mismo sistema de recomendación.

- » **Sistemas Recomendadores basados en contexto:** Un reto importante en los Sistemas Recomendadores es que en diferentes dominios como el tiempo, el lugar, o círculo social, existen datos que son críticos para la recomendación, de ahí que también se consideren estos datos en este modelo.

Approach	Conceptual Goal	Input
Collaborative	Give me recommendations based on a collaborative approach that leverages the ratings and actions of my peers/myself.	User ratings + community ratings
Content-based	Give me recommendations based on the content (attributes) I have favored in my past ratings and actions.	User ratings + item attributes
Knowledge-based	Give me recommendations based on my explicit specification of the kind of content (attributes) I want.	User specification + item attributes + domain knowledge

Tabla 1 Objetivos conceptuales de Sistemas Recomendadores Básicos. Aggarwal (2016)

Usualmente la pregunta que surge es si aplicar un sistema Colaborativo o uno basado en contenidos de ahí que muchos autores nos hablan de los pros y los contras de utilizar uno y otro.

Colaborativo		Basado en Contenido	
Pros	Contras	Pros	Contras
Son sistemas más precisos	Requiere de ratings implícitos o explícitos presenta el problema del "cold start"	Nos evitamos el problema cold start	Requiere tener suficiente información del Item
Las recomendaciones no son obvias para el usuario		Las representaciones de contenido son variadas	Puede ser que la recomendación sea obvia
		Es sencillo hacer un sistema más transparente con el usuario	

Tabla 2 Pros y Contras Filtrado Colaborativo contra Basado en Contenido

1.4 Sistemas Recomendadores basados en Vecindad

Es el sistema colaborativo más sencillo, tiene dos tipos de algoritmos primarios:

1. Filtrado basado en usuarios: Se consideran los ratings que usuarios similares dan a los ítems y en base a eso se hace la recomendación.

2. Filtrado basado en ítems: Se consideran los ratings otorgados a los ítems para hacer la recomendación.

La construcción de un Sistema Recomendador requiere de la construcción de una matriz de ratings que puede formularse de dos formas:

- 1. Prediciendo el valor del rating para una combinación usuario-ítem.**
- 2. Determinando los top-k ítems más relevantes para un usuario o los top-k usuarios para un ítem determinado.**

Los ratings pueden ser: continuos, intervalos, valores ordinales, binarios, o tener un sólo valor y en base a esto se establece un algoritmo adecuado para definir la matriz de similitud que se utilizará para hacer la recomendación.

Los sistemas de recomendación suponen tres pasos a seguir para su desarrollo. Según Pasquale Lops, Marco de Gemmis y Giovanni Semeraro (2011) esta es la arquitectura de los Sistemas Recomendadores basados en contenidos, pero bien se aplica para el filtrado basado en Vecindad:

Analizador de Contenido: Es una fase que corresponde a la fase que Charu Aggarwal (2016) llama: "Preprocessing and feature extraction" se utiliza cuando la información no es estructurada, para pre-procesarla y dejarla en una forma adecuada para su utilización en los procesos posteriores. Este paso es esencial para que el Sistema Recomendador funcione óptimamente.

Perfil de Aprendizaje: Recolecta toda la información representativa de las preferencias del usuario que ha mostrado anteriormente y trata de generalizar estos datos para construir el perfil de usuario que se utilizará para intentar predecir sus intereses basado en su historial de ratings o compras. Muchas veces esta construcción se realiza usando técnicas de inteligencia artificial como machine learning.

Componente de Filtrado: Utiliza el perfil del usuario creado en el paso anterior para sugerir ítems que pueden ser de interés para el usuario en cuestión.

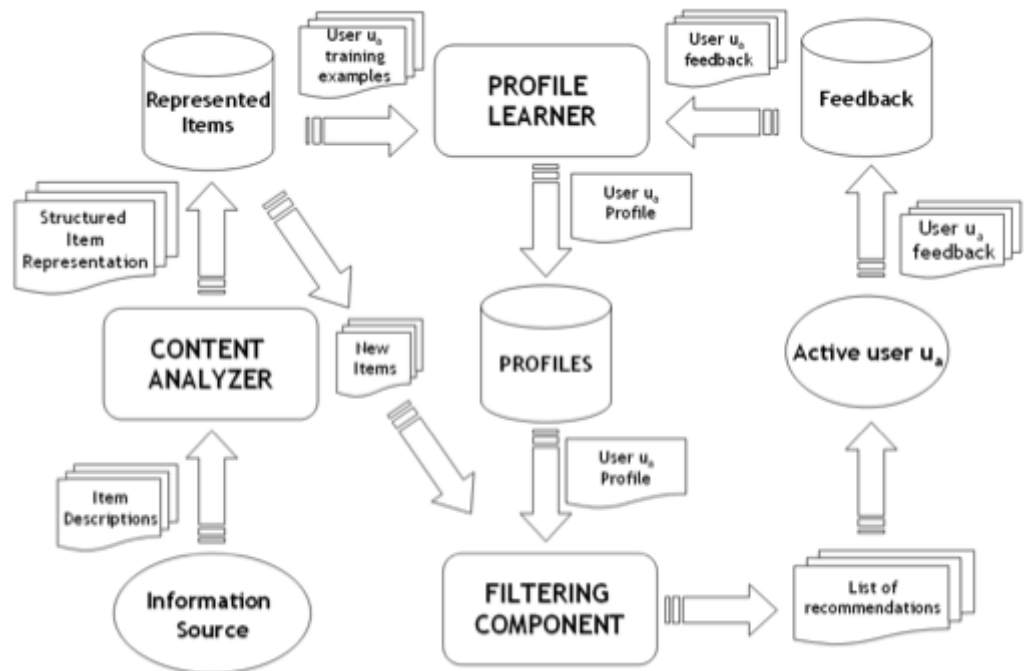


Figura. 1-2 Arquitectura de alto nivel de un Sistema Recomendador basado en contenido.
Pasquale Lops, Marco de Gemmis y Giovanni Semeraro (2011)

También durante el Máster de Análisis y Visualización de Datos Masivos de UNIR en la clase de Inteligencia Artificial vimos que además de definir el filtrado, se debe utilizar una **función de emparejamiento** que permita sugerir en base al cálculo del grado de emparejamiento entre estereotipo y contenidos para el caso de filtrado demográfico, entre el perfil de usuario y el contenido en el filtrado de contenido o entre perfiles de usuario para el caso de filtrado colaborativo. En el caso del Máster utilizamos la función de similitud coseno para este fin, aunque también existen otras como la función de Pearson o pueden usarse técnicas de inteligencia artificial que ayudan a clasificar los usuarios o ítems para permitir la recomendación, como lo son los árboles de decisión, clustering o redes neuronales.

1.5 Business Intelligence

Antes de hacer un análisis sobre las herramientas de BI me gustaría definir lo que es Business Intelligence, he aquí que encontré las siguientes definiciones:

» **Pentaho.com** [\[1\]](#)

Business Intelligence (BI): Colección de herramientas, infraestructura, y procesos que permiten a la compañía darle sentido a los datos brutos que reúne.

Información de ventas, inventarios, satisfacción del cliente contribuye al BI. Incluso los servidores que contienen la información son parte de la BI de la compañía.

El software de BI hábilmente mezcla datos de diferentes aplicaciones de BI para crear una vista que abarca todo el desempeño de la compañía.

Se puede pensar en los datos sin procesar a través de diferentes sistemas como una cascada de información, mientras que el software de BI canaliza todos esos datos en un flujo de conocimientos digeribles.

Business analytics: Es un aspecto de BI que es la suma de todas las herramientas de investigación e infraestructura de información. Estrictamente hablando se enfoca en el análisis estadístico de la información que provee el BI.

Los datos en sí mismos son a menudo ambiguos. Le corresponde al Business analytics determinar qué información sugiere y conduce a un impacto en el negocio y proveer reportes sintetizados que indiquen claramente retos y oportunidades.

Los programas de Business analytics organizan y despliegan datos en forma familiar como visualizaciones, modelos de predicción analítica y dashboards interactivos.

Estas aplicaciones ayudan a las compañías a entender su desempeño histórico, identificar obstáculos de crecimiento y encontrar oportunidades para el desarrollo del producto. El objetivo es entregar información útil tan rápido como los datos lo permitan.

» **Conesa (2010)**

Se entiende por Business Intelligence al conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización.

Algunas de las tecnologías que forman parte de Business Intelligence son:

- » Data warehouse
- » Reporting
- » Análisis OLAP
- » Análisis visual
- » Análisis predictivo
- » Cuadro de mando
- » Cuadro de mando integral
- » Minería de datos
- » Gestión de rendimineto
- » Previsiones
- » Reglas de negocio
- » Dashboards
- » Integración de datos (que incluye ETL, Extract, Transform and Load)

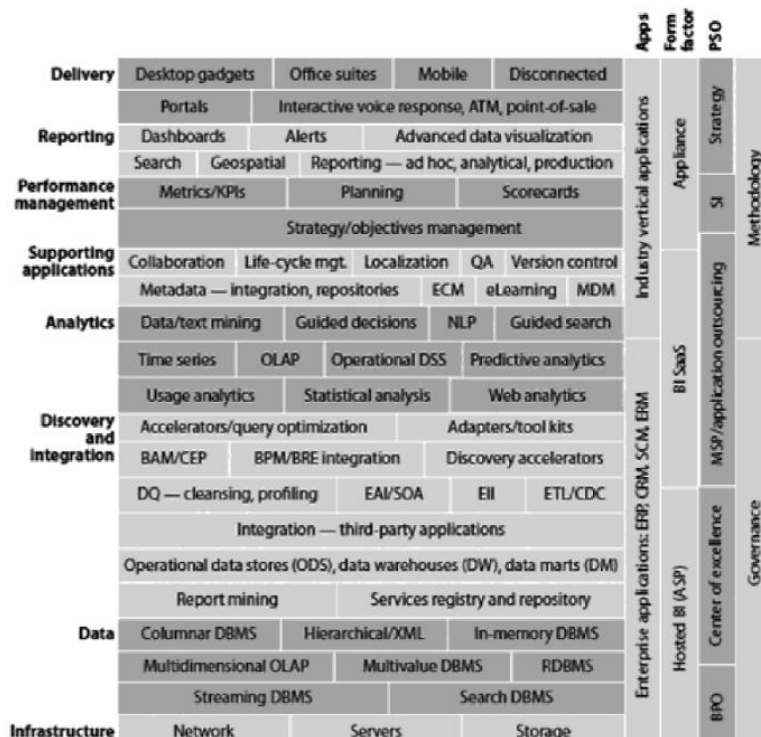


Figura 1-3 Diferentes Tecnologías y aspectos que forman parte de la Inteligencia de Negocio creado por Forrester Research

1.6 Herramientas de BI

» Pentaho [2]

Es una herramienta de BI que se utiliza para tener una vista completa del usuario y analizar la información de distintas fuentes para hacer para cada rol de la compañía una vista del usuario adaptada a sus necesidades de análisis.

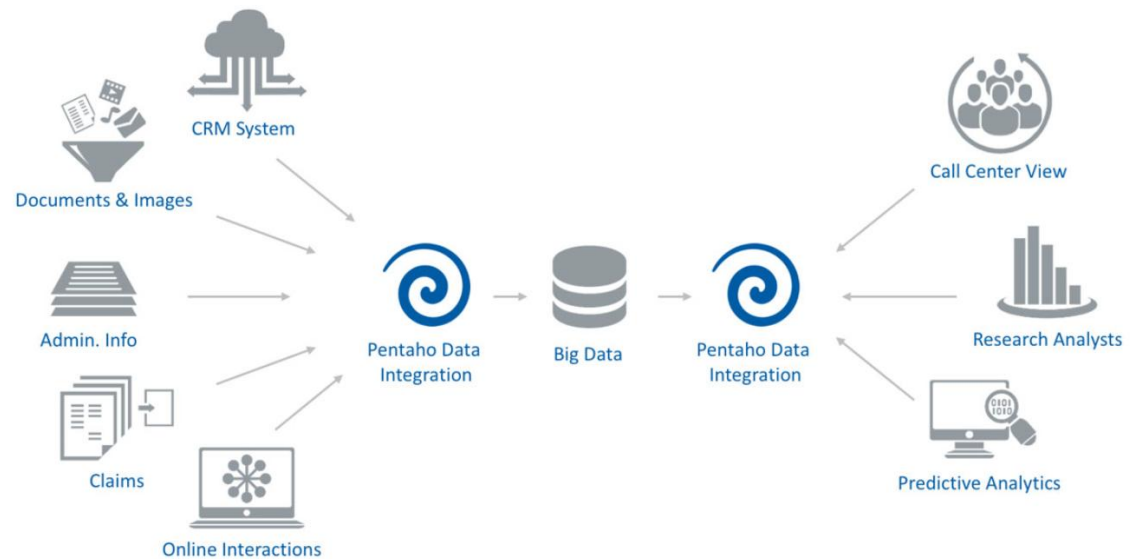


Figura 1-4 Pentaho integración (pentaho.com)

Como podemos ver integra todas las fuentes de datos, una de ellas son los análisis predictivos, investigando más a fondo encontré que estos análisis constan de predecir futuras ventas o el comportamiento del negocio en el futuro.

» Knowage PA[3]

Es también una compañía que provee un producto con una versión open source y otra empresarial que ofrece productos para trabajar con Big Data, Con metadatos, análisis predictivos, con inteligencia de locación y para analizar el éxito de la empresa permitiendo que el usuario cree sus propios dashboards con las visualizaciones que le parezcan clave.

» WebFOCUS[4]

También la herramienta de BI WebFOCUS permite el análisis de los datos del negocio para crear dashboards con el resultado de los análisis deseados.



Figura 1-5 Dashboard WebFOCUS

» **Spagobi [5]**

Es otra herramienta de BI que permite hacer lo siguiente:

- **Reportes:** Permite visualizar datos a través de reportes personalizados.
- **Análisis Multidimensional:** Enfocado a OLAP permite ver los datos con distintos niveles de profundidad.
- **Gráficas:** Permite graficar datos de manera simple o interactiva y relacionando gráficos.
- **Kpi:** Permite crear, manejar, ver y buscar en modelos hereditarios KPI a través de diferentes métodos, cálculo de reglas y alarmas.
- **Cockpits Interactivos:** Permite agregar diferentes análisis a una vista y explorar las visualizaciones dinámica y gráficamente.
- **Reporte "ad-hoc":** Permite crear reportes multi-sheets con tablas, tablas cruzadas y gráficos.
- **Location Intelligence:** Permite mostrar datos en mapas.
- **Libre Consulta:** Permite hacer exploraciones intuitivas de los datos.
- **Data mining:** Permite extraer conocimiento de grandes volúmenes de datos.
- **Análisis de redes:** Para visualizar e interpretar relaciones entre entidades animadas o inanimadas.
- **ETL:** Permite integrar datos dentro de la data warehouse y manejarlos.

- **Colaboración:** Permite que cualquier usuario pueda agregar notas y comentarios a las visualizaciones en un ambiente colaborativo.
- **Oficina automatizada:** Permite publicar documentos personales como Open Office.
- **Masterdata management:** Permite escribir a la base de datos.
- **External procesos:** Permite manejar los procesos analíticos en segundo plano y agendar su inicio o fin.

También como nos explica JAYANTHI RANJAN (2009) las técnicas utilizadas en BI Las capacidades del BI incluyen el soporte a decisiones, procesamiento analítico en línea, el análisis estadístico, previsión y minería de datos y resume en la siguiente tabla los alcances del BI.

TECHNIQUE	DESCRIPTION
Predictive modeling	Predict value for a specific data item attribute
Characterization and descriptive data mining	Data distribution, dispersion and exception
Association, correlation, causality analysis (Link Analysis)	Identify relationships between attributes
Classification	Determine to which class a data item belongs
Clustering and outlier analysis	Partition a set into classes, whereby items with similar characteristics are grouped together
Temporal and sequential patterns analysis	Trend and deviation, sequential patterns, periodicity
OLAP (OnLine Analytical Processing)	OLAP tools enable users to analyze different dimensions of multidimensional data. For example, it provides time series and trend analysis views.
Model Visualization	Making discovered knowledge easily understood using charts, plots, histograms, and other visual means
Exploratory Data Analysis (EDA)	Explores a data set without a strong dependence on assumptions or models; goal is to identify patterns in an exploratory manner

Tabla 3 Técnicas actuales de BI - JAYANTHI RANJAN (2009)

1.7 Herramientas de BI Vs Sistemas Recomendadores

Después de analizar las definiciones y algunas herramientas y suites de BI llegué a la conclusión de que el BI se enfoca en explotar la información de las empresas para su análisis, se encarga de hacer análisis estadísticos y puede hacer predicciones del comportamiento del negocio en el futuro o de la venta de productos, o mostrar visualizaciones de la aceptación del producto o marca, pero me parece que la única manera de hacer recomendaciones de productos basados en ítems o en usuarios es a través de Técnicas de Inteligencia Artificial, como los Sistemas Recomendadores. Esa es seguramente la forma más estudiada y probada para resolver este tipo de problemas, de una manera tan adaptable a las necesidades específicas de cada empresa.

En los libros y artículos que he leído hasta el momento he encontrado que los sistemas de BI van creciendo hacia los ambientes colaborativos o móviles que permitan a todos los interesados estar al pendiente del BI/A de la empresa en miras de la toma de decisiones clave en el momento adecuado. Mientras que los Sistemas Recomendadores van evolucionando hacia mejores recomendaciones al usuario en ambientes Web y hacia el filtrado colaborativo basado en redes sociales, como comentan Félix Pascual-Miguel, Julián Chaparro-Peláez y Antonio Fumero-Reverón (2011).

Realmente han ido creciendo y evolucionando tanto los estudios sobre SR y BI que tal vez en algún momento convergerán ya sea para acoplarse o para desarrollar soluciones similares, pienso que incluso este trabajo podría adaptarse a BI para hacer Sistemas Recomendadores que no sean Web ni para el usuario final sino de escritorio y para las empresas, buscando más información útil para la toma de decisiones.

2. Piloto Experimental Sistema Recomendador para el área de Ventas de Yamaha Motor de México

2.1 Problemática

La Empresa Yamaha Motor de México no hace ningún análisis de sus datos, por lo tanto, toda la labor de ventas, tanto las ventas de motocicletas como las de refacciones y accesorios, se realiza de manera intuitiva y basada en la experiencia de cada vendedor, por esta razón, los niveles de venta de cada distribuidor son muy diferentes unos de otros y basados en la habilidad que cada distribuidor tiene para lograr las ventas hacia el usuario final.

Un factor que la empresa persigue es el aumento de las ventas en el mercado. Se observa las ventas de la competencia que aventajan muchísimo a las de Yamaha y se observan también las ventas de algunos distribuidores que venden mucho más que otros, un factor que se considera clave en el fenómeno que observan es la falta de Inteligencia Artificial en el negocio, dado que al no conocer dinámicamente el comportamiento de compras de los clientes, es imposible tomar decisiones clave para el negocio que satisfagan las necesidades cada vez más cambiantes de sus clientes.

Actualmente otra de las preocupaciones de la empresa es disminuir el tiempo que pasan las motocicletas en el área de servicio cuando se trata de algún servicio o de algún reclamo de garantía, ya que esto impacta directamente en la satisfacción del cliente, punto medular del éxito del negocio y, se ha visto que muchas veces se debe a la falta de refacciones para la reparación de las unidades. Esto tiene un impacto mayor cuando algunas de las partes que se requieren son de importación y lleva meses hacerlas llegar a México. La falta de Inteligencia Artificial en el negocio, les impide saber qué refacciones deberían tener en almacén los distribuidores, para que cuando el cliente se presente puedan resolver el problema en un tiempo considerablemente más reducido.

Tratar de conducir un negocio sin hacer análisis de su información, es como querer conducir un auto sin ver el camino, por lo que se considera ésta la raíz principal del problema y es también ésta la que se pretende atacar con el desarrollo de este piloto experimental.

2.2 Objetivos

El objetivo principal de este piloto experimental es atacar la raíz de la problemática enunciada anteriormente a partir de la Inteligencia Artificial por medio de Sistemas de Recomendación y de la visualización dinámica de su información para proporcionar al área de ventas información procesada, dinámica, real, actualizada y certera como una herramienta eficaz para realizar su labor de ventas.

Para conseguir esto se realizaron las siguientes iniciativas:

Desarrollar un Sistema Recomendador que permita al asesor de ventas tener información procesada, dinámica, real y certera mostrada en una visualización eficaz para poder recomendar a los distribuidores las compras calve para aumentar sus ventas, basado en lo que distribuidores similares han comprado, con la finalidad de que se aumenten las ventas del negocio a mediano plazo.

Hacer una visualización que a partir de una motocicleta comprada se muestre las refacciones que más se desgastan por cualquier causa, para que el asesor de ventas inste al distribuidor a comprarlas y tenerlas en almacén disponibles, para dar al cliente final un servicio eficiente en tiempo y forma.

Para lograr los objetivos anteriormente mencionados se deberá conseguir los siguientes objetivos específicos:

- Analizar la información de Yamaha Motor de México para obtener la información que servirá para el Sistema Recomendador y para la recomendación de refacciones.
- Tratar los datos obtenidos para darles la forma adecuada para trabajar con ellos.
- Construir el algoritmo para el filtrado y la función de emparejamiento.
- Utilizar datos de prueba para observar el comportamiento del Sistema Recomendador.
- Implementar el Sistema en la empresa.
- Evaluar el Sistema Recomendador.

La novedad de este proyecto es que no está dirigido al cliente final de Yamaha, sino que se dirige al asesor de ventas como una herramienta que le permita visualizar las compras que puede sugerir al distribuidor basado en lo que ha comprado y a lo que otros distribuidores han comprado. Además los sistemas recomendadores suelen ser Web y estar públicos en línea dado que generalmente se encuentran en tiendas en línea, este proyecto estará en una página web privada para la empresa en beneficio directo para el asesor de ventas en su labor de ventas y buscando también beneficiar al distribuidor y al cliente final de un modo un tanto indirecto a través de afinar los procesos internos de la compañía.

2.3 Resultados esperados

La implementación de este piloto experimental en la empresa Yamaha Motor de México brindará al área de ventas no sólo más información para realizar su labor de ventas sino que brindará la información clave y además ya procesada, que le permita hacer recomendaciones eficaces que lleven al distribuidor a hacer compras inteligentes. Se espera motivar a área de ventas para realizar su labor de ventas de uno modo más informado y con un objetivo claro.

También se espera que estas recomendaciones impacten positivamente al negocio, logrando que los distribuidores incrementen sus ventas de motocicletas considerablemente y mejoren sus tiempos de respuesta en el área de servicio, dado a su capacidad de respuesta inmediata en lo que se refiere a refacciones, logrando la satisfacción de sus clientes.

El fin último de este piloto experimental es aumentar el mercado de Yamaha Motor de México a partir de la venta de productos clave y de la satisfacción del cliente.

2.4 Metodología

Para conseguir los resultados esperados de este piloto experimental se seguirá la siguiente metodología:

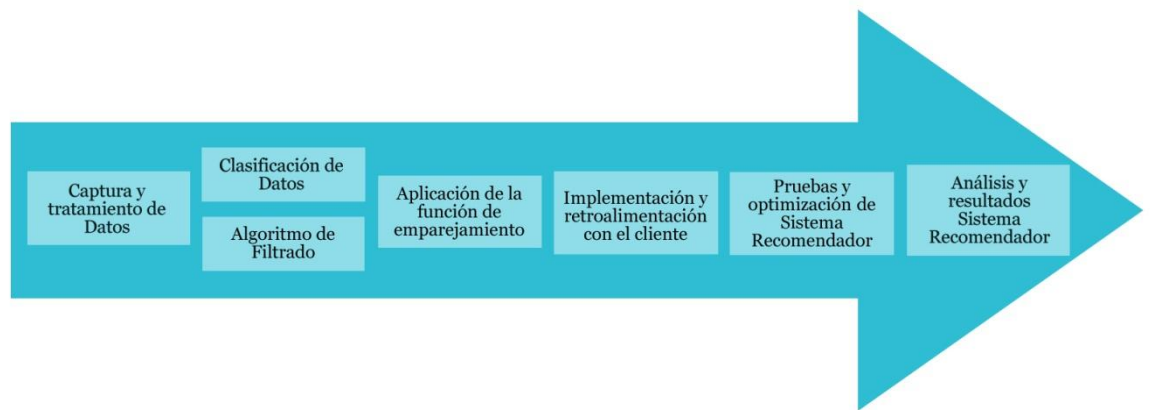


Figura 2-1 Metodología de Piloto experimental

- a. Captura y Tratamiento de Datos. En este paso se obtendrán los datos necesarios para desarrollar el Sistema Recomendador. La empresa tiene datos de venta de motocicletas, venta de refacciones y accesorios, datos de servicio y datos de garantías.

Para hacer el Sistema Recomendador de motocicletas se utilizarán los datos de venta de motocicletas, que contienen las motocicletas vendidas, con fecha, numero de modelo, distribuidor que la ha comprado, entre otros datos relacionados.

Para hacer la lista de refacciones y accesorios relacionados al modelo comprado se utilizarán los datos relacionados a servicios y reclamos de garantía, de ellos se obtendrán las partes cambiadas para el modelo consultado. De estos datos se puede saber cuántas y cuáles partes se cambiaron, en qué fecha y de qué modelo.

Para el tratamiento de datos se deberán filtrar los caracteres especiales o japoneses, así como los campos que no sean relevantes para el trabajo. Se unirán las tablas relacionadas y se les dará la forma necesaria para cada procedimiento, en forma matricial para desarrollar el Sistema

Recomendador y en forma de tabla para la recomendación de refacciones y accesorios.

Se deberán generar los datos necesarios para los dos procedimientos a realizar y de estos datos se deberá poder ordenar los datos por numero de modelo en ambos casos.

Para la parte de venta de motocicletas deberá poderse filtrar las motocicletas que se facturaron de las que realmente se compraron, también deberá poderse identificar el distribuidor que realizó la compra y la fecha de compra.

Para la parte de refacciones es importante que la tabla nos permita conocer los servicios y reclamos de garantía que se realizaron y en cada uno qué partes tuvieron que cambiarse y la cantidad de partes. Deberá ser posible identificar la fecha del servicio/reclamo y el modelo de motocicleta al que se le realizó.

- b. Clasificación de Datos. Para este trabajo no se clasificaron los datos propiamente, ya que se utilizaron todos los modelos que vende la empresa así como también se consideraron para la recomendación todas las compras hechas por todos los distribuidores.
- c. Algoritmo de Filtrado. En el caso de este trabajo se utilizará el filtrado colaborativo basado en ítems, que consiste en obtener la matriz de ratings que considera la relación usuario-ítem. Esta matriz tendrá los distintos modelos de motocicletas de Yamaha Motor de México y la cantidad de veces que han sido compradas cada una.
- d. Aplicación de la función de emparejamiento. Una vez obtenida la matriz del paso anterior nos será posible realizar la función de emparejamiento cuya función es encontrar los ítems más similares al que ha comprado el usuario por la cantidad de veces que ha sido comprado por él o por otros distribuidores, dado que no se tienen valoraciones para los distintos modelos (ratings) se utilizarán las compras realmente hechas.

Para encontrar la similitud entre modelos se utilizará la función coseno, que nos permitirá obtener una lista de los 10 modelos que se parecen más al ingresado, en relación a las compras que han hecho los distribuidores de los distintos modelos que vende la marca.

- e. Implementación y retroalimentación con el Cliente. Se deberá revisar lo obtenido con la empresa para conocer su opinión sobre el sistema y revisar si funciona correctamente y si realmente ayuda al proceso de venta de la empresa.
- f. Pruebas y Optimización de Sistema Recomendador. Estas pruebas deberán hacerse previamente a la visita a la empresa, pero también en la empresa con el personal del área de ventas. En estas pruebas se comprobará que efectivamente está recomendando correctamente, que en todos los casos se tiene información y que las visualizaciones propuestas son adecuadas.
- g. Análisis y resultados Sistema Recomendador. En conjunto con la empresa se analizará la funcionalidad del piloto experimental, se revisará si cumple con el objetivo planteado y se elaborará un informe de lo observado.
- h. El fin último del piloto experimental es lograr ganar más mercado, lo cual sólo podría evaluarse si la empresa decide emplear este Sistema de forma definitiva como parte de su estrategia de negocio y deberá decidirse el tiempo adecuado para evaluarse los resultados obtenidos desde la perspectiva financiera. Este trabajo no podrá evaluar el resultado financiero debido a que los tiempos de entrega de este trabajo y los tiempos de la empresa en la autorización de aplicación de este piloto de manera constante no se acoplan así como tampoco se ha empalmado con la estrategia de la empresa para definir el tiempo que se esperará antes de evaluar el impacto financiero.

2.5 Desarrollo

Para el desarrollo del piloto experimental elegí R por ser una herramienta que ayuda para la manipulación y tratamiento de datos que aunque no es BigData tampoco son pocos y están en tablas planas, también lo elegí porque permite hacer proyectos y porque tiene muchos paquetes para la aplicación de funciones matemáticas y estadísticas, de visualización e interfaz con el usuario y porque es open source. Los beneficios encontrados de utilizarlo en este trabajo han sido los siguientes:

- Está disponible para Linux, Windows y Mac lo que asegura que funcionará en ambiente de desarrollo, producción y real sin problemas.
- Soporta extensiones que permiten realizar distintas funcionalidades desde tratar los datos, implementar funciones estadísticas, de filtrado de datos, funciones que permiten implementar técnicas de Inteligencia Artificial, y también otros que permiten la graficación de los resultados. Esto se realiza a través de la importación de diversos paquetes disponibles en formato open source para descargar e instalar.
- Un paquete muy útil en el desarrollo de este proyecto es Shiny utilizado en este trabajo para crear una interfaz con el usuario, este paquete permite hacer un programa Web para que el usuario ingrese datos a partir de la función Input, que permite insertar campos de texto, listas de selección, botones checkbox, etc. Estos datos de entrada se procesan a partir de programas insertados en la función Server y los resultados son mostrados a partir de la función output que tiene también distintas representaciones como tablas, gráficas, imágenes, texto, etc.

ui.R

```
library(shiny)

# Define UI for miles per gallon application
shinyUI(pageWithSidebar(

  # Application title
  headerPanel("Miles Per Gallon"),

  sidebarPanel(),

  mainPanel()
))
```

Figura 2-2 App Template - Shiny Articles

- Hay mucha información en tutoriales y ayudas para aprender a implementar o resolver errores, así que esto permite que los errores que se cometen pueden resolverse, también permite conocer e implementar paquetes para la solución del problema, incluso pueden encontrarse distintas formas de solucionarlo lo que permite elegir la que ayuda más óptimamente a la solución del problema específico.

El Sistema Recomendador se realizó con el modelo de filtrado basado en vecindad que es un tipo de filtrado colaborativo, se utilizó éste dado que se cuenta con la información de ventas de motocicletas que será utilizada como las valoraciones implícitas del ítem, además considerando que los sistemas colaborativos son más eficientes que los basados en contenido y que hacen mejores recomendaciones. También se eligió debido a que una de las razones para realizar este piloto fue que se notó que los distribuidores hacen muy diversas compras y tienen muy diversos niveles de compra y se quiso motivar a la compra de los más pequeños en función de lo que los distribuidores más grandes han comprado.

La función de emparejamiento que se utilizó es la similitud coseno que es la que se ha estudiado a detalle durante el Máster en la materia de Inteligencia Artificial y se ha comprobado su funcionalidad en el caso de Sistemas Recomendadores colaborativos, como mencionan Jannach, D., Zanker, M., Felferning, A. & Friedrich,

G. (2011). "Para las recomendaciones basadas en ítem, la similitud coseno es la métrica estándar, como ha sido mostrado que produce los resultados más precisos".

La función coseno que estamos utilizando se expresa matemáticamente de la siguiente manera:

$$similitud = \cos(\vec{A}, \vec{B}) = \frac{\vec{A} \cdot \vec{B}}{||A|| \cdot ||B||} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \cdot \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Ecuación 1 Similitud Coseno Wikipedia

Donde: A y B son los componentes del vector A y B respectivamente que se están comparando.

Para la realización de visualizaciones se siguen las recomendaciones y se aplican las "best practices" estudiadas a detalle en la materia del Máster Visualización Interactiva de la información.

Para la realización de este proyecto se siguió la metodología descrita en el capítulo anterior, el código y los resultados de explicarán a detalle en los siguientes apartados.

2.5.1. Captura de Datos

Para este piloto se utilizaron los datos de la compañía referentes a motocicletas, refacciones y accesorios, datos de ventas de motocicletas, datos de venta de refacciones, datos del área de servicio y garantías, datos de distribuidores y atributos del artículo estos datos fueron exportados a archivos de texto separados por espacios.

Se hicieron los diagramas entidad-relación para entender los datos y sus relaciones y una vez corregidos errores y entendidos los datos se determinaron los datos relevantes para solucionar la problemática planteada.

2.5.2. Filtrado y Pre-procesamiento de Datos

Una vez entendida la relación entre sus tablas, procedí a revisarlas para saber si los datos son consistentes, si hay basura en las tablas, para conocer la cantidad de registros, etc.

El primer proceso que se realizó con los datos fue convertir los archivos de txt a separados por tabuladores para facilitar la importación a R.

Lo siguiente fue hacer algunos filtrados de caracteres porque contenía caracteres como: [# , " ' ´] y algunos caracteres japoneses que hacían que al ser importados los datos la tabla fuera inconsistente.

```
CY50BMX0,4MN6,50,0,CY-50 Jog
D 41.52 RS, ,0,0,"D 41.5"" R-SWIN 17 #4"
DBLADERIFL, ,0,0,DBLADE RIFLE #3-9PAS
DBLADER3-9, ,0,0,"DBLADE #3-9,A,P,SW R"
DBLADES3-9, ,0,0,"DBALDE #3-9,P,A,SW S"
DFOCUS 17R, ,0,0,DEMO FOCUS 17 R #2
DFOCUS 17S, ,0,0,DEMO FOCUS 17 S #2
DFOCUS 20R, ,0,0,DEMO FOCUS 20 R #3
DFOCUS 20S, ,0,0,DEMO FOCUS 20 S #3
DLM-001, ,0,0,DEMO LM-001
DLM-002, ,0,0,DEMO LM-002
DR2AQ1,J0B1,0,2017,DR2AQ1 GAS 2017
```

Figura 2-3 Filtrado de caracteres

También se filtraron columnas que no eran relevantes y se unieron tablas relacionadas.

Al final se obtuvieron dos documentos, uno para el desarrollo del Sistema Recomendador y otro para la recomendación de refacciones.

2.5.3. Desarrollo de Sistema Recomendador con R

Primero trabajé en el Sistema Recomendador. Para poder trabajar con los datos tratados anteriormente y obtenidos en el paso anterior fue necesario hacer lo siguiente:

- Había paquetes necesarios que no estaban incluidos en la versión básica por lo que fue necesario instalarlos. (data.table y lsa)
- Se importaron las tablas y se establecieron los campos llave para que quedara ordenada por el campo necesario.
- Se hizo un tratamiento sobre el campo de fecha para poder filtrar por fecha para obtener solo las ventas de este año.
- Se obtuvieron también las ventas realmente realizadas porque en la tabla se guardan los datos de todos los pedidos, pero no todos se concretan, algunos se quedan como cotización, otros se cancelan y otros se concretan y facturan.
- Se requirió de hacer también una suma de las ventas por modelo y por distribuidor.

Todo este tratamiento se realizó con la finalidad de tener los datos de compras por modelo porque el Sistema Recomendador será ítem a ítem y se consideran las compras como una valoración implícita del producto y el objetivo es crear una matriz de las compras por ítem de la siguiente forma:

Distribuidor	Modelo	Compras
A	A	0
B	B	1
C	C	10

Tabla 4 Tabla tipo para construir matriz de ratings

Primero instalé paquetes necesarios `install.packages("data.table")`

```
library(data.table)
```

El siguiente paso fue importar las tablas necesarias

```
pedidodet<-read.table("ventas/pedidosdetail.TXT",sep = "\t",fill=TRUE,  
fileEncoding="latin1")  
pedidos<-read.table("ventas/pedidos.TXT",sep = "\t",fill=TRUE,  
fileEncoding="latin1")
```

Establecemos las tablas como `data.table` y creamos los índices para poder trabajar con los datos

```
pedidodet<-data.table(pedidodet)
setkey(pedidodet,V1)
pedidos<-data.table(pedidos)
setkey(pedidos,V1)
```

Para este piloto consideraremos el año actual para lo cual tomaremos sólo el año para hacer un análisis de este año solamente y lo actualizamos en la tabla de pedidos:

```
pedidos$V4<-as.numeric(year(as.Date(pedidos$V4,"%d/%m/%Y")))
```

Creé una nueva tabla considerando los datos sólo de este año:

```
pedidosnow<-pedidos[(pedidos$V4>2016),]
```

Tomamos los datos de detalle del pedido que contiene los datos de los modelos comprados. Considerando sólo los pedidos de este año y luego consideramos sólo los pedidos que no tienen cancelación y sí fueron enviados:

```
modxped<-pedidodet[(pedidodet$V1 %in% pedidosnow$V1),]
modxped<-modxped[(modxped$V6 == 0) & (modxped$V5 > 0),]
```

Escribimos el usuario (distribuidor) según el numero de pedido en `modxped`. Sólo consideramos las columnas necesarias:

```
modxped<-modxped[,c(1:2,5)]
pedidosnow<-pedidosnow[,c(1,5)]
```

Hacemos un left outer join in R para agregar distribuidor a la tabla de detalle de pedidos:

```
df<-merge(x=modxped,y=pedidosnow,by="V1",all.x=TRUE)
```

Eliminamos columna pedido pues es irrelevante y Sumar las compras de cada distribuidor por modelo comprado:

```
colnames(df)[2] <- "Mdl"
colnames(df)[3] <- "Qty"
colnames(df)[4] <- "Dlr"
df<-data.table(df)
setkey(df,Dlr)
df<-aggregate(df$Qty, by=list(Dlr=df$Dlr,Mdl=df$Mdl), FUN=sum)
```

En este punto hemos logrado llegar a la tabla tipo con los ratings que necesitamos para hacer la recomendación.

Dlr	Mdl	x
<int>	<fctr>	<int>
10138	AF115SCD	1
10998	DR2AQ2	2
40096	DR2AQ2	2
40103	DR2AQ2	1
40106	DR2AQ2	1
40107	DR2AQ2	1
10111	DR2EQ4	5
10118	DR2EQ4	1
10998	DR2EQ4	7
30120	DR2EQ4	1

1-10 of 1,634 rows

Previous 1 2 3 4 5 6 ... 100 Next

Figura 2-4 Tabla tipo con ratings

Para poder aplicar la función de emparejamiento requerimos tener nuestros datos en una tabla donde aparezcan todos los distribuidores y todos los modelos, que reflejen las compas que han hecho y que en los modelos donde no han comprado aparezca en ceros, porque ahora es necesario hacer lo siguiente:

- Necesitamos saber cuáles son los diferentes modelos que existen y se han comprado y cuál es su posición en la tabla.
- También requerimos cambiar la forma en que se encuentra de long a wide, para que aparezca la información tipo matriz de la siguiente forma:

Ítem/Compras	A	B	C	D
1	0	2	1	10
2	1	2	0	0
3	3	0	1	1
4	5	2	1	0
5	0	0	1	1

Compras hechas por los distribuidores que se consideran como ratings dados a los diferentes modelos.

Tabla 5 Tabla tipo con formato para tabla de ratings

- Después, requerimos eliminar al distribuidor porque no deseamos que se incluya en el cálculo dado que lo que deseamos comparar son los ítems.
- Luego, convertimos la tabla a tipo matriz que es la estructura de datos requerida para aplicar la función coseno.
- Finalmente, aplicamos la función coseno a la matriz resultante del paso anterior.

Obtengo los modelos comprados, para poder obtener su posición en la matriz y para comprobar que fueron añadidos todos los modelos

```
dfmode<-df[,c(2:3)]
colnames(dfmode)[1] <- "Mdl"
dfmode<-data.frame(dfmode)
dfxmod<-aggregate(dfmode$x, by=dfmode["Mdl"], FUN=sum)
colnames(df)[3] <- "Total"
```

Cambiando la formade long a wide para que tenga la forma deseada para aplicar la similitud coseno:

```
dfwide<-reshape(df, idvar = "Dlr", timevar = "Mdl", direction = "wide")
```

Antes de aplicar el coseno quitamos el distribuidor de la tabla para que no lo incluya en el cálculo:

```
dfwide2<-dfwide[,c(2:109)]
```

Creando funciones para pasar parámetros y obtener resultados y aplicando la funcion coseno a la matriz wide creada.

Primero cambiamos campos NA por 0: install.packages("lsa")

```
dfwide2[is.na(dfwide2)] <- 0
library(lsa)
## Warning: package 'lsa' was built under R version 3.3.3
## Loading required package: SnowballC
```

Convertimos a matriz la tabla y aplicamos el coseno:

```
dfwide2<-as.matrix(dfwide2)
mtrxsim<-cosine(dfwide2)
```

Hemos obtenido la matriz de similitud resultante del procesamiento de la tabla de ratings con los resultados de la aplicación de la similitud coseno:

	Total.AF115SCD	Total.DR2AQ2	Total.DR2EQ4	Total.DR2EQ4 AC	Total.EF1000IS	Total.EF2000IS	Total.EF2000ISV2	Total.EF2600D	Total
Total.AF115SCD	1.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	
Total.DR2AQ2	0.000000000	1.000000000	0.515406100	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	
Total.DR2EQ4	0.000000000	0.515406100	1.000000000	0.046524211	0.065795169	0.088273483	0.000000000	0.000000000	
Total.DR2EQ4 AC	0.000000000	0.000000000	0.046524211	1.000000000	0.471404521	0.527046277	0.000000000	0.000000000	
Total.EF1000IS	0.000000000	0.000000000	0.065795169	0.471404521	1.000000000	0.596284794	0.000000000	0.000000000	
Total.EF2000IS	0.000000000	0.000000000	0.088273483	0.527046277	0.596284794	1.000000000	0.000000000	0.000000000	
Total.EF2000ISV2	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	1.000000000	0.000000000	
Total.EF2600D	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	1.000000000	
Total.EF2600FW	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000	
Total.EF3000ISE	0.000000000	0.23210354	0.043863446	0.314269681	0.222222222	0.447213595	0.000000000	0.000000000	

Showing 1 to 12 of 108 entries

Figura 2-5 Resultado de la aplicación de la similitud coseno

Ya en esta matriz se pueden obtener manualmente y con la sola observación los modelos más recomendados a partir de los más similares a cada modelo, pero para ver aplicada la recomendación se hará un programa que permita ingresar un modelo que supondremos que es el comprado por el distribuidor, para que calcule la similitud de ese modelo con el resto y nos devuelva los más similares.

Creamos una función que permita interactuar con el usuario para obtener un modelo y devolver las recomendaciones del Sistema:

```

buscaritem<-function(val){
  item<-which(dfxmod$Mdl == val)
  vector<-mtrxsim[item:item,]
  vector[[item]]<-0
  vector<-vector[vector>0]
  vector<-sort(vector,decreasing=TRUE)
  head(vector,10)
}

```

El resultado de este programita de interfaz con el usuario ingresando el modelo XT12U7 es el siguiente:

Esta función nos permite ingresar un modelo y obtener las primeras 10 recomendaciones de la matriz de similitud creada anteriormente:

```

buscaritem("AF115SCD")
## Total.XTZ125EBF Total.XTZ250BG Total.YB125BH Total.YB125ZRBG
## 1.000000000 0.40824829 0.15778550 0.08340286
## Total.XC115BG Total.YBR125CBH Total.YB125ZRBH Total.FZSCH

```

```
##          0.07557940          0.07408531          0.06786116          0.04546223
## Total.YZFR15CGES    Total.MTN320BJ
##          0.03533326          0.02658409
```

Como la vista no es muy fácil de interpretar la pondremos en una gráfica:

```
library(data.table)

buscaritem<-function(val) {
  item<-which(dfxmod$Mdl == val)
  vector<-mtrxsim[item:item,]
  vector[[item]]<-0
  vector<-vector[vector>0]
  vector<-sort(vector,decreasing=TRUE)
  vector<-head(vector,10)
  vector
  par(oma=c(2,5,3,5))
  barplot(vector, names.arg=dimnames(vector), las=2, adj=1,
oma=c(20,2,2,2), mgp=c(3,0.5,0) )
}

buscaritem("AF115SCD")
```

El resultado de la recomendación para el modelo consultado y ya graficado es el siguiente:

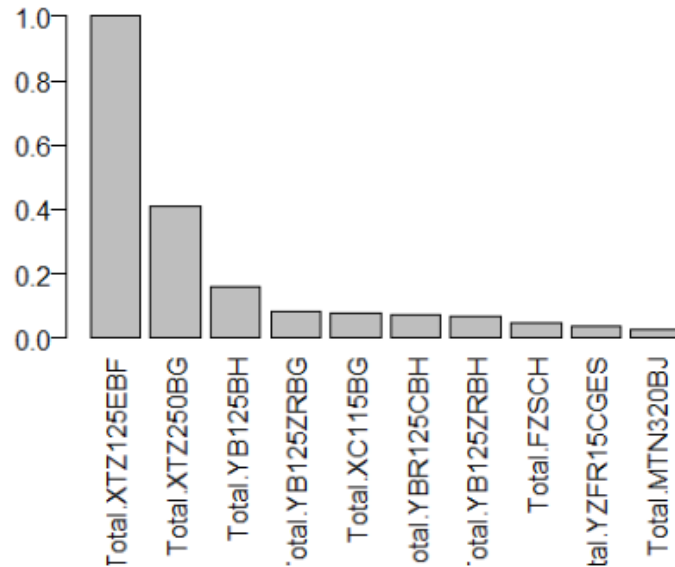


Figura 2-6 Resultado de la recomendación graficado

Podemos comprobar fácilmente viendo la matriz de similitud que efectivamente los resultados de la recomendación mostrados en la gráfica, corresponden a los modelos más similares al consultado.

2.5.4. Desarrollo de visualización de refacciones necesarias con R

Otra funcionalidad importante para apoyar al área de ventas de refacciones es mostrar para cada modelo ingresado las refacciones necesarias según los registros de servicio y garantías.

Para poder mostrar esta visualización se requiere hacer lo siguiente:

- Importar las tablas de servicio, garantías, partes y distribuidores.
- Tratamos las tablas haciendo pre-procesamiento para eliminar caracteres y darles formato.
- Unimos las tablas de reclamos con la de modelos y la de distribuidores con la finalidad de saber qué modelos son objeto del reclamo de servicio o garantía y también para saber qué distribuidor fue al que el cliente llegó para hacer su servicio o garantía.
- Ya teniendo la tabla anterior se filtrarán los campos que no sea relevantes y se le dará formato a la tabla.
- Se hará una última unión, ahora con la tabla de partes para poder saber qué partes se cambiaron producto del servicio o garantía realizada, esto nos permitirá saber qué partes se cambian para cada modelo y se hace una sumatoria.

Se consideró importante incluir las Refacciones necesarias para el modelo ingresado, por lo que importamos datos de Servicio para obtener las partes más demandadas en servicio o garantía

claimWAuMdluDlr

PartnfW2_tab

Filtramos los campos para quedarnos sólo con los necesarios de cada tabla:

```
claimWAuMdluDlr<-  
read.table("Servicio/claimWAuMdluDlr.txt",header=TRUE,sep = "\t",fill=TRUE,  
fileEncoding="latin1")  
PartnfW2_tab<-read.table("Servicio/PartnfW2 tab.txt",sep =  
"\t",fill=TRUE, fileEncoding="latin1")  
claimWAuMdluDlr<-claimWAuMdluDlr[,c(1:2,5:6,15:16,23)]  
PartnfW2_tab<-PartnfW2_tab[,c(1:5,8)]
```


Ponemos nombre a los campos para poder unir las tablas

```
colnames(PartnfW2_tab)[1] <- "Poliza"  
colnames(PartnfW2_tab)[2] <- "primaria"  
colnames(PartnfW2_tab)[3] <- "Parte"  
colnames(PartnfW2_tab)[4] <- "Desc"  
colnames(PartnfW2_tab)[5] <- "Qty"  
colnames(PartnfW2_tab)[6] <- "Reclamo"
```

Unimos tablas de claim y partes para poder obtener las partes más reclamadas en servicio o garantía:

```
partes<-  
merge(x=claimWAuMdluDlr,y=PartnfW2_tab,by=c("Poliza","Reclamo"),all.x=TRUE)
```

- Finalmente hacemos un programa que permita la interacción del asesor de ventas para que escriba el modelo que desea consultar y le mostraremos las partes que se cambian para el modelo seleccionado en una gráfica ordenada.

Sumamos las partes que se requirieron en Servicios y/o garantías para obtener los totales de partes, le ponemos nombre a la columna de total de partes y lo establecemos como llave para que ordene la tabla por cantidad de partes. Esto nos será útil al graficar:

```
partesSum<-aggregate(partes$Qty,  
by=list(modelo=partes$modelo,Parte=partes$Parte,Desc=partes$Desc), FUN=sum)  
  
colnames(partesSum)[4] <- "Qty"  
partesSum<-data.table(partesSum)  
setkey(partesSum,modelo)
```

Tomamos un modelo para ver cómo se presenta la gráfica:

```
parts<-which(partesSum$modelo == "AF115SCD")  
vector<-partesSum[parts,]  
vector<-vector[,c(2,4)]  
par(oma=c(0,5,3,5))  
barplot(vector$Qty, names.arg=vector$Parte, horiz=1, las=1,  
mgp=c(2,0.5,0), main = "Most Spare Parts Needed",xlab="Spare Parts Amount"  
)
```

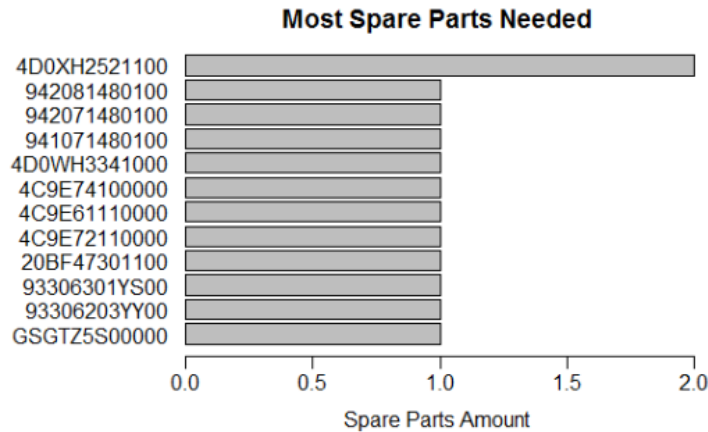


Figura 2-7 Resultado graficado de las partes requeridas

En este caso se ingresó para el ejemplo el modelo AF115SCD y se obtuvieron las 12 partes que se requieren para hacer servicios o garantías a este modelo.

2.5.5. Desarrollo de interfaz de Sistema Recomendador con Shiny

Dado que este piloto experimental tiene por objetivo las áreas de ventas y mercadotecnia de Yamaha Motor de México se implementó este Sistema Recomendador en Shiny, que permite crear una interfaz web para que el usuario interactúe con el programa realizado en R y obtenga los resultados en pantalla con un formato intuitivo y sencillo de entender.

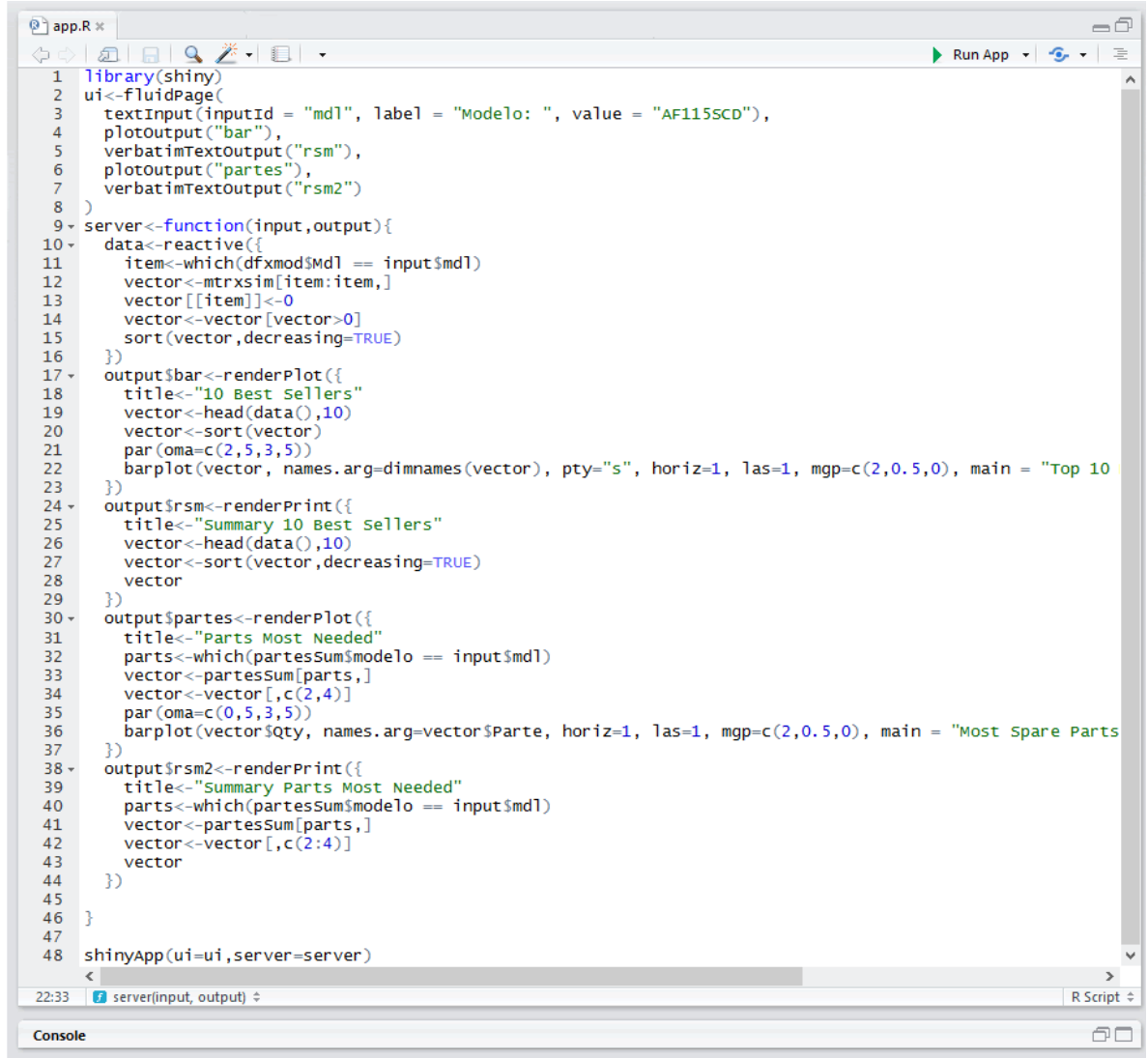
Para lo cual implementaremos:

1 entrada de texto donde el asesor de ventas escribirá el modelo a consultar y

4 salidas:

- Gráfica para mostrar las recomendaciones de compra de motocicletas.
- Visualización de texto con un resumen de partes y su nivel de similitud con el modelo consultado.
- Gráfica para mostrar las refacciones requeridas para el modelo consultado.
- Visualización en una tabla para mostrar las refacciones por código, su descripción y la cantidad de partes que se requieren.

El código desarrollado para construir esta interfaz con Shiny del Sistema Recomendador es el siguiente:



```
1 library(shiny)
2 ui<-FluidPage(
3   textInput(inputId = "mdl", label = "Modelo: ", value = "AF115SCD"),
4   plotOutput("bar"),
5   verbatimTextOutput("rsm"),
6   plotOutput("partes"),
7   verbatimTextOutput("rsm2")
8 )
9 server<-function(input,output){
10   data<-reactive({
11     item<-which(dfxmod$mdl == input$mdl)
12     vector<-mtrxsim[item:item,]
13     vector[[item]]<-0
14     vector<-vector[vector>0]
15     sort(vector,decreasing=TRUE)
16   })
17   output$bar<-renderPlot({
18     title<-"10 Best Sellers"
19     vector<-head(data(),10)
20     vector<-sort(vector)
21     par(oma=c(2,5,3,5))
22     barplot(vector, names.arg=dimnames(vector), pty="s", horiz=1, las=1, mgp=c(2,0.5,0), main = "Top 10")
23   })
24   output$rsm<-renderPrint({
25     title<-"Summary 10 Best Sellers"
26     vector<-head(data(),10)
27     vector<-sort(vector,decreasing=TRUE)
28     vector
29   })
30   output$partes<-renderPlot({
31     title<-"Parts Most Needed"
32     partes<-which(partessum$modelo == input$mdl)
33     vector<-partessum[partes,]
34     vector<-vector[,c(2,4)]
35     par(oma=c(0,5,3,5))
36     barplot(vector$Qty, names.arg=vector$Parte, horiz=1, las=1, mgp=c(2,0.5,0), main = "Most Spare Parts")
37   })
38   output$rsm2<-renderPrint({
39     title<-"Summary Parts Most Needed"
40     partes<-which(partessum$modelo == input$mdl)
41     vector<-partessum[partes,]
42     vector<-vector[,c(2:4)]
43     vector
44   })
45 }
46 }
47
48 shinyApp(ui=ui,server=server)
```

Figura 2-8 Código de Aplicación de Shiny

El resultado de la aplicación de esta aplicación es el que se genera en formato html y se puede observar enseguida:

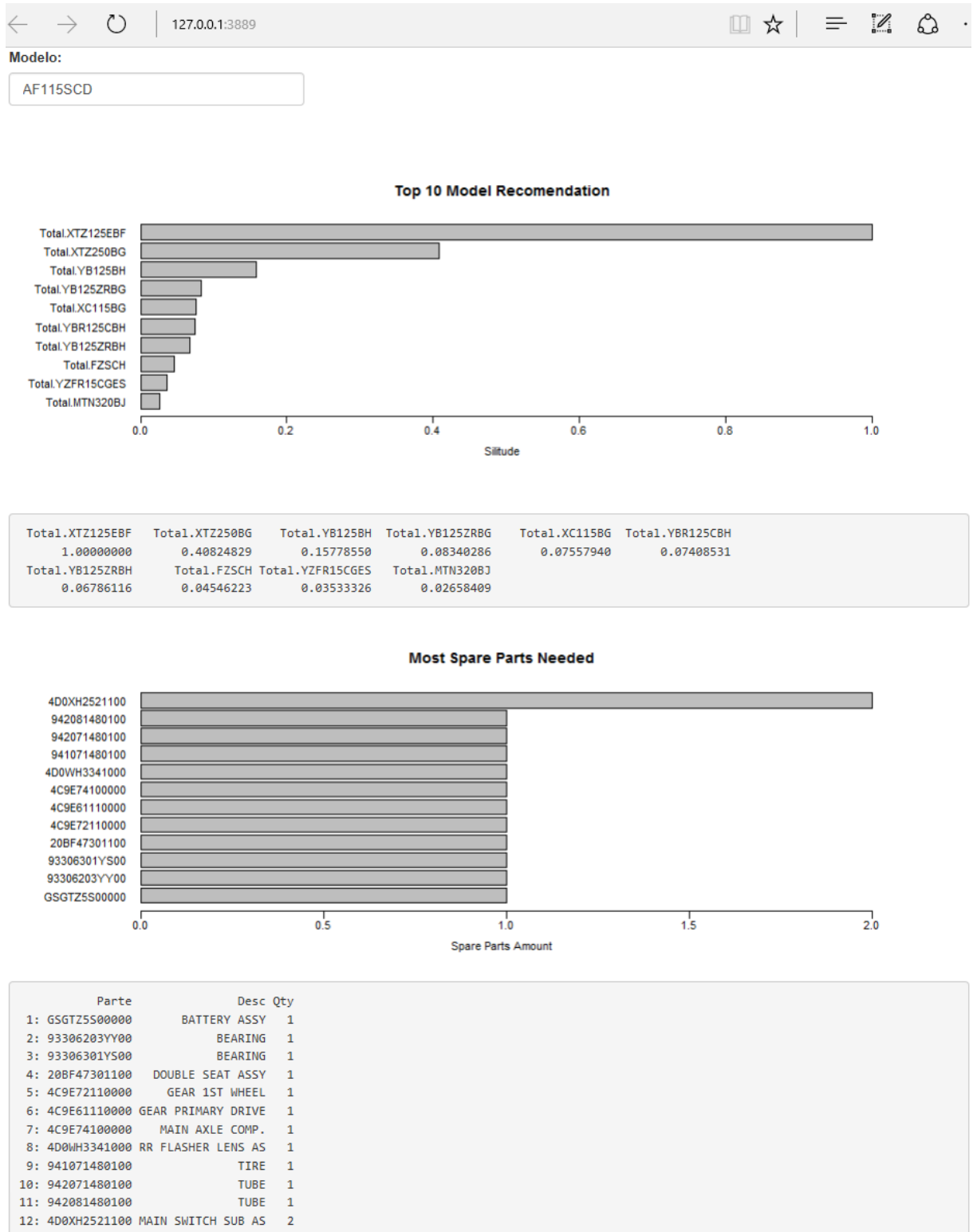


Figura 2-9 Resultado de Aplicación Shiny

2.5.6. Visualización de compras por distribuidor y modelo

Como una ayuda adicional al agente de ventas se desarrolló una visualización de las compras de los distribuidores del último mes que se tiene registro con la finalidad de mostrarle los modelos comprados y que no tenga que hacer ninguna consulta adicional para obtener este dato y para poder alimentar con el dato de modelo al Sistema Recomendador seleccionando el distribuidor deseado.

Para proporcionar esta funcionalidad al asesor de ventas se realizó lo siguiente:

- Se filtró la tabla de ventas por distribuidor para obtener las ventas del mes más reciente con que se cuenta información.
- Se realizó la suma de compras por distribuidor y modelo.
- Se realizó la visualización de la información obtenida.

Se explicará a detalle el procedimiento enseguida:

Ayudaremos al asesor de ventas mostrándole las compras de un distribuidor deseado. Para eso haremos lo siguiente:

Para tomar los últimos dos meses de ventas de motocicletas consideramos el año en curso

```
library(lubridate)
## Warning: package 'lubridate' was built under R version 3.3.3
##
## Attaching package: 'lubridate'
## The following objects are masked from 'package:data.table':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday,
##     week, yday, year
## The following object is masked from 'package:base':
##
##     date
pedidosnow<-pedidos[(pedidos$V4>2016),]
pedidosnow<-
pedidosnow[(as.numeric(year(as.Date(pedidosnow$V3,"%d/%m/%Y")))==2017),]
```

El último mes capturado es marzo por lo que consideramos los datos cuyo mes es mayor a 02 (febrero)

```
pedidosnow<-  
pedidosnow[(as.numeric(month(as.Date(pedidosnow$V3,"%d/%m/%Y")))>02),]
```

Tomamos los datos de detalle del pedido que contiene los datos de los modelos comprados. CONsiderando sólo los pedidos de este año

```
modxped<-pedidodet[(pedidodet$V1 %in% pedidosnow$V1),]
```

Consideramos sólo los pedidos que no tienen cancelación y sí fueron enviados

```
modxped<-modxped[(modxped$V6 == 0) & (modxped$V5 > 0),]
```

Escribimos el usuario (distribuidor) según el numero de pedido en modxped. Sólo consideramos las columnas necesarias

```
modxped<-modxped[,c(1:2,5)]  
pedidosnow<-pedidosnow[,c(1,5)]
```

Haciendo un left outer join in R para agregar distribuidor a la tabla de detalle de pedidos y eliminamos campos innecesarios

```
df<-merge(x=modxped,y=pedidosnow,by="V1",all.x=TRUE)  
df<-df[,c(2:4)]
```

Sumamos las compras de cada distribuidor por modelo comprado

```
colnames(df)[1] <- "Mdl"  
colnames(df)[2] <- "Qty"  
colnames(df)[3] <- "Dlr"  
df<-data.table(df)  
setkey(df,Dlr)  
  
df<-aggregate(df$Qty, by=list(Dlr=df$Dlr,Mdl=df$Mdl), FUN=sum)
```

Damos formato al dlr

```
regs<-which(as.numeric(as.character(df$Dlr)) < 10)  
df$Dlr[regs]<-paste("0000",df$Dlr[regs])  
regs<-which(as.numeric(as.character(df$Dlr)) < 100)  
## Warning in which(as.numeric(as.character(df$Dlr)) < 100): NAs  
introducidos  
## por coerción
```

```
df$Dlr[regs]<-paste("000",df$Dlr[regs])
regs<-which(as.numeric(as.character(df$Dlr)) < 1000)
## Warning in which(as.numeric(as.character(df$Dlr)) < 1000): NAs
introducidos
## por coerción
df$Dlr[regs]<-paste("00",df$Dlr[regs])

df$Dlr<-gsub(" ", "", df$Dlr)
```

Importamos CLientes y filtramos para obtener solo el nombre

```
clientes<-read.table("ventas/clientes.TXT",sep = "\t",fill=TRUE,
fileEncoding="latin1")
clientes<-clientes[,c(1:3)]
```

Obtenemos nombre de dealer de tabla cliente

```
colnames(clientes)[1] <- "Dlr"
colnames(clientes)[2] <- "Nm"
colnames(clientes)[3] <- "Srt"
```

Damos formato al campo Dealer y lo escribimos en la tabla df

```
regs<-which(as.numeric(as.character(clientes$Dlr)) < 10)
clientes$Dlr[regs]<-paste("0000",clientes$Dlr[regs])
regs<-which(as.numeric(as.character(clientes$Dlr)) < 100)
## Warning in which(as.numeric(as.character(clientes$Dlr)) < 100): NAs
## introducidos por coerción
clientes$Dlr[regs]<-paste("000",clientes$Dlr[regs])
regs<-which(as.numeric(as.character(clientes$Dlr)) < 1000)
## Warning in which(as.numeric(as.character(clientes$Dlr)) < 1000): NAs
## introducidos por coerción
clientes$Dlr[regs]<-paste("00",clientes$Dlr[regs])
clientes$Dlr<-gsub(" ", "", clientes$Dlr)

df<-merge(x=df,y=clientes,by="Dlr",all.x=TRUE)
```

Obtengo los distintos distribuidores que han comprado para selectInput vector para shiny, y lo ordenamos ascendentemente

```
df2<-data.table(df)
keycols<-c("Dlr", "Mdl")
setkeyv(df2, keycols)
dlrs<-unique(df2[list(Dlr, Nm, Srt)])
dlrs<-data.table(dlrs)
```

Con estos datos ya podemos graficar

```
parts<-which(df$Dlr == "00001")
dlrW<-which(dlrs$Dlr == "00001")
dlrNm<-dlrs[dlrW,]
vector<-df[parts,]
vector<-vector[,c(2,3)]
par(oma=c(0,5,3,5))
barplot(vector$x, names.arg=vector$Mdl, horiz=1, las=1, mgp=c(2,0.5,0),
main = dlrNm$Nm, xlab="Total Models Bought" )
```

Obtenemos la gráfica que muestra los resultados de la siguiente forma:

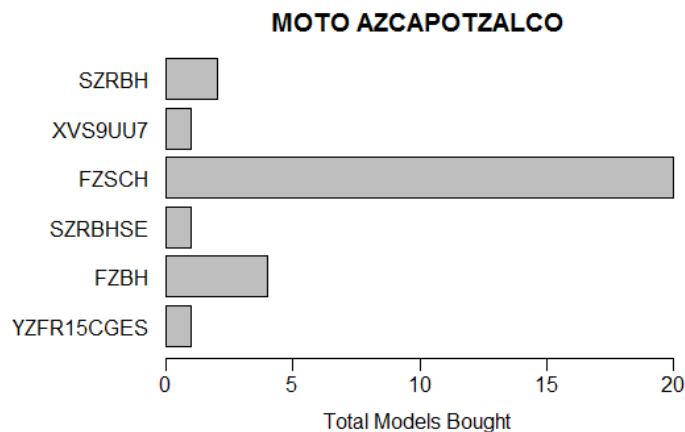


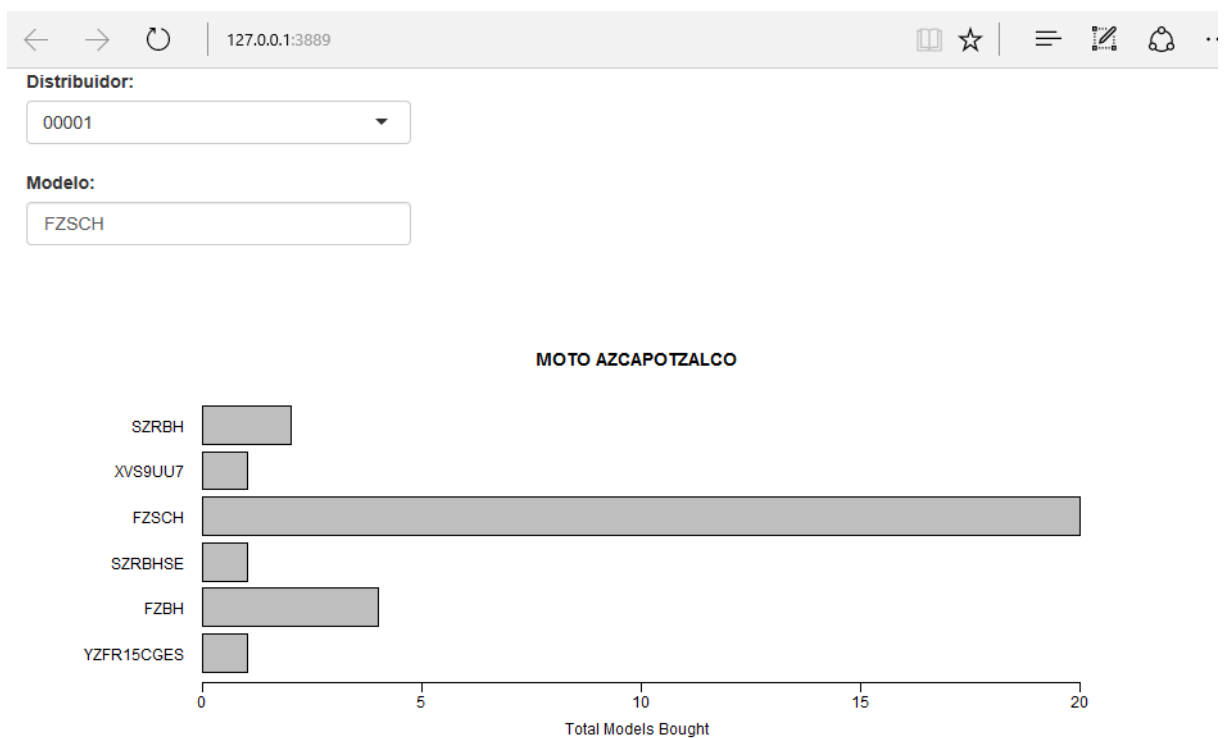
Figura 2-10 Resultado graficado modelos comprados por distribuidor

Podemos observar que en esta gráfica es posible reconocer fácilmente qué modelos son los comprados por el distribuidor en el último mes y en qué medida ha comprado cada uno en base a esto el asesor de ventas podrá elegir alguno para obtener sus recomendaciones.

Pasamos esto a Shiny junto con el desarrollo anterior para que aparezca en la misma pantalla. Para ello creamos:

- Una nueva entrada Distribuidor donde el asesor de ventas seleccionará de una lista a alguno de los distribuidores que han comprado motocicletas durante el mes. Esta aparecerá adicional a la entrada de modelo que sigue apareciendo debajo de esta nueva entrada.
- Un nuevo gráfico con las compras del Distribuidor seleccionado que muestra las etiquetas de los modelos comprados para facilitar la tarea de elegir modelo a consultar como entrada para el Sistema Recomendador.

En el caso del ejemplo aquí expuesto se dejó el primer distribuidor (aparecen ordenados en forma ascendente) y tomé como modelo el de mayor compras para este distribuidor "FZSCH" para hacer la recomendación y los resultados son los siguientes:



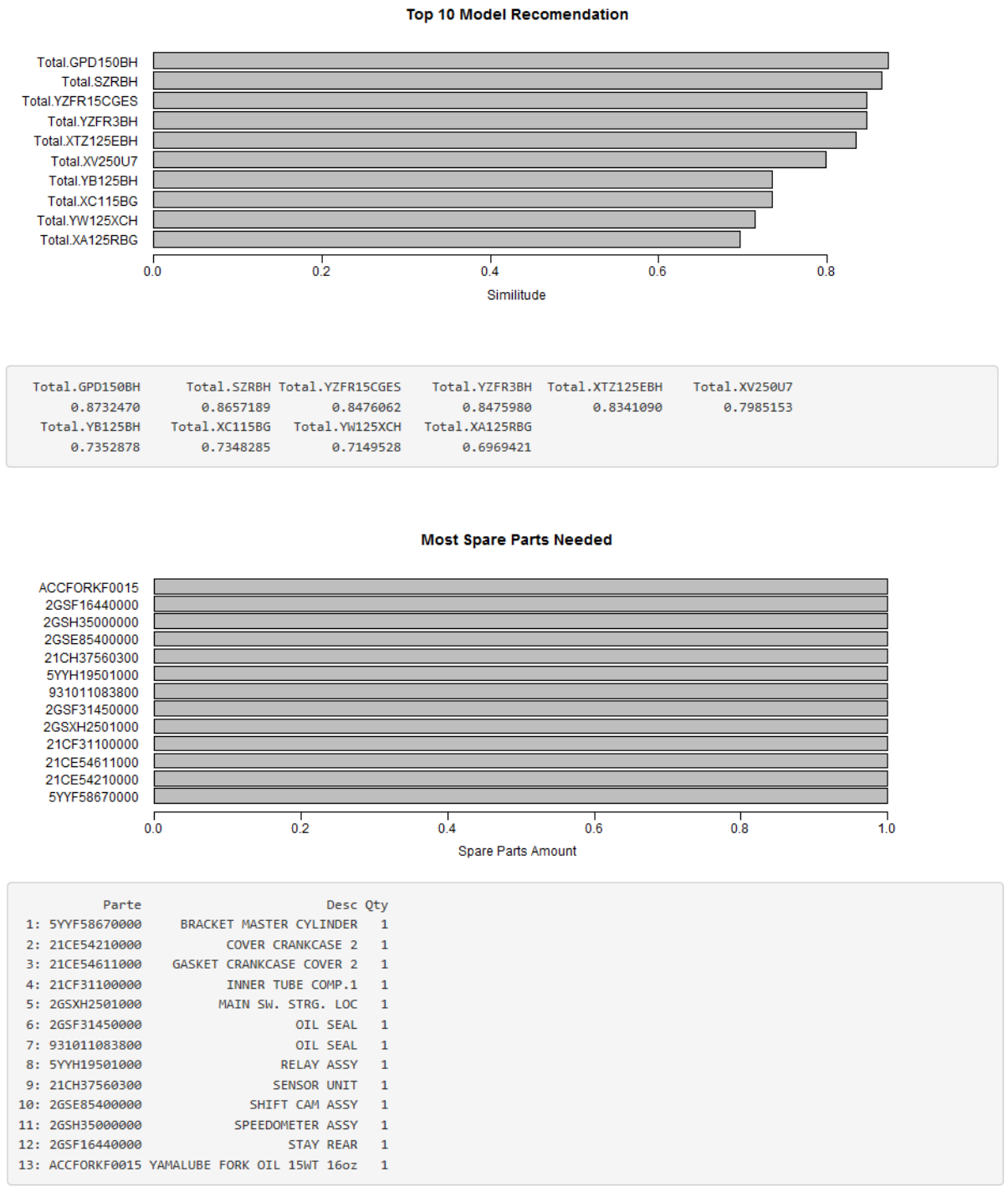


Figura 2-11 Resultado final de Aplicación Shiny

2.5.7. Publicación de resultados en sitios públicos

El desarrollo de la documentación de este proyecto se realizó en Notebook de RStudio y la Aplicación en Shiny como se ha comentado con anterioridad.

Los resultados de este proyecto están publicados en sitios públicos donde se puede ver su comportamiento y/o utilizarlos como referencia para futuros trabajos o estudios. Tome en cuenta que algunos datos fueron anonimizados a solicitud de la empresa por la confidencialidad de su información.

Código de desarrollo de Sistema Recomendador disponible en el sitio público:

http://rpubs.com/Gabriela_Cueva/TFM

Aplicación Shiny para interfaz con el usuario disponible en el sitio público:

<https://gabrielacueva.shinyapps.io/tfmqcm/>

2.6 Resultados

Para la implementación de este piloto experimental se reunió a las personas clave del negocio que utilizarán este piloto como una herramienta para su trabajo.

Al aplicar este piloto Experimental en la empresa Yamaha Motor de México, se hizo una explicación del Piloto, también de la forma de utilizar la aplicación en Shiny, se realizaron algunas pruebas guiadas y después se permitió a los usuarios seguir probando por su cuenta.

Durante la prueba se recogieron las impresiones y comentarios y al final se les invitó a responder la siguiente encuesta:

Preguntas
¿Le parece útil el Sistema?
¿Cree que sea una herramienta que facilite la labor de venta?
¿Le parece fácil de usar?
¿Considera que la información que entrega es útil?
¿Le es sencillo interpretar los resultados de la consulta?
¿Cree que estas recomendaciones puedan inducir al distribuidor a la compra?
De aplicarse en su labor de ventas ¿Considera que mejora su seguridad al recomendar productos al distribuidor?
¿Piensa que cambiaría positivamente su modo de trabajo?
Del 1 al 5 ¿qué calificación le daría a esta herramienta?

Tabla 6 Preguntas realizada en la encuesta de satisfacción

Esta información se obtuvo a través de una encuesta en línea con survey monkey realizada a **2** asesores de ventas de motocicletas, **2** asesores de ventas de refacciones y **2** representantes del área de TI que evaluaron el piloto experimental como software de apoyo desde su área de experiencia y conocimiento de negocio, siendo estas áreas las más directamente interesadas en el caso que estudia este trabajo.

Se habría querido evaluarlo con más personas, pero hay que considerar que los asesores de ventas no son muchos como tampoco los interesados del área de TI, y dado que debe estar siempre alguno en el área para no dejarla descuidada al menos uno de cada área no pudo participar en el piloto, los demás estuvieron presentes.

Los resultados de la encuesta fueron los siguientes:

Resultado de Encuesta

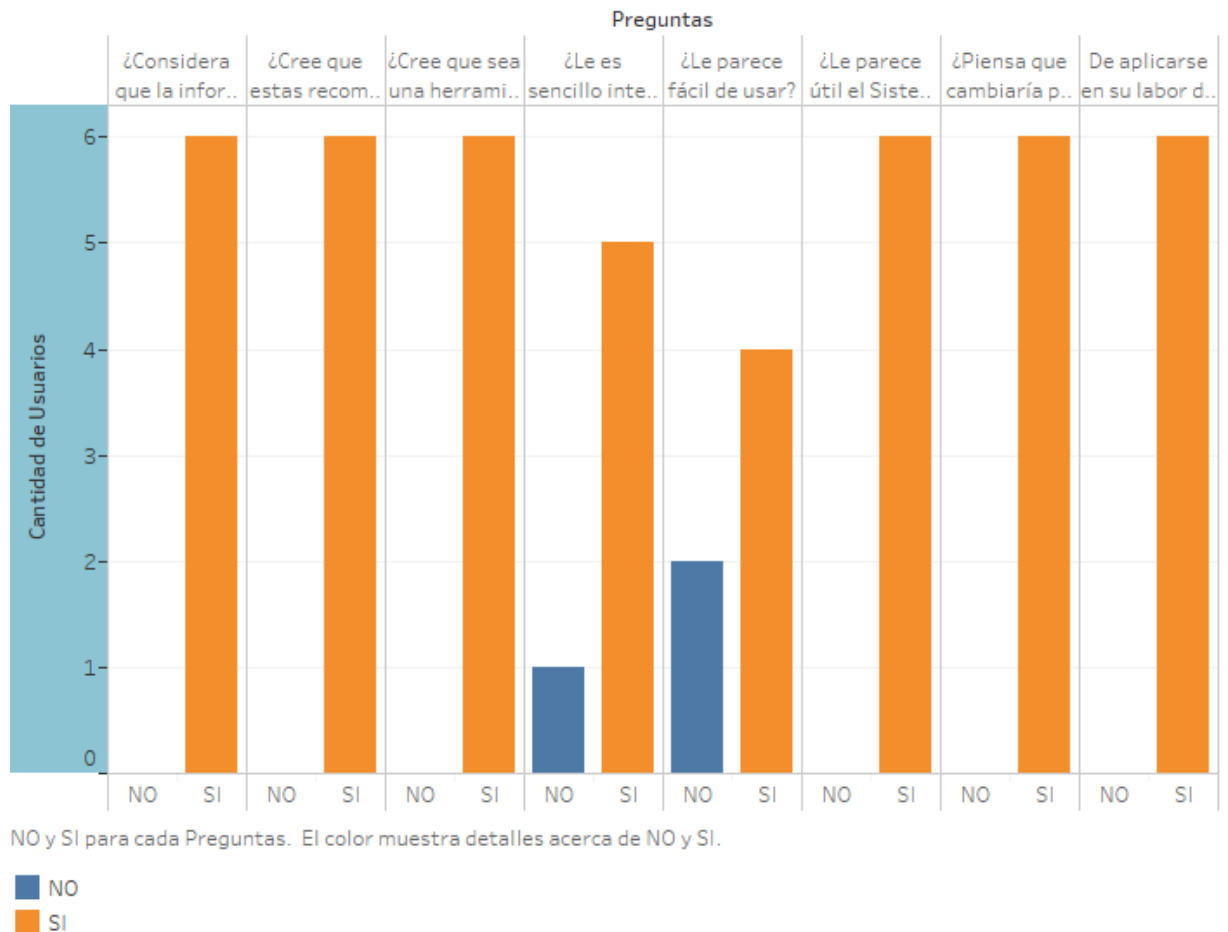


Figura 2-12 Resumen Resultados Encuesta Preguntas 1 a 8



Figura 2-13 Resumen Resultados Encuesta Pregunta 9

2.7 Discusión

En la prueba del Sistema Recomendador el sistema funcionó adecuadamente, pero algunos usuarios se resistieron a su uso porque no realizan ninguna labor de análisis y les pareció un poco complicado de entender al principio, o tal vez por la simple resistencia normal de aprender algo nuevo. Más tarde se mostraron más interesados en la herramienta e incluso sugirieron hacer otro Sistema Recomendador para que el distribuidor pueda consultarlo directamente y le sea personalizado.

Los usuarios que se mostraron más interesados en el sistema desde el principio son los que se dedican a la venta de refacciones porque ellos hacen un análisis manualmente con muy poca información y que además obtienen de manera manual por lo que no los consideran actualizados ni exactos, por esto les pareció una herramienta que facilita su trabajo y que les arroja información útil, más completa, exacta y actualizada que la que ellos utilizan. Ellos planean también pedir autorización a su gerente de área para utilizarla de manera permanente.

El área de TI estuvo siempre interesada desde el planteamiento del piloto y les pareció una herramienta muy útil y un principio para el análisis, visualización y aprovechamiento de sus datos en áreas clave del negocio. Ellos pretenden ir avanzando hacia la transformación

del negocio tomando en cuenta sus datos y también en este sentido les pareció que pueden aprovechar este proyecto como un inicio de este proceso, dado que buscaban despertar el interés que ha despertado en las áreas clave del negocio.

2.8 Conclusiones

Este piloto experimental abrió la puerta a la incorporación de Inteligencia Artificial en Yamaha Motor de México, les aportó información que consideran útil y que desean seguir explorando.

Las diferentes áreas interesadas se vieron motivadas por tener esta herramienta a la mano dado que les facilita su trabajo e incluso lo mejora puesto que muchos de ellos no venían haciendo ninguna recomendación de compra a los distribuidores.

La información que recibieron les ha parecido muy interesante así como también la manera de interactuar con ella y la forma de visualizarla, las áreas clave de la empresa son las más interesadas en el piloto dado que ven la utilidad desde la perspectiva del negocio.

Hasta este punto se ha logrado cumplir con los objetivos del piloto experimental de Brindar información clave para el negocio y Motivar al área de ventas para utilizar este piloto experimental.

Los expertos en el área consideran según los resultados de la encuesta realizada que este piloto logrará incrementar las ventas por lo que basados en su conocimiento y experiencia concluimos que efectivamente se cumplirán los siguientes objetivos planteados:

- Impactar al negocio
 - Aumentando las compras
 - Mejorando los tiempos de respuesta
 - Aumentando la satisfacción de los clientes
 - Aumentando el mercado

Resulta un poco complicado empatar los tiempos de este trabajo con los de la empresa y dado que las autorizaciones requieren tiempo sólo obtuvimos el permiso de implantar este piloto en la parte de desarrollo de la empresa y no en la parte productiva.

Aún falta presentarlo a los Gerentes comerciales y de mercadotecnia que hasta el momento no han podido recibirnos, y es necesario replantear los planes estratégicos de la empresa para lograr la implementación de este piloto en productivo y como parte de una estrategia de negocio que impacta prácticamente todas las áreas clave del negocio, por lo que ahora resta esperar a ver si se aplica definitivamente en la empresa y si a mediano plazo se ve reflejado desde su perspectiva financiera.

Estoy segura, al igual que los expertos en el negocio y por lo que indica la experiencia en otros negocios incluso del mismo ramo, que han implementado diversas técnicas de Inteligencia Artificial y diversos tipos de Sistemas Recomendadores en su negocio, que siendo un cambio tan favorable y radical dado que no se hacía ninguna clase de análisis sobre sus datos, y basados en estas mismas tecnologías, seguramente la implementación en productivo de este piloto experimental en la empresa, tendrá un impacto no solamente en su plan estratégico de negocio que replanteará las áreas clave de su negocio, sino también en sus ventas a distribuidores, sus ventas a clientes finales e incluso en el incremento de mercado que han venido buscando por varios años.

2.9 Líneas Futuras de Investigación

Sería muy interesante desarrollar un Sistema Recomendador para el usuario final que visita la página web de Yamaha y enlazarlo con este sistema para recomendar productos y publicidad de los productos que el usuario final está consultando.

También sería muy bueno hacer otros Sistemas Recomendadores complementarios a éste que consideren la ubicación de los distribuidores para hacer la recomendación, incluso podría mostrarse directamente al distribuidor como sugerían los usuarios.

Otra línea de investigación que me parece interesante desarrollar más adelante es la de las redes sociales, actualmente la compañía no tiene información a este respecto, pero si la desarrollaran a partir de una campaña de marketing estratégico, sería muy interesante alimentar sus Sistemas de Recomendación con las preferencias de los usuarios expresadas en la red, hacia este rumbo tienden los Sistemas Recoemendadores actualmente por lo que sería muy interesante aplicarlo a las necesidades de Yamaha Motor de México también.

También pienso que se podría hacer un sistema que detalle el tiempo que dura en promedio las refacciones para un modelo dado calculándolo de la fecha de compra a la fecha de falla por la aplicación de una garantía o por la realización de un servicio. De este modo se podría tener también un análisis del rendimiento de las partes.

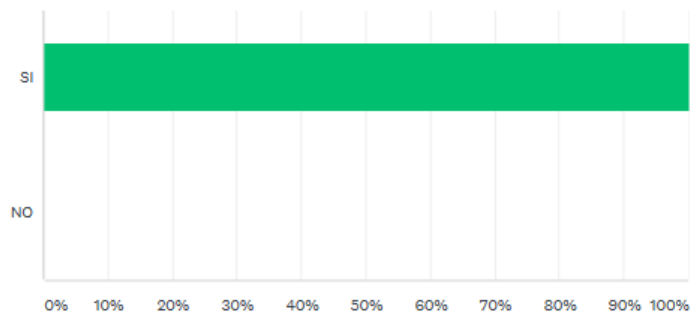
Siendo que la empresa no hacía ningún análisis de sus datos las líneas futuras de investigación que quedan por desarrollar son muchas y muy variadas.

C. Anexos

Resultados de la Encuesta obtenidos de Survey Monkey a detalle para cada pregunta realizada.

¿Le parece útil el Sistema?

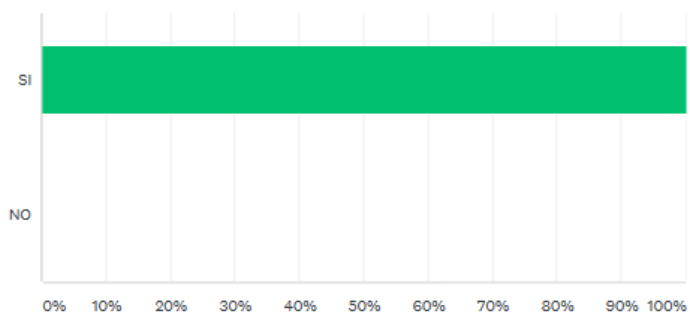
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES
SI	100.00% 6
NO	0.00% 0
Total	6

¿Cree que sea una herramienta que facilite la labor de venta?

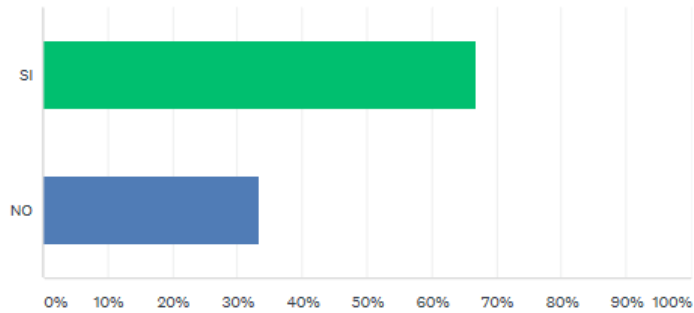
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES
SI	100.00% 6
NO	0.00% 0
Total	6

¿Le parece fácil de usar?

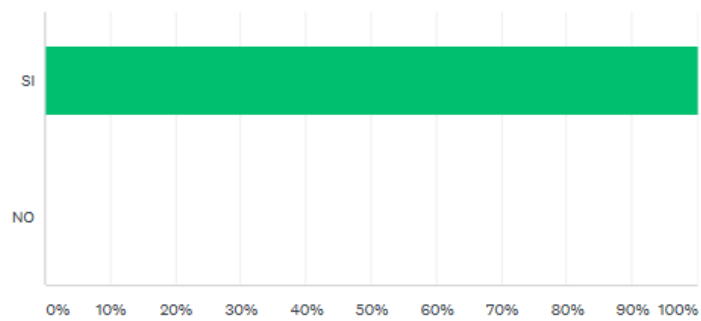
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES	
SI	66.67%	4
NO	33.33%	2
Total		6

¿Considera que la información que entrega es útil?

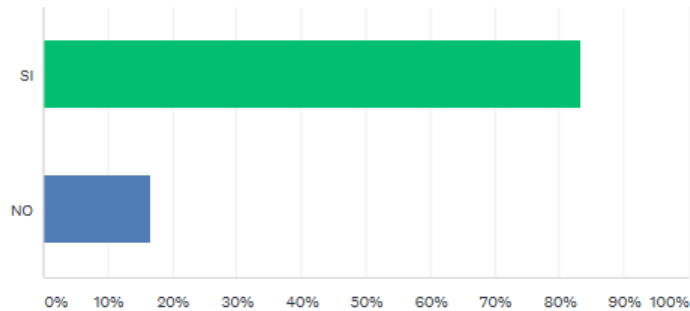
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES	
SI	100.00%	6
NO	0.00%	0
Total		6

¿Le es sencillo interpretar los resultados de la consulta?

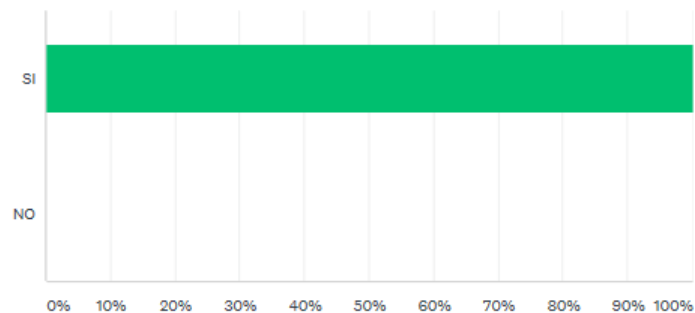
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES	
SI	83.33%	5
NO	16.67%	1
Total		6

¿Cree que estas recomendaciones puedan inducir al distribuidor a la compra?

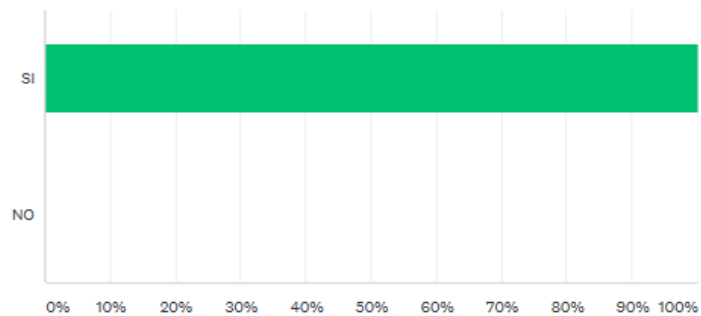
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES	
SI	100.00%	6
NO	0.00%	0
Total		6

De aplicarse en su labor de ventas ¿Considera que mejora su seguridad al recomendar productos al distribuidor?

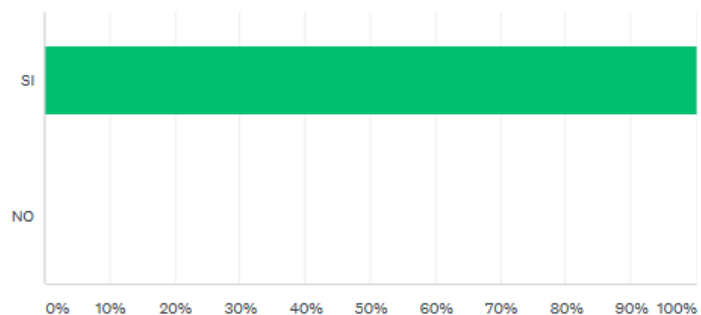
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES
SI	100.00% 6
NO	0.00% 0
Total	6

¿Piensa que cambiaría positivamente su modo de trabajo?

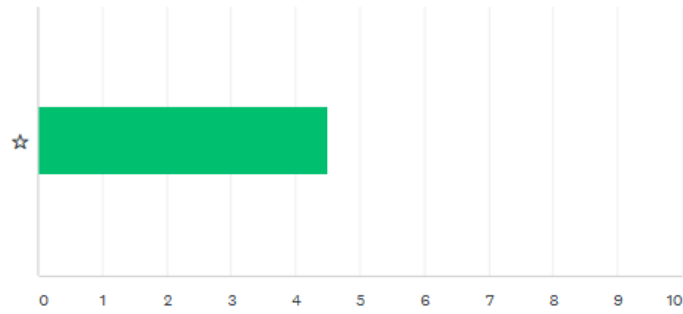
Answered: 6 Skipped: 0



ANSWER CHOICES	RESPONSES
SI	100.00% 6
NO	0.00% 0
Total	6

¿Qué calificación le daría a esta herramienta?

Answered: 6 Skipped: 0



	1	2	3	4	5	TOTAL	WEIGHTED AVERAGE
☆	0.00% 0	0.00% 0	0.00% 0	50.00% 3	50.00% 3	6	4.50

Figura C-1 Figuras a detalle de resultados de evaluación

D. Bibliografía

Libros

1. Charu C. Aggarwal (2016). *Recommender Systems: The Textbook*. Ed. Springer International Publishing Switzerland.
2. Pasquale Lops, Marco de Gemmis and Giovanni Semeraro (2011). *Recommender systems handbook*. Ed. Springer International Publishing Switzerland.
3. Francesco Ricci · Lior Rokach · Bracha Shapira · Paul B. Kantor (2011). *Recommender Systems Handbook*. Ed. Springer Science+Business Media (www.springer.com).
4. J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen (2007) *Lecture Notes in Computer Science Conference Paper. Collaborative Filtering Recommender Systems*.
5. Jannach, D., Zanker, M., Felfernig, A. & Friedrich, G. (2011). *Recommender Systems: An Introduction*. Cambridge University Press.
6. Jordi Conesa Caralt (Coord.) Josep Curto Díaz (2010). *Introducción al Business Intelligence*. Ed. UOC.
https://books.google.com.mx/books?hl=es&lr=&id=iU3RAXYQXMkC&oi=fnd&pg=PA140&dq=business+intelligence+definici%C3%B3n&ots=Nngijv73ds&sig=LJfgv4_9suFymzzWQBONZ22OhSc&redir_esc=y#v=onepage&q&f=false

Noticias

1. <http://www.elfinanciero.com.mx/empresas/italika-domina-el-60-del-mercado-de-las-motos-en-mexico.html>

Artículos

1. Ricardo Moya (2013). ¿Que son los Sistemas de Recomendación?
<https://jarroba.com/que-son-los-sistemas-de-recomendacion/>
2. Pascual Miguel, Félix José; Chaparro Peláez, Jose Julián y Fumero Reverón, Antonio (2011). *Presente y futuro de los Sistemas Recomendadores en la web 2.0*. Revista: "El Profesional de la Informacion" Biblioteca UNIR:

http://biblioteca.unir.net/documento/presente-y-futuro-de/FETCH-crossref_primary_10_3145_epi_2011_nov_073

3. Brad Power (2017). How Harley-Davidson Used Artificial Intelligence to Increase New York Sales Leads by 2,930%. https://hbr.org/2017/05/how-harley-davidson-used-predictive-analytics-to-increase-new-york-sales-leads-by-2930?utm_campaign=Revue%20newsletter&utm_medium=Newsletter&utm_source=revue
4. Xavier Amatriain and Justin Basilico (Personalization Science and Engineering) (2012). Netflix Recommendations: Beyond the 5 stars (Part 1) <https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429>
5. Estela Gil (Consultado en Julio 2017) La importancia de la comunidad en sistemas de recomendación para eCommerce <https://www.brainsins.com/es/blog/la-importancia-de-la-comunidad-en-sistemas-de-recomendacion-para-ecommerce/109111>
6. JAYANTHI RANJAN (2009) BUSINESS INTELLIGENCE: CONCEPTS, COMPONENTS, TECHNIQUES AND BENEFITS. Institute of Management Technology, Ghaziabad, Uttar Pradesh, India.

Enlaces

1. <http://www.pentaho.com/what-is-business-analytics>
2. <http://www.pentaho.com/customer-360-degree-view#tab-2>
3. <https://www.knowage-suite.com/site/home/>
4. WebFOCUS: <http://www.informationbuilders.es/products>
5. <http://www.spagobi.org/product/>
6. https://en.wikipedia.org/wiki/Cosine_similarity

Bibliografía para el Desarrollo

1. G. Adomavicius and Y. Kwon. (2007) New recommendation techniques for multicriteria rating systems. IEEE Intelligent Systems.
2. <http://shiny.rstudio.com/>
3. <https://github.com/jumanbar/Curso-R>
4. <http://shiny.rstudio.com/articles/build.html> - Template de App Shiny-