

[191033] Desarrollo de un modelo de recomendación de prendas de vestir basado en información local

Ricardo Enrique Jara Chacón^{a,c}, María Camila Rodríguez Patiño^{a,c}, Natalia Andrea Romero Arrieta^{a,c}.

Camilo Ernesto Martínez Erazo^{b, c}

^aEstudiante de Ingeniería Industrial

^bProfesor, Director del Proyecto de Grado, Departamento de Ingeniería Industrial

^cPontificia Universidad Javeriana, Bogotá, Colombia

Engineering Design Summary.

It is an overwhelming experience for the client to find content of his interest within a wide range of existing options, therefore, companies have seen the need to implement new methodologies that allow them to suggest products that fit their customers' preferences. However, in the fashion industry it is a challenge to suggest products of interest, considering the large number of attributes involved in the garments and the subjectivity factor present at the time of making a purchase. The problem will be addressed by developing a garment recommendation system for a local clothing brand. The objective of this research is to ease decision making, improve the customer perception of the store's experience and allow the company to meet the needs of the market. Consequently, the recommendation system considers three different moments during the shopping experience in the local clothing brand, the vendor will suggest garments depending on the point of the route where the client is: welcome, checkout line, register.

In defining the rules for each moment, on one hand, important factors were considered for the brand, such as inventory turnover and promotions. On the other hand, tools such as data analysis, Pareto analysis, RFM analysis, market segmentation and association rules, were used to give an idea of the data and transform it into usable actions for system design. Python was used to determine the association rules, the recommendations selected were those that met established criteria for its confidence, support, and lift value. In order that the seller can clearly visualize the recommendations, an application was developed in Visual Basic language, which receives the SQL database as input and executes the designed recommendation system. The application was developed in batch considering the standards of ISO 25010 and ISO 25012. This generates recommendations in less than 3 seconds and does not generate product recommendations that are not available. Furthermore, this is for the exclusive use of the brand, considering that the recommendations generated respond to past behavior patterns of their customers. The application does not consider the size of the garments to make recommendations and, in case of requiring updates, the input parameters must be modified.

Finally, to evaluate the results, on the one hand, tests were conducted that allowed the measuring the quality of the association rules and, on the other hand, a focus group was made with some of the brand's sellers, to assess the perception of the mark with respect to the designed application. The results were analyzed to extract relevant information and identify opportunities for improvement, according to the suggestions received..

1. Justificación y planteamiento del problema.

Hoy en día se presentan situaciones donde la toma de decisiones sencillas como escoger una canción, una película, un destino turístico o incluso qué ropa comprar, resulta ser un trabajo complejo cuando se consideran la gran cantidad de opciones disponibles para el usuario. Un ejemplo claro se encuentra en la industria de la moda en la que el alto nivel de subjetividad y la amplia variedad de características involucradas en las prendas como estacionalidad, estilo, material y color, convierten la selección en un gran reto. Ocasionando que los

clientes puedan llevarse una experiencia poco satisfactoria producto de tomar una mala decisión (Kang, Fang, Wang, & McAuley, 2017).

Antes de la era tecnológica, la tendencia general era que los individuos tomaran este tipo de decisiones basadas en opiniones dadas por personas de su alrededor. A raíz de esto, surge el concepto de sistemas de recomendación para hacer sugerencias personalizadas a los usuarios aplicando algoritmos matemáticos (Vaidya & Khachane, 2017). Dichos algoritmos son actualmente implementados en plataformas como Google y Facebook para clasificar contenido de sus resultados de búsqueda y noticias, y por plataformas digitales como Netflix y YouTube quienes los usan para sugerir nuevos videos (Hao, 2019). Las estadísticas de los estudios de McKinsey y Tech Emergence, consultoras especializadas en analítica, revelan que implementar sistemas de recomendación representa hasta el 75% del consumo de video en Netflix y el 60% de las visitas en YouTube (Xing et al., 2018).

De la misma manera, el grupo Alibaba obtuvo 6,7 billones de dólares en un evento de compras en línea tras implementar un sistema de recomendación personalizado (Faggella, 2017). Dicho sistema conocido como *FashionAI* ofrece a los consumidores recomendaciones de prendas personalizadas con el objetivo de fidelizar a los clientes y mejorar la eficiencia operacional (Rebholz, 2019). Así como esta empresa, existe el retail de prendas de vestir Zalando, el cual ha generado un aumento del 15% en las ventas virtuales de los países donde se encuentra, gracias a su sistema de recomendación de estilos en su página web (The Economist, 2016). Otras marcas como Levis, H&M, Victoria's Secret y ASOS, utilizan estos sistemas para mostrar los “recién llegados”, “también compraron”, “promociones”, “vistos recientemente” y “mejor valorados”, mejorando la experiencia del cliente e incrementando sus beneficios (Kottage, Keshawa, Chankuma, Upeksha, & Sandanayake, 2018).

Otro de los objetivos de la implementación de sistemas de recomendación es fidelizar al cliente. En un estudio realizado en Estados Unidos y el Reino Unido se revela que el 73% de 2000 consumidores prefieren experiencias de compra personalizadas (Tamurk, 2017). Los sistemas de recomendación permiten a las empresas aprovechar el conocimiento sobre sus clientes y crear una experiencia personalizada adaptada a las necesidades individuales (Walter, Battiston, Yildirim, & Schweitzer, 2012).

Finalmente, un último objetivo de la implementación de estos sistemas es reducir el volumen de devoluciones. En Estados Unidos se devolvieron 260 billones de dólares en mercancía a los minoristas en 2015, según la Federación Nacional de Minoristas. Cuando el cliente no logra adquirir un producto que llena sus expectativas las compañías se ven enfrentadas a este problema, las tasas de retorno son del 30% o más para la mercancía que se compra en línea y las devoluciones de ropa pueden estar cerca del 40% (Reagan, 2016). En el caso de Colombia el panorama no resulta ser más alentador. El observatorio e-Commerce estima que la categoría de moda es una de las que más devoluciones genera (Observatorio eCommerce, 2019).

Ahora bien, el reto actual es adaptar estos sistemas a campos donde su implementación representa un gran desafío, con el fin de lograr la consecución de los objetivos mencionados. El campo de la moda hace parte de este grupo ya que la mayoría de las marcas que se dedican a la venta exclusiva de prendas de vestir no implementan sistemas de recomendación por la dificultad que representa identificar los gustos de las personas. Es importante abordar este ámbito porque las compras de los elementos que abarcan este campo competen a la mayoría de la población humana.

Cada año, se compran 80 billones de prendas de vestir en todo el mundo, una cifra 400% superior a la de hace trece años (Hargreaves, 2015). Solo en Colombia el consumo de vestuario total en 2016 fue de aproximadamente 13 billones (Potafolio, 2017). Mientras en el 2005 el consumidor promedio compraba siete prendas al año, para el 2016 este número alcanzó las 22. Expertos explican este aumento como el resultado del aumento del poder adquisitivo que a su vez beneficia a la industria textil (Semana, 2017).

Estos resultados surgen como respuesta a la creencia de que la ropa es una de las señales más informativas de la apariencia humana. Es por esto por lo que es indispensable que las tiendas que ofertan artículos de moda procuren resaltar sus diferencias y se cuestionen acerca de qué le van a brindar al consumidor además de un producto con calidad y un buen diseño. Ahora bien, ¿Cómo se pueden abordar las temáticas expuestas anteriormente de manera práctica en un retail de moda?

Una de las posibles respuestas es la utilización de estrategias comerciales como el cross-selling. La utilizan los vendedores para crear relaciones duraderas entre el cliente y la empresa, obtener una mayor valoración del cliente, reducir la rotación y aumentar el número de clientes leales (Akçura & Srinivasan, 2005). Consiste en la práctica de vender un producto o servicio adicional a un cliente existente y se conoce como una de las principales prioridades estratégicas para muchas industrias (Johnson, Jeff & Friend, 2015). Cabe destacar que los sistemas de recomendación sirven como una base estructurada para el desarrollo de esta técnica.

El retail de moda en el que se quiere implementar esta estrategia es una marca local de la que, por cuestiones de confidencialidad no será mencionado su nombre. Para esta marca es importante encontrarse a la vanguardia de las tendencias del mercado con el fin de cumplir con su estrategia organizacional, lo que se evidencia en sus principios organizacionales, dentro de los que se encuentran:

“La rapidez para reaccionar a los cambios de la moda, la capacidad para adaptarse a cada nuevo mercado, la habilidad para gestionar las relaciones con los clientes, conocerlos y adelantarse a sus necesidades, y mantenerse en la cresta de la ola tecnológica en la gestión de la empresa, las tiendas y los medios virtuales”.

Otro de los objetivos que persigue esta compañía es la búsqueda de una ventaja competitiva que le permita aumentar la participación en ventas a través de un factor diferenciador. La participación en ventas de esta compañía en relación con sus principales competidores fue de un 2,7% durante el año 2018 (Orbis, 2018).

Finalmente, con el fin de interpretar el comportamiento de consumo de sus clientes, se realizó un análisis sobre la base de datos recopilada hasta el mes de marzo 2019, donde se identificó que el 53% de los usuarios registrados realizó tres o menos compras desde la apertura de los locales hasta la fecha. Este resultado permite deducir que un alto porcentaje de los clientes no se encuentra fidelizado con la marca y, además, representa la oportunidad de implementar estrategias que permitan aumentar el volumen de compra de los clientes.

De estos planteamientos surge el cuestionamiento: ¿Cómo lograr identificar las preferencias de los usuarios de esta marca para llegar a cumplir las expectativas de los clientes y permitir que la organización cumpla con sus objetivos estratégicos? De lograr resolver esta pregunta se podrían aumentar las ventas, aumentaría la satisfacción de los clientes y con ello, su experiencia comprando la cual mejoraría significativamente.

2. Antecedentes.

Los sistemas de recomendación surgieron a mediados de la década de los 90 con el objetivo de brindar a los usuarios resultados de búsqueda de información cercanos a sus necesidades (Rodríguez, Duque, Ovalle, Peluffo, & Salazar, 2014). Formalmente, los sistemas de recomendación son herramientas de software que se valen de técnicas de filtrado para entregar a los usuarios sugerencias sobre ítems, medios o información de su interés (Bolívar & Giraldo, 2017). Existen varios tipos de sistemas de recomendación y cada uno puede ser usado en diferentes contextos de acuerdo con las necesidades (Rodríguez et al., 2014).

Las recomendaciones generalmente se hacen por medio de tres métodos. El primero de ellos se conoce como filtro de contenido. Este calcula la similitud entre elementos y recomienda los relacionados con aquellos en los

que el usuario ha expresado interés (Golbeck, 2006). En él se usan palabras clave que describen elementos y se construye un perfil de usuario que identifique el tipo de estos que le gustan (Omran, 2018).

Algunas de las aplicaciones de este método se ven reflejadas en el área educativa, donde han sido creados para la recomendación de patrones de diseño en recursos educativos (Terry González, Estrada Sentí, & Arteaga Gómez, 2016). También han sido implementados en recomendación de música para ayudar a los usuarios a descubrir sus preferencias musicales a través de calificaciones (Ja-Hwung, Wei-Yi, & Tseng, 2017) y en la recomendación de contenido de televisión (Gutta et al., 2000).

El segundo método de recomendación es conocido como enfoque de filtro colaborativo. Este construye un modelo a partir del comportamiento pasado del usuario, así como decisiones similares tomadas por otros usuarios (Linden, Smith, & York, 2003). Lo anterior se basa en la suposición de que las personas que estuvieron de acuerdo en el pasado estarán de acuerdo en el futuro y, por lo tanto, les gustaran ítems parecidos (Mukherjee & Herbrich, 2016).

Los sistemas de recomendación por filtros colaborativos son importantes para servicios como Amazon, Netflix y otros proveedores en línea similares. Por un lado, Amazon cuenta con un modelo de recomendación conocido como *Filtro colaborativo ítem a ítem*. A diferencia de otros, este algoritmo se escala independientemente del número de clientes o artículos, y genera recomendaciones en tiempo real de alta calidad (Linden et al., 2003). Por otro lado, Netflix usa un sistema conocido como SVD, que toma en cuenta las calificaciones que proporcionan los usuarios para encontrar relaciones entre grupos, agregando sesgos como la calidad, serendipia, y comparación de la calificación con el promedio general (Gower, 2014).

El tercer método de recomendación es conocido como asociación de reglas de minería. Es una técnica de minería de datos que identifica la relación entre varios elementos en un conjunto de datos (Siswanto & Thariqa, 2018). Consiste en averiguar las reglas de asociación que satisfacen *la confianza* y *el soporte mínimo* predefinidos de una base de datos determinada y con ello busca extraer correlaciones interesantes, patrones frecuentes y asociaciones ocasionales entre conjuntos de elementos en esta (Siswanto & Thariqa, 2018).

La asociación de reglas de minería fue usada en un estudio para determinar reglas importantes que podrían ser útiles en la detección de factores de riesgo del cáncer de mama, con el fin de que los médicos tomaran una decisión informada de prevención o de iniciación en la etapa temprana de la progresión de la enfermedad (Kabir, Ludwig, & Abdullah, 2018). Otra aplicación se evidencia en una investigación la cual establecía reglas de asociación que permitieran determinar los ingredientes más utilizados en videos de recetas de cocina de Indonesia que estaban publicados en YouTube (Siswanto & Thariqa, 2018). A su vez, para comprender la estructura y localizar información relevante dentro de una colección de textos, un estudio propuso filtrar reglas de asociación para detectar temas de un grupo de documentos (Lopes, Pinho, Paulovich, & Minghim, 2007).

Las reglas de asociación del tipo $X \rightarrow Y$ donde X y Y son un conjunto de ítems, tienen ciertos criterios de medición que permiten su evaluación. A continuación, se explican algunas de las más relevantes.

Support. Indica la frecuencia con la que aparece un elemento en todas las transacciones de los registros del conjunto de datos. Corresponde a la proporción de transacciones que contienen tanto a X como a Y , y se puede considerar como la probabilidad conjunta de X y Y , o sea $P(X \cap Y)$. Una regla con *support* S se interpreta como una regla que se cumple en $S\%$ de las transacciones (Nahm & Mooney, 2000).

Confidence. El valor de confianza indica qué tan fuerte es la relación entre dos elementos en una regla. Corresponde a la probabilidad de que Y aparezca en una transacción dado que X ya ocurrió, y se puede considerar como la probabilidad condicional de Y dado X , o sea $P(Y|X)$. Una regla con *confidence* C se interpreta como: si X ocurrió, entonces también estará Y en un $C\%$ de los casos (Nahm & Mooney, 2000).

Lift. Es la relación de confianza y probabilidad **condicional**. Un lift superior a 1, muestra que X y Y son dependientes una de otra. Se calcula de la forma $\frac{supp(XUY)}{supp(X) \times supp(Y)}$ (Ahmadon & Yamaguchi, 2018).

Además de evaluar estas métricas otro factor importante a considerar es que las recomendaciones generadas por el sistema deben ser no solo desconocidas, sino también sorprendentes para los usuarios, a esto se le conoce como *serendipia* (Mohamed, Khafagy, & Ibrahim, 2019). Este generalmente se encuentra acompañado por un factor de *diversidad* de elementos que puede determinarse mediante la relación del número de ítems diferentes recomendados sobre los disponibles.

Así como la implementación de sistemas de recomendación tiene muchas ventajas, requiere abordar distintas dificultades que se pueden presentar en el desarrollo de cada método. En la tabla 1 se presentan algunos de los retos que conlleva desarrollar estos algoritmos.

Tabla 1.

Retos de la implementación de sistemas de recomendación.

Problemas	Descripción	Sistemas en los que aplican
Cold start problem	Surge principalmente cuando se tiene un nuevo usuario o cuando se añade un nuevo ítem al sistema. Esto ocasiona que no se conozcan los intereses del usuario o que no se pueda recomendar un nuevo producto (Mohamed, Khafagy, & Ibrahim, 2019).	Filtrado por contenido Filtrado colaborativo
Scalability	La escalabilidad mide la capacidad del sistema de trabajar eficazmente y con precisión mientras crece la cantidad de información (Ma & Uchyigit, 2008).	Filtrado por contenido Filtrado colaborativo Reglas de asociación
Sparcity	Este problema se genera cuando los usuarios no proporcionan una calificación, en consecuencia se genera una escasez de información que dificulta la recomendación (Neamah & El-Ameer, 2018).	Filtrado colaborativo
Big data problem	El sistema de recomendación debe ser capaz de funcionar con la gran variedad y volumen de datos (Jariha & Jain, 2018).	Filtrado por contenido Filtrado colaborativo Reglas de asociación
Over-fitting	Los usuarios están restringidos a obtener recomendaciones similares a las que ya se conocen o están definidas, de esta manera se está evitando que el usuario descubra nuevas opciones (Sharma & Gera, 2013).	Filtrado por contenido
Ranking of recommendation	Consiste en identificar las variables estratégicas que permitan clasificar la recomendación obtenida (Jariha & Jain, 2018).	Filtrado por contenido Filtrado colaborativo Reglas de asociación
Gray sheep	Se produce cuando las opiniones de un usuario no coinciden con ningún grupo y en consecuencia no se logra obtener una recomendación (Mohamed, Khafagy, & Ibrahim, 2019).	Filtrado colaborativo

Fuente. Autoría propia.

Para mejorar el rendimiento de los métodos y mitigar las consecuencias que tiene no abordar adecuadamente los retos anteriores existe un hilo común de investigación orientado a combinar múltiples técnicas de recomendación para lograr cierta sinergia entre ellas, a este se le conoce como sistemas de recomendación híbridos (Burke, 2002).

Contextualizando el uso de los sistemas de recomendación en el tema de interés del proyecto, se encontró que en el campo de la moda crear sistemas de recomendación efectivos es un desafío (Kang et al., 2017). Diferentes autores han abordado este campo desarrollando modelos como los presentados a continuación.

Una empresa de moda coreana desarrolló un algoritmo denominado K-RecSys, el cual amplía el algoritmo de filtrado colaborativo, combinando la información generada en las páginas web de la compañía y sus datos de ventas ponderados para reflejar las preferencias de los clientes. Para su implementación se tuvieron en cuenta

los siguientes aspectos: las tiendas virtuales y físicas de la compañía venden los mismos productos, los productos suelen ser estacionales y los clientes suelen comprar artículos para reemplazar los previamente preferidos o comprar artículos para complementar aquellos que ya compraron. También, adopta una función de reducción de preferencias para reflejar los cambios en estas a lo largo del tiempo, y finalmente, recomienda productos sustitutos y complementarios (Hwangbo, Kim, & Cha, 2018).

A su vez, Zalando enfoca sus esfuerzos en recoger información de sus clientes para identificar tendencias, uso de anuncios dirigidos y adaptar sus respuestas a los cambios climáticos o los gustos de la moda (The Economist, 2016). De esta forma, logró desarrollar una aplicación para buscar estilos, en la cual el cliente puede satisfacer todas sus necesidades de moda a partir de los consejos ofrecidos por la aplicación (Fry, 2018). Otros ejemplos de modelos aplicados en diferentes enfoques de este campo se enuncian en la siguiente tabla.

Tabla 2.
Otros modelos aplicados al campo de la moda.

Fuente	Modelo	Metodología
(Bodaghi & Homayounvala, 2018)	Sistema de recomendación basado en contenido	Aplicación interactiva que rastrea las interacciones de los usuarios y les asigna un nivel de experiencia basado en la selección que realizan de diferentes prendas de vestir.
(Zeng, Koehl, Wang, & Chen, 2013)	Sistema de recomendación basado en inteligencia artificial	Recomienda la ropa tomando en cuenta las medidas del cuerpo y las percepciones humanas.
(Liu, Gao, Bian, Wang, & Li, 2018)	Sistema de recomendación por filtros colaborativos	Incluye la detección de la zona de uso de la ropa en el cuerpo humano y a partir de esto sugiere la combinación del mejor par.
(Ramesh & Moh, 2018)	Sistema de recomendación basada en eventos	Usa técnicas de procesamiento de imagen para identificar eventos y escenarios a los que los usuarios podrían asistir.
(He, Zhang, & Meng, 2019)	Sistema de recomendación basada en reconocimiento visual	Parten del principio de que la gente suele comprar ropa con estilos similares, por tanto, usó el reconocimiento visual para lograr recomendaciones más precisas.
(Burke, 2007)	Sistema de recomendación por Filtros Colaborativos	Usa un algoritmo de recomendación con filtros colaborativos involucrando atributos como moda, estilo, material, color, calidad y estacionalidad.
(Wakita, Oku, Huang, & Kawagoe, 2015)	Reglas de asociación	Utiliza tanto las reglas de asociación de marcas de ropa como las características de estas, con el fin de proporcionar al usuario una nueva marca de moda que sea similar a su favorita.

Fuente. Autoría propia.

Para el desarrollo del proyecto se cuenta con información histórica de las ventas realizadas desde 2017 hasta 2019 en 10 sucursales de la marca distribuidas a nivel nacional. Para cada transacción se tiene la fecha y el local, las referencias de los productos comprados junto con su valor comercial, además de información sobre el color, material, temporada, y el departamento y sección a los que pertenecen las prendas, por ejemplo, existe el macro grupo blusa (departamento) que contiene todas las referencias de blusas existentes y existe el subgrupo blusas manga larga (sección).

La información consultada permite conocer los tipos de sistemas de recomendación existentes, el cómo han sido aplicados y evaluar las ventajas y desventajas de usar cada uno de ellos. Teniendo en cuenta esta información y realizando un análisis de la base de datos existente, se determinó que se realizará un modelo de recomendación basado en reglas de asociación, partiendo de dos razones principales: la primera de ellas es que

la dispersión de los datos que se genera es producto de la gran cantidad de referencias existentes en relación con el promedio de artículos comprados en una transacción. Esta imposibilita la aplicación de filtros colaborativos. La segunda, es que no se cuenta con caracterización suficiente de los artículos para poder desarrollar adecuadamente un sistema basado en contenido. Sin embargo, la información sobre los artículos permitirá realizar la categorización de las prendas para así evitar problemas de sinonimia. El sistema de recomendación que se diseñará es de tipo no supervisado ya que se centra en el análisis exploratorio en lugar del análisis predictivo, es decir, se van a encontrar patrones o tendencias sobre un conjunto de variables (Ruchika, Singh, & Sharma, 2017).

Las reglas de asociación que se generen serán evaluadas bajo las métricas *support*, *confidence* y *lift* ya que son las apropiadas para la implementación del tipo de sistema de recomendación escogido para el desarrollo del proyecto. Además de estas se asegurará que exista serendipia en las recomendaciones y se evaluará la diversidad de productos, midiendo la cantidad de elementos distintos que se recomiendan al usuario.

3. Objetivos.

Desarrollar un sistema de recomendación que sugiera ítems a partir de combinaciones de prendas de vestir en un retail de moda

- Desarrollar un sistema de recomendación basado en asociación de reglas de minería que sugiera prendas de vestir relacionadas con las preferencias del cliente y evaluarlo bajo las métricas de *support*, *confidence*, *lift*, *diversidad* y *serendipia*.
- Diseñar un aplicativo en bache que permita visualizar las recomendaciones y con ellas establecer un canal de comunicación entre las recomendaciones del negocio y sus clientes.
- Medir la percepción de la marca sobre el aplicativo y las recomendaciones generadas.

4. Metodología.

La metodología Cross-Industry Standard Process for Data Mining más conocida como CRISP-DM, define un proyecto como un proceso cíclico donde se usan varias iteraciones para asegurar que los datos estén en la forma correcta y así, cumplir con los objetivos del negocio y modelación del proyecto (Moro, Laureano, & Cortez, 2011). De esta manera, orienta trabajos de minería de datos al ofrecer una visión estructurada y robusta del ciclo de vida de un dato para su planificación y ejecución. Dicho ciclo se desglosa en 6 fases: Comprensión del problema, comprensión de los datos, preparación de los datos, modelado, evaluación e implementación (Wirth, 2000).

Para consolidar el proyecto se desarrollaron las fases de esta metodología. En la fase I: comprensión del problema, fueron establecidos los objetivos del proyecto, su alcance, evaluación de la situación y en términos generales se determinó el plan del proyecto.

4.1. Comprensión y preparación de los datos.

El retail de ropa actualmente cuenta con 10 franquicias en Colombia, para el análisis se tomaron las ventas de cada local desde su apertura hasta el 30 de junio del 2019, esto se hizo extrayendo del software ICGManager los registros. Las franquicias no tienen la misma participación en ventas debido a que algunas fechas de apertura son muy recientes, el local V59 abrió sus puertas el 3 de mayo del 2019 y tiene sobre el total de transacciones una participación del 1,09%. El B32, local más antiguo, abierto el 5 de septiembre del 2017, tiene una participación del 19,29% sobre las transacciones totales. Aunque los locales no tengan la misma participación en ventas, la distribución de los productos en estas no presenta un patrón diferenciador para ningún local en

relación con los otros como se puede evidenciar en la figura 1. Por esta razón, se decidió tomar todas las compras como un mismo conjunto de datos.

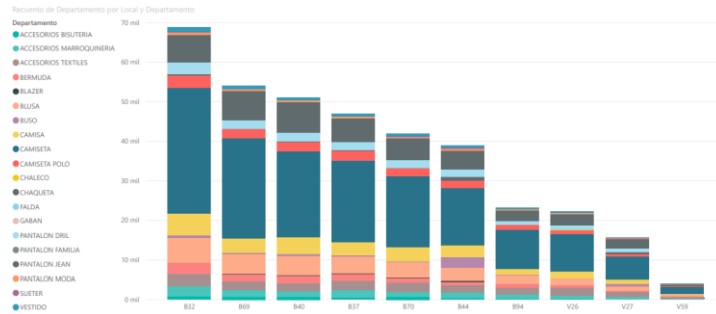


Figura 1. Proporción de ventas por departamento en cada local.

Al recolectar la información de las ventas en todos los locales se consolidó la base de datos, esta cuenta con un total de 190.031 transacciones y 11 atributos correspondientes a el indicativo de factura, fecha, referencia del ítem, descripción, departamento, sección, precio, temporada, género y características físicas como color y material.

Para la limpieza de datos fueron eliminados registros con inconsistencias, faltantes y atípicos, estos correspondieron a venta de bolsas, costos de envío y devoluciones, además de aquellos que no tenían los campos referencia, descripción y características físicas. El volumen de registros eliminados frente al total fue de 28,80%, dejando un total de 187.660 transacciones, un 1,25% menos con relación a las iniciales.

Continuando con el preprocesamiento de datos, se efectuó una transformación en el atributo color. En la base de datos el atributo color contaba con un código que representaba determinado color, al identificar el color asociado a cada código se evidenció que códigos diferentes representaban el mismo color en varias ocasiones, por ello se le asignó un único código a cada color, reduciendo el número de colores diferentes de 86 a 15.

Finalmente, analizando la base de datos se identificó que las referencias de este retail duran en promedio 120,30 días en exhibición. Adicional, al realizar una prueba de independencia chi cuadrado en el software R se pudo identificar, rechazando la hipótesis alterna, que con una significancia del 5% las prendas tienen estacionalidad lo que significa que la venta por departamentos tiene relación con la temporada de las prendas. Las hipótesis planteadas se pueden apreciar a continuación.

H_0 : Hay independencia entre las ventas por departamentos y la temporalidad.

H_1 : No hay independencia entre las ventas por departamentos y la temporalidad.

A pesar de tener estacionalidad, se realizó el proyecto asumiendo que las prendas no presentaban temporalidad, esto representa una limitación para el proyecto, teniendo en cuenta que se disminuye la probabilidad de acierto en algunas ocasiones.

4.2. Modelado.

Durante la experiencia de compra existen varios momentos donde es oportuno recomendar artículos con el fin de mejorar las ventas del negocio. Para cada uno de estos momentos se cuenta con cierto conocimiento del cliente y sus gustos, por lo que es necesario definir diferentes metodologías para ser más precisos en las recomendaciones, estas se presentan a continuación.

4.2.1. Reglas de bienvenida.

Tabla 3.

Información reglas de bienvenida.

Estado del cliente:	Exploración.
Número de productos seleccionados:	Ninguno.
Otros datos:	Género.

Fuente. Autoría propia.

En esta etapa el cliente llega a la tienda con una intención clara o no de lo que necesita. Dado el conocimiento mínimo del cliente se hace uso del conocimiento agregado que se tiene de los muchos clientes que ya han visitado la tienda o alguna necesidad de inventarios del negocio.

4.2.1.1. Reglas de selección productos estrella

El producto estrella es aquel que la mayoría de los clientes asocian más fuertemente con la marca, y así mismo es aquel que tiene mayor impacto sobre la rentabilidad de la empresa (Magne, Oivind, & Leif, 2004) Para saber si un producto es de este tipo, es necesario identificar aquellos productos que representan el mayor volumen de ventas o aquellos que generan las mayores ganancias. Como el retail obtiene el mismo porcentaje de ganancia por cualquier producto que venda, se decidió entonces, utilizar como técnica el análisis de ventas globales y por categorías aplicando el principio de Pareto que permite separar los aspectos significativos de aquellos triviales (Walid, 2018).

Para hacer estas recomendaciones es necesario evaluar las unidades vendidas de las referencias que están a la venta en el momento, por ello no se puede usar la base de datos histórica, esta se dividió en una ventana de tiempo ajustada a la temporalidad que tienen los productos en promedio, correspondiente a cuatro meses. Se tomaron las ventas realizadas del 1 de marzo del 2019 al 30 de junio del 2019 y con ellas se obtuvieron las referencias estrella que debía sugerir el aplicativo según los intereses del cliente. En la figura 2 se puede apreciar el proceso que se realizó para definir las referencias estrella.

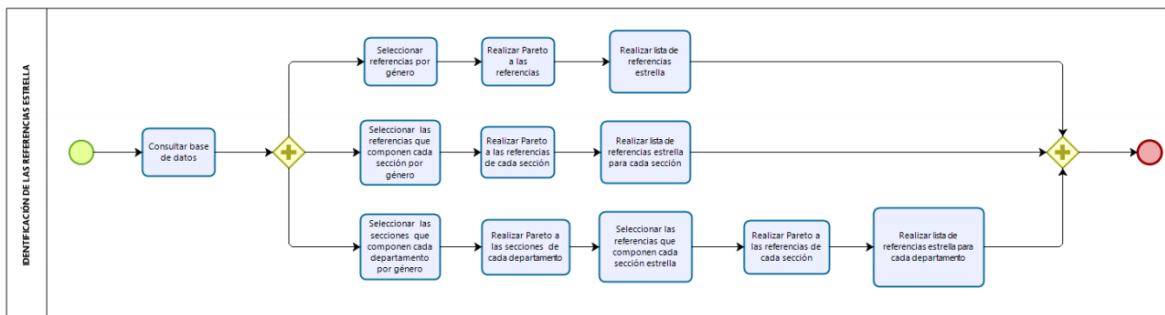


Figura 2. Proceso de definición de referencias estrella.

El proceso fue realizado para las ventas por departamento, sección y referencia con el fin de reflejar para cada uno de estos el 20% de las referencias que representan el 80% de las ventas y así determinar las referencias estrella para cada categoría. Al usuario finalmente se le recomendarán, según su género, referencias que hacen parte de ese 20%. A continuación, se puede observar que al identificar las referencias estrella de la sección camiseta manga sisa para el género femenino, según la figura 3, de 13 referencias de camisetas que componen esta sección, seis de ellas representan el 80% de las ventas de esta y por ello estas serán las recomendadas como productos estrella de esta sección.

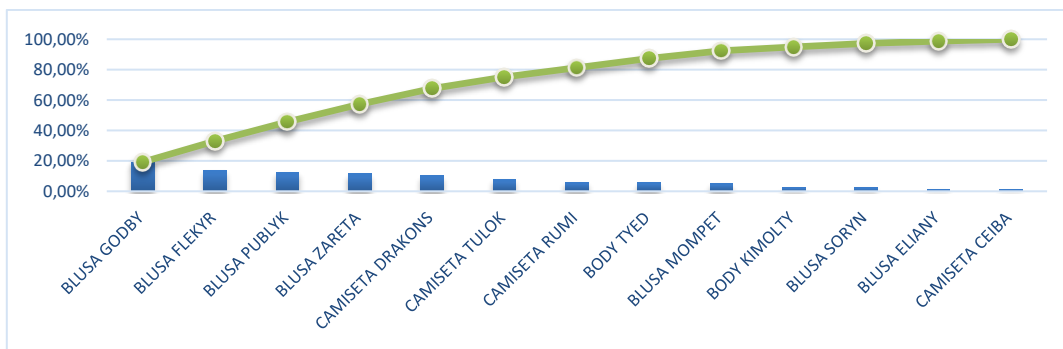


Figura 3. Pareto referencias estrella de la sección camiseta manga sisa.

4.2.1.2. Inventario

Muchas empresas han optado por incluir los niveles de rotación de inventarios como parte de sus estrategias para mitigar concentraciones de ventas en departamentos específicos, esto les ha permitido reducir los costos operativos por acumulación de inventarios, ampliar la venta de su portafolio de productos y reducir la probabilidad de una escasez futura, factor importante pues el usuario no estará satisfecho si el producto que desea no se encuentra disponible (Demirezen & Kumar, 2016). En la marca local se presenta el problema de concentración de ventas por departamentos, esto se evidencia en el análisis estadístico, donde se extrajo que el 72,33% de las ventas se encuentran concentradas en cuatro departamentos. Para aminorar esta problemática, se optó por añadir una estrategia de inventarios.

Se considera que una referencia tiene alto inventario si se tiene de esta una cantidad de unidades en el almacén mayor a los niveles aceptados, por ello, al usuario, sin importar lo que desee llevar se le recomendarán según su género aleatoriamente referencias que tienen el valor máximo de unidades hasta menos cinco, de tal forma que, si el valor máximo de unidades en inventario de una referencia es 30, se le recomendarán referencias con stock entre 25 y 30 unidades. Para el desarrollo del proyecto no se contaba con información sobre el número de unidades en inventario de cada referencia, por ello se le asignó a cada una un número aleatorio entre 0 y 30.

4.2.2. Reglas independientes de los productos.

Tabla 4.

Información reglas independientes de los productos.

Estado del cliente:	Con intención de compra.
Número de productos seleccionados:	Independiente de los productos.
Otros datos:	Comportamiento histórico del negocio.

Fuente. Autoría propia.

En esta etapa el cliente ya ha seleccionado al menos un producto y se encuentra realizando la fila de pago. Dada la información histórica de las ventas del negocio, se realizó un análisis del conjunto de secciones que podrían ser propensas a impulsar en fila con el fin de realizar cross selling.

La personalización de las ofertas y la venta de complementos son consideradas como las principales fuentes para convertir una transacción en única y mantener una relación duradera con el cliente (Johnson & Friend, 2015). El cross selling se ha convertido en una práctica frecuente entre las empresas porque permite obtener mayores beneficios, atraer nuevos clientes, mejorar las relaciones con clientes antiguos, reducir los costos y generar una diferenciación competitiva (Bernard & Pawel, 2010). Las estadísticas indican que las tasas de respuesta para el cross selling son de dos a cinco veces mayores que las ventas en frío (Kamakura, 2008).

4.2.2.1. Análisis RFM.

Para definir las secciones a impulsar en fila se utilizó el análisis *Recency, Frequency, Monetary* conocido por sus siglas en inglés como RFM (Chan, 2008). Esta técnica utiliza información histórica de los clientes para predecir comportamientos, bajo el supuesto de que los patrones futuros de los consumidores se parecen a los pasados y actuales, por ello, para el análisis fue necesario utilizar toda la información recolectada. El enfoque tradicional se orienta a identificar el comportamiento de los clientes, sin embargo, para el tema de estudio se ha optado por enfocar esta estrategia hacia el producto, basando las categorizaciones en el método sugerido por (IBM, 2017).

Se modela desde tres dimensiones, lo reciente, lo frecuente y lo monetario. La primera dimensión, indica el tiempo promedio transcurrido entre la compra de una misma sección, teniendo en cuenta que las adquisiciones recientes tienen más probabilidad de volver a ser adquiridas. Se clasificaron los productos en cinco categorías, teniendo en cuenta el promedio de días que pasan desde que se compra una sección, hasta que vuelve a ser comprada. De tal forma que se clasificaron los datos por quintiles como se puede apreciar en la tabla 5. Se le asignó el valor de cinco a las secciones con menor promedio de días hasta que vuelven a ser vendidas y uno a aquellas que tienen un mayor promedio.

Tabla 5.
Valores máximos y mínimos para cada quintil de la recencia.

	Recency				
Quintiles	5	4	3	2	1
Mínimo	1,01	1,15	1,65	2,68	5,32
Máximo	1,14	1,55	2,62	5,11	31,33

Fuente. Autoría propia.

La segunda dimensión es la frecuencia e indica cuántas veces fue comprada una sección. Al igual que en la dimensión anterior se clasificaron los datos por quintiles, como se puede apreciar en la tabla 6, teniendo en cuenta la cantidad de veces que han sido comprados. Se asignó un valor de cinco las secciones más vendidas y un valor de uno a las menos vendidas.

Tabla 6.
Valores máximos y mínimos para cada quintil de la frecuencia.

	Frequency				
Quintiles	5	4	3	2	1
Mínimo	5542	1600	519	140	4
Máximo	142489	4498	1417	515	128

Fuente. Autoría propia.

Finalmente, la tercera dimensión evalúa la participación monetaria de una sección en las ventas. Los productos en los que los clientes más han invertido en el pasado tienen más posibilidad de ser comprados que aquellos en los que han invertido menos. Se clasificaron los productos por sección en sus cinco quintiles como se puede ver en la tabla 7. El valor de cinco fue asignado a las secciones con mayor participación en los ingresos y uno a las que menos.

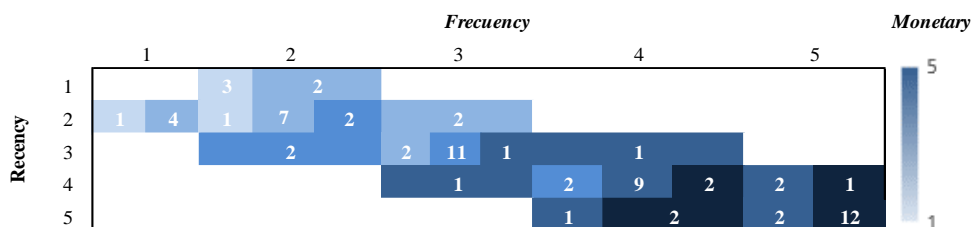
Tabla 7.
Valores máximos y mínimos para cada quintil del valor monetario.

	Monetary				
Quintiles	5	4	3	2	1
Mínimo	\$ 155.274.794	\$ 54.809.806	\$ 25.691.529	\$ 6.891.412	\$ 113.067
Máximo	\$ 2.510.774.252	\$ 138.485.040	\$ 54.578.180	\$ 19.776.871	\$ 5.852.847

Fuente. Autoría propia.

Una vez asignadas las puntuaciones a todas las secciones se realizó una tabla cruzada para identificar aquellas con mayor calificación en las tres categorías. La tabla 8 muestra la cantidad de secciones que se encuentran en las diferentes categorías. Para seleccionar las secciones a impulsar se tuvieron únicamente en cuenta aquellas que tuvieran puntuación entre cuatro y cinco para las tres categorías.

Tabla 8.
Segmentación de las secciones.



Fuente. Autoría propia.

Se obtuvo que existen 31 de 85 secciones que cumplen con el criterio de puntuaciones escogido, sin embargo, por pertinencia, fueron escogidas siete posibles secciones mostrada en la tabla 9. Según las secciones sugeridas, el retail es quien decide cuales de estas desea impulsar según sus limitaciones de espacio, inventario, costos, ganancias, entre otras. Así mismo, es quien decide las referencias de cada sección que desea ubicar cercanas a los puntos de pago. Sin embargo, se hizo un análisis de las referencias que generaron mayores ganancias y de las que menos unidades se vendieron de cada una de estas secciones en los últimos cuatro meses, sugiriendo que sean estas las que se impulsen por medio de cross selling para esta temporada.

Tabla 9.
Análisis RFM de las secciones seleccionadas.

Sección	Recency (días)	Rscore	Frecuency (unidades)	Fscore	Monetary (pesos)	Mscore
Bermuda cinco bolsillos	1,28	4	3568	4	\$ 182.172.206	4
Cachucha	1,23	4	3038	4	\$ 54.809.806	4
Cinturon	1,13	4	3363	5	\$ 59.294.512	5
Gafas	1,06	4	4360	5	\$ 83.981.921	5
Medias	1,15	5	11837	4	\$ 75.062.223	4
Short cinco bolsillos	1,12	5	6256	5	\$ 286.916.710	5
Tenis	1,17	5	8788	4	\$ 596.863.786	4

Fuente. Autoría propia.

4.2.3. Reglas según productos elegidos.

Tabla 10.
Información reglas independientes de los productos.

Estado del cliente:	Comprando.
Número de productos seleccionados:	Mínimo 1.
Otros datos:	Transacciones históricas.

Fuente. Autoría propia.

En esta etapa el cliente se encuentra en el punto de pago. Dada la información histórica de las transacciones del negocio, se diseñó un algoritmo de reglas de asociación con el fin de encontrar relaciones entre los artículos de la tienda y los que el cliente va a adquirir. Adicional, se hizo un análisis sobre las promociones que ofrece el retail y con ellas se generaron otro tipo de recomendaciones.

4.2.3.1. Reglas de asociación.

Se diseñó un algoritmo en Python usando las librerías *numpy*, *mlxtend* y *pandas*. El procesamiento de la información para generar las reglas de asociación fue lento por el consumo de memoria que requiere, por ello fue necesario ejecutar el código en una máquina virtual creada en Google Cloud Platform.

El algoritmo se desarrolló basado en reglas de asociación y sugiere prendas de vestir relacionadas con las preferencias del consumidor. Contextualizado al tema de estudio, una regla tipo Camiseta → Pantalón se interpreta como: las personas que compran camisetas también compran pantalones.

Para la creación del algoritmo, se analizó el número de ítems que componían cada transacción en la base de datos, obteniendo que 100.931 transacciones, es decir, el 53,78% contenían un solo ítem. En la figura 4 se puede apreciar la cantidad de transacciones con determinado número de ítems.

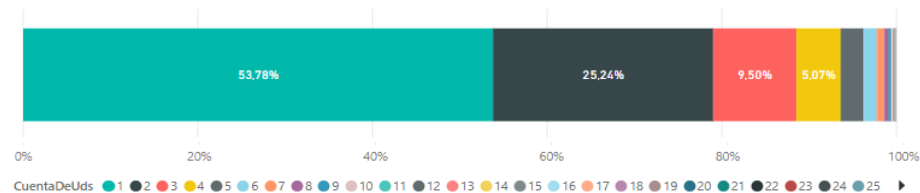


Figura 4. Cantidad de transacciones con determinado número de ítems.

Con base en el dato obtenido anteriormente y teniendo en cuenta las condiciones que se necesitan para establecer reglas de asociación, en la ejecución del algoritmo se tomaron de la base de datos histórica solo aquellas transacciones que contenían más de un ítem, dejando 86.729 transacciones y pasando de tener un promedio de productos por transacción de 1,96 a 3,07 unidades.

Los datos de entrada del algoritmo son tablas en Excel con dos columnas, la primera de ellas hace referencia a los indicativos de las facturas y la segunda al elemento del atributo adquirido sobre el cual se quieren obtener las reglas de asociación. Partiendo de la tabla de entrada el algoritmo crea una matriz binaria encabezada así: las filas están constituidas por cada transacción y las columnas por cada uno de los elementos del atributo seleccionado. Por ejemplo, si se quisieran obtener reglas de asociación con respecto al departamento, las dimensiones de la matriz serían $86.729 * 20$ (cantidad de departamentos existentes). La matriz asigna *true* en caso de que la factura contenga el elemento del atributo y *false* de lo contrario.

Se fijaron los valores de los umbrales mínimos de *support* y *confidence* en 0,01 y 0,5 respectivamente, teniendo en cuenta que son los umbrales comúnmente usados en modelos de clasificación de reglas de asociación como CBA, CMAR, CPAR y TFPC (Hernández, Carrasco, Martínez, & Hernández, 2010). De tal forma que, por un lado, se utilizó el algoritmo a priori obteniendo los ítems frecuentes, es decir, aquellos que están en más de un 1% de las transacciones, y por el otro lado, las reglas de asociación generadas serán aquellas para las que los ítems frecuentes satisfagan la probabilidad del 50% de ocurrencia. De las reglas obtenidas las finalmente seleccionadas fueron aquellas para las que el *lift* era superior a 1. Este indicador representa la independencia entre el antecedente y el consecuente de una regla. Cuando el *lift* es superior a 1 quiere decir que el antecedente y el consecuente ocurren frecuentemente. Es decir, existe dependencia entre ambos artículos y uno permite predecir el comportamiento del otro (Walid, 2018).

Se ejecutó el algoritmo con el fin de encontrar reglas de asociación para los atributos departamento, sección y color. Adicional, para profundizar en el comportamiento del cliente y así encontrar reglas de asociación más alineadas a este, se segmentó la base de datos por género (Devi & Rajagopalan, 2012). De tal manera que, fueron evaluados los mismos atributos, pero bajo el comportamiento de los compradores de prendas femeninas y los de masculinas. En la figura 5 se pueden observar algunas de las reglas más fuertes encontradas para los

departamentos según el género, por ejemplo, la regla pantalón dril, accesorios de marroquinería → camisa para el género masculino tiene un *lift* de 2,37 y la regla camiseta polo → camiseta para el género femenino tiene un *lift* de 1,07.

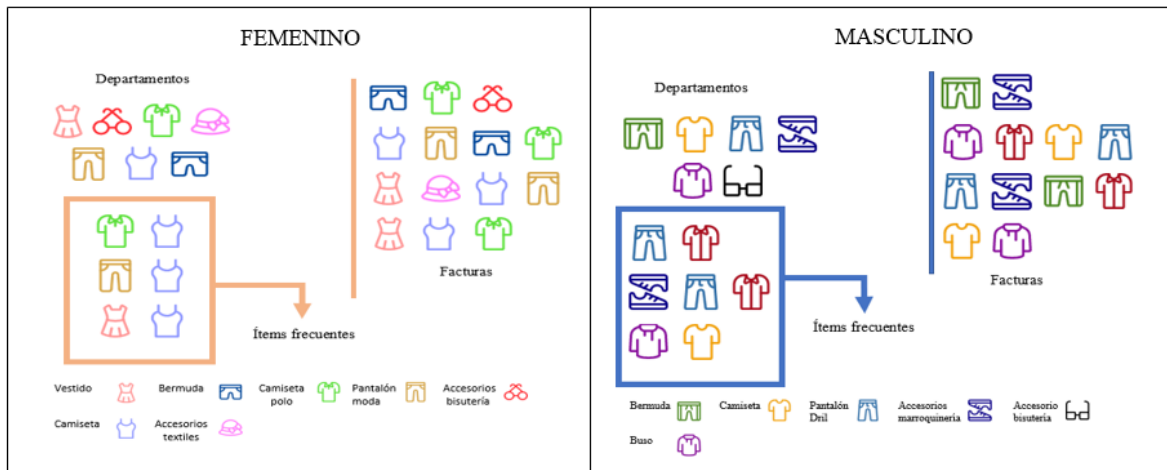


Figura 5. Reglas de asociación por departamento para cada género

Desde otra perspectiva, en el trabajo de (Omiecinski, 2003) se propone incluir *bond* como una medida de lo interesante de una asociación, siendo esta similar al *support*, pero respecto a un subconjunto de datos y no sobre el total de los datos. De la misma manera en el trabajo de (Kotsiantis & Kanellopoulos, 2013), se definen subconjuntos dentro de la base de datos que satisfacen limitaciones de tiempo. En ambos casos la idea es encontrar conjuntos de elementos frecuentes en un conjunto de intervalos de tiempo. En este caso, las características de los datos definen los subconjuntos, no el usuario final. Basado en estos estudios, se decidió establecer subconjuntos de datos de acuerdo con la duración promedio que tiene una prenda en el retail y así ejecutar el algoritmo para cada subconjunto de datos, obteniendo 44 nuevas reglas de asociación.

4.2.3.1. Reglas de promociones.

El retail con el cual se está desarrollando el proyecto está enfocado en ofrecer colecciones de alta calidad a precios muy competitivos, el segmento del mercado hacia el cual se enfocan sus estrategias son personas jóvenes que desean vestir a la moda. Luego, el mercado potencial son hombres y mujeres entre 14 y 28 años. A partir de las estadísticas ofrecidas por el DANE hasta el año 2018, segmentando la población total colombiana según los municipios donde se encuentran las diferentes franquicias de la marca y según la tasa global de participación de la población joven, se obtiene que el mercado potencial del retail es el 0,47% de la población colombiana (Departamento administrativo nacional de estadística, 2018).

Este segmento se caracteriza por ser sensible al precio y por ello aprovecha las oportunidades de comprar un producto a bajo precio a través de promociones (Muchardie, Gunawan, & Aditya, 2019). Las promociones tienen como propósito aumentar el tráfico de clientes en la tienda, aumentar la venta de los productos con mucho inventario y aumentar la rentabilidad de la categoría (Blattberg & Neslin, 2010). De esta forma, en busca de añadir recomendaciones enfocadas en promociones, se le harán al cliente dos recomendaciones. La primera de ellas utiliza el principio de normalización estadística donde se transforman valores medidos en diferentes escalas a valores adimensionales, con el objetivo de poder comparar elementos de distintas variables y distintas unidades de medida (Rodo, 2017) y la segunda utiliza una de las estrategias que aplica la marca actualmente.

Para la primera, se establecieron dos indicadores adimensionales, cuyo cociente permitiera evaluar la relación porcentaje de promoción - precio final del producto. El primer indicador es el cociente entre el porcentaje de

descuento que tiene un producto y el porcentaje promedio de descuento de todos los productos, mientras el segundo indicador es el cociente entre el precio del producto luego de aplicar el descuento y el promedio de los precios finales del **cuartil** al que pertenezca el artículo, tomando como referencia los **cuartiles** 25, 50 y 75 de acuerdo con los precios finales de los artículos. Los productos en promoción que se recomendarán serán aquellos para los que el cociente entre ambos indicadores sea mayor a 1, es decir el valor del primer indicador está al menos una vez en el segundo, y cuyo valor en inventario sea superior o igual al valor promedio de unidades en inventario.

Teniendo en cuenta que no se contaba con la información de los descuentos aplicados a las prendas en el horizonte de tiempo seleccionado, estos valores se fijaron extrayendo información de la página web de la marca y estableciendo los porcentajes mínimo y máximo de descuento que se aplicaban a los artículos del retail para cada departamento. Se determinó que solo el 25% de las referencias tuvieran un porcentaje de descuento aleatorio entre el máximo y mínimo establecido por la marca para cada departamento.

Para la segunda, el retail ofrece al cliente descuentos por acumular un monto específico de dinero en compras, a este se le otorga un bono por cierto porcentaje de la compra realizada para adquirir otros productos en la tienda. Por ejemplo, si un cliente acumula \$100.000 en compras obtiene el 30% del valor de su compra en bonos. Se diseñó un algoritmo que permite sugerir a los clientes productos que les permitan acceder a este tipo de promociones. Si el cliente cumple con el monto mínimo se da por terminada la compra, si por el contrario el monto de su compra es inferior al valor mínimo establecido en la promoción, al cliente se le recomendarán productos que le permitan alcanzarla.

El algoritmo considera el monto actual de la compra y busca dentro de las referencias existentes aquellas con el precio mínimo para lograr alcanzar la promoción, con una tolerancia superior de \$4.000. Retomando el ejemplo anterior, si la compra actual del cliente es de \$50.000, el algoritmo busca y sugiere todas las referencias que tengan precio entre \$50.000 y \$54.000.

4.2.4. Oportunidad de implementación.

Se consideraron diferentes escenarios posibles en el contexto real del negocio con el fin de usar las reglas establecidas anteriormente durante la permanencia del cliente dentro del local. En la figura 6 se puede apreciar el trayecto que podría seguir un cliente potencial al entrar en el establecimiento. Dependiendo de la fase donde este se encuentre, se aplicará una estrategia compuesta por alguna regla.

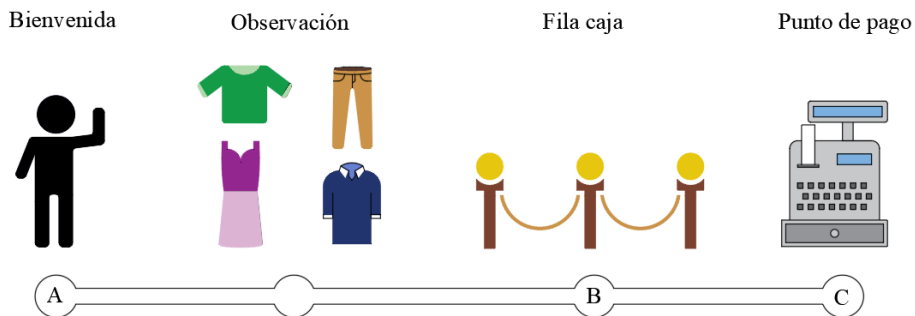


Figura 6. Trayecto del cliente en el establecimiento.

4.2.4.1. Escenario A.

Bienvenida. Es el momento donde el vendedor tiene su primer contacto con el cliente, en esta estrategia se usarán las reglas de bienvenida. El vendedor puede decidir según las necesidades del establecimiento, entre

sugerir al cliente la compra de un producto estrella o sugerir un producto con numerosas existencias en inventario.

Existen dos escenarios, en el primero de ellos, el cliente expresa lo que necesita y en el segundo aún no ha definido qué desea llevar. Para cada uno de estos escenarios se implementa la estrategia del producto estrella de una forma diferente, en el primer caso, se le recomendará al usuario una referencia estrella del departamento o sección que este desea adquirir según su género o del que manifieste será el dueño de la prenda. En el segundo caso, según el género del cliente, se le recomendará alguna de las referencias estrella.

Recomendar elementos populares genera un alto nivel de precisión en la recomendación pero se pueden presentar sesgos de concentración, quiere decir que el modelo recomienda artículos populares generando sugerencias repetitivas (Adamopoulos & Tuzhilin, 2014). Con el fin de resolver este sesgo y mejorar la diversidad en las recomendaciones, el aplicativo mostrará de manera aleatoria, y con posibilidad de actualización, un máximo de tres referencias estrella que pueden ser sugeridas al cliente.

Para el caso en el que el vendedor desee hacer uso de la estrategia de inventario, al inicio del día este deberá actualizar en el aplicativo el inventario que se tiene de cada referencia, si ingresa al módulo de inventario en el aplicativo podrá visualizar las primeras 24 referencias que cumplen con los criterios propuestos para este tipo de regla.

4.2.4.2. Escenario B.

Fila caja. En este momento no hay un contacto directo entre el cliente y el vendedor, sin embargo, a través de los productos que se encuentran cercanos a la caja, se aplica la estrategia cross selling, por ello para este momento se usarán las reglas independientes de los productos. En el aplicativo se puede acceder a las referencias sugeridas en el módulo de la marca según si se desea saber las cinco referencias que más ingresos representan o las referencias que menos unidades vendidas tienen, para cada sección escogida.

4.2.4.3. Escenario C.

Punto de pago. En este momento el vendedor nuevamente tiene un rol de intermediario, esta vez para hacer cross selling a través de las reglas de productos elegidos según los artículos que el cliente está comprando. Una vez el cliente se encuentra en el punto de pago, según los artículos comprados se le recomendará el artículo consecuente, correspondiente a la regla de asociación que aplique.

El artículo complementario está determinado por el departamento o la sección de los productos adquiridos. En caso de que el cliente compre una camiseta y no se encuentre una regla de asociación, se le recomendará llevar otra teniendo en cuenta que, en el análisis estadístico realizado, se identificó que este era el departamento más vendido y que en el 41,36% de las transacciones con más de un ítem adquirido, se venden dos o más unidades de camisetas. Adicional, en el 61,81% de las transacciones están presentes camisetas manga corta, en consecuencia, las referencias de esta sección serán las sugeridas aleatoriamente. El color de la referencia seleccionada para recomendar estará determinado por las reglas de asociación por color establecidas, de tal manera que se le recomendará una camiseta manga corta del color consecuente al adquirido.

Por otra parte, en caso de que el cliente adquiera un producto diferente a camiseta y no exista una regla de asociación, se le recomendará aleatoriamente un producto en promoción según las reglas de promociones establecidas.

A continuación, para cualquiera de los dos casos, se verifica si la compra supera el umbral mínimo de dinero en compras establecido para lograr la promoción adicional y obtener los bonos. En caso de no hacerlo al cliente se le sugiere algunos de los productos que arroja el algoritmo.

4.3. Evaluación.

Las métricas utilizadas para evaluar las reglas de asociación fueron support, confidence, lift, serendipia y diversidad. Las reglas consideradas apropiadas fueron aquellas cuyo lift era superior a 1, y su confidence y support superaban el umbral mínimo establecido inicialmente. Tomando como ejemplo la regla generada al ejecutar el algoritmo para departamento del género masculino, correspondiente a pantalón dril → camiseta, su support indica que esta combinación se encuentra en el 7,17% de las transacciones, siendo este valor superior al 1% establecido. De la misma manera, su confidence es 55,42%, lo que sugiere que, si se compró pantalón dril, en un 55,42% de los casos se comprará también camiseta, satisfaciendo el confidence mínimo del 50%. Finalmente, relacionando los dos indicadores anteriores se obtiene que la regla tiene un lift de 2,02 lo que permite que esta y todas aquellas reglas que cumplen con estos mismos criterios sean fiables y puedan ser recomendadas al usuario.

Para garantizar que las recomendaciones sean diversas, las referencias sugeridas siempre son elegidas de manera aleatoria dentro del conjunto de referencias que cumplen con los criterios establecidos en cada estrategia. Adicional, es posible si el usuario lo desea, actualizar las referencias mostradas. El aplicativo recomienda el 24% de las referencias existentes y muestra una cantidad de recomendaciones mayor a uno en cada estrategia, siempre y cuando el diseño de esta lo permita y haya disponibilidad en el retail.

Con el fin de evaluar la serendipia en las recomendaciones se debe tener en cuenta que esta incluye una dimensión emocional que la hace subjetiva, en consecuencia, no existe un método estandarizado para evaluarla (Iaquinta et al., 2008; Manca et al., 2018) . Sin embargo, (Kotkov, Wang, & Veijalainen, 2016) afirma que cuanto menor es la probabilidad de que el usuario conozca un elemento, mayor será la probabilidad de que este pueda dar lugar a una recomendación fortuita, relacionando esta probabilidad con la popularidad de los artículos, de tal forma que si un artículo es popular es más probable que los usuarios estén familiarizados con este. De la misma manera, (Zheng, Chan, & Ip, 2015) sugiere que un elemento es novedoso si es impopular y difiere con respecto al perfil de los usuarios. En el aplicativo, además de sugerir ubicar productos poco vendidos cercanos a las cajas de pago, se implementó un botón “sorpréndeme” que genera recomendaciones aleatorias entre un conjunto de productos impopulares o poco vendidos. El vendedor puede hacer uso de esta opción en el momento que lo desee.

4.4. Implementación.

Se ejecutó la lógica del sistema de recomendación en Visual Basic, fueron definidos los parámetros de entrada y salida del aplicativo, se llevaron a cabo sus pruebas y se evaluaron los requerimientos de las normas ISO 25010 y 25021.

Con el propósito de evaluar la percepción de los vendedores frente al aplicativo e identificar oportunidades de mejora, se desarrolló una investigación cualitativa mediante un focus group, para su desarrollo inicialmente fueron reclutados cinco individuos, específicamente usuarios que interactuarían directamente con el aplicativo y que cumplieran con el perfil establecido, es decir, vendedores de la marca local con una experiencia superior o igual a seis meses. Fueron formuladas nueve preguntas abiertas y fáciles de comprender con el fin de obtener respuestas amplias y que brindaran información relevante sobre la postura del usuario. Finalmente se analizaron los resultados obtenidos.

5. Componente de Diseño en ingeniería.

Declaración de diseño.

Se diseñó una aplicación en bache que permite visualizar el sistema de recomendación. Los vendedores tendrán a su disposición el aplicativo para que durante la permanencia del cliente en la tienda y de acuerdo con las

estrategias propuestas puedan transmitir recomendaciones.

Proceso de diseño.

Para el desarrollo del aplicativo, se usó únicamente la información de las referencias existentes durante los últimos cuatro meses, con ellas se aplicaron cada una de las estrategias propuestas y con los resultados obtenidos se generaron tablas que sirvieran como entrada para el aplicativo. Para su desarrollo, se creó un servidor local y en él una base de datos a través del Software SQL Server Management Studio donde a su vez se importaron las tablas mencionadas anteriormente. Desde Visual Basic se accedió al servidor local y a continuación, se creó un formulario principal y se programó la lógica del aplicativo. Para los botones y los fondos del formulario se diseñaron imágenes en el software Inkscape y se agregaron como recursos en la aplicación.

En el siguiente enlace <https://www.youtube.com/watch?v=7Xo50SEFeAI> se puede visualizar el funcionamiento del aplicativo, en general el tiempo que el aplicativo tarda en generar una recomendación es inferior a tres segundos y depende de la estrategia a la que acceda el usuario, adicional este no recomienda referencias que no tengan inventario existente dentro del retail. Se diseñó un instructivo del aplicativo que se puede apreciar como anexo.

Restricciones de diseño.

- Solo se cuenta con la base de datos de transacciones realizadas en la marca local, por ello el aplicativo es de uso exclusivo para esta.
- No se tiene en cuenta el tallaje de las prendas para generar recomendaciones.
- El aplicativo fue diseñado en bache, por ello en caso de requerir una actualización del razonamiento del aplicativo se deben actualizar la base de datos y las tablas para realizar un nuevo análisis.
- El aplicativo no es fácilmente portable. Para poder acceder a él se debe tener conexión con el servidor local en SQL.

Cumplimiento del estándar.

El aplicativo se desarrolló bajo los estándares de las normas ISO/IEC 25010 y la Norma ISO/IEC 25021.

- Norma ISO/IEC 25010 - *System and software quality models*: Se consideraron las ocho características principales contenidas en la norma, estas son: adecuación funcional, eficiencia de desempeño, compatibilidad, usabilidad, fiabilidad, seguridad, mantenibilidad y portabilidad (ISO 25000, 2019).
- Norma ISO/IEC 25012 - *Data Quality model*: Se evaluó la calidad de los datos a partir de las características exactitud, completitud, existencia y credibilidad (ISO 25000, 2019).

6. Resultados.

6.1. Pruebas computacionales.

Del código realizado en Python fueron obtenidas en total 134 reglas, de las cuales el 56,72% cumplen con las métricas establecidas y por ello se seleccionaron para implementar en el aplicativo. Se realizaron pruebas [durante el entrenamiento](#) para verificar el grado de acierto que tenían las reglas y comprobar el funcionamiento correcto del aplicativo. De un total de 16.578 transacciones realizadas en los últimos cuatro meses, fueron seleccionadas 105 que tuvieran entre tres y cinco ítems comprados aleatoriamente.

Las pruebas realizadas consistían en eliminar un artículo en las transacciones que tuvieran tres, dos en las que tuvieran cuatro y tres en las que tuvieran cinco, agregar la información restante a la aplicación y evaluar para cada transacción si dentro de las recomendaciones generadas por el aplicativo se encontraban los artículos

eliminados. A la vez que se evaluaron las recomendaciones se verificó también el funcionamiento del aplicativo y se obtuvo que en todos los casos este funcionaba adecuadamente, permitía añadir los artículos al carrito de compras, actualizar el inventario de las referencias y en general se cumplían cada una de las declaraciones de diseño establecidas.

De estas pruebas se obtuvo que en el 54,29% de los casos las recomendaciones realizadas por el aplicativo lograban predecir las compras realizadas por un cliente. De este porcentaje se logró identificar que en 32 de las transacciones la recomendación acertaba con el departamento del artículo comprado, en 9 transacciones acertaba usando las reglas de asociación por sección y consecuentemente en 13 transacciones se podía predecir el comportamiento de los clientes a partir de reglas de asociación con cualquiera de los dos atributos.

Para las transacciones con tres ítems se obtuvo que en el 58,10% de los casos se logró predecir el ítem faltante, mientras para las transacciones con cuatro ítems en el 47,06% se identificó al menos uno de los ítems faltantes y de esa proporción en el 30% se lograron predecir ambos ítems. Finalmente, para las transacciones con cinco ítems en el 11,11% de los casos se encontraron los tres ítems faltantes, en el 11,11% se identificaron dos ítems, en el 33,33% se encontró solo uno de los ítems faltantes y en el porcentaje restante no se encontró ninguno.

6.2. Focus group.

Para dar cumplimiento al tercer objetivo propuesto en el proyecto se realizó un focus group. Al ser una técnica cualitativa, fue grabada en audio y posteriormente transcrita, para así, poder ser analizada en el Software Textalyzer el cual permite identificar las palabras y frases más frecuentes usadas por los usuarios en cada pregunta. En términos generales se obtuvieron resultados tanto para la interfaz como para las estrategias implementadas.

Por un lado, los usuarios consideraron pertinente la creación de una interfaz más llamativa en cuanto a color y estilo, a su vez expresaron que una interfaz con fotos de los artículos facilitaría la comunicación de las recomendaciones y su uso sería más sencillo. Por otro lado, tres de ellos opinaron que sería más útil si la usara el cliente directamente por medio de la página web de la marca pues el tiempo que le tomaría a el vendedor usarla podría perjudicar sus ventas y generar conflicto con el cliente. Así mismo, fue unánime la opinión de simplificar las funciones que tiene el aplicativo para hacer ver este menos complejo, los usuarios reemplazarían algunas de las funciones actuales por otras que ellos consideran facilitarían la atención al cliente, dentro de estas se encontraba una opción de filtrado en una ventana donde se pudiera apreciar un listado de todas las referencias existentes junto con sus atributos, añadir una sección donde se puedan visualizar todas las ofertas, adicionar un atributo del cuidado que se debe tener para el lavado de la prenda y finalmente una notificación para identificar que un producto se ha agotado o no se está vendiendo.

Haciendo énfasis en las estrategias propuestas para implementar en el aplicativo, los usuarios expresaron que estas se alineaban a su día a día y coincidían en que es importante brindarle una buena experiencia al cliente donde se sienta cómodo y más allá de generar una venta, crear una experiencia única que le genere un deseo de volver. En general todos expresaron que aplican y consideran importante la estrategia de cross-selling en el momento de asesorar a los clientes. Por último, dentro de los conceptos más relevantes se encuentran las recomendaciones de productos estrella, productos con alto inventario y colores complementarios. Como sugerencia de mejora, dos de ellos concordaron en que no discriminarían los momentos del cliente dentro del local para aplicar las estrategias, sino que identificarían el mejor momento para aplicar cada una de ellas.

A lo largo del desarrollo del proyecto se evidenció que la estrategia de ventas del retail está enfocada en la creación de atuendos para sus clientes, muestra de esto es que el 66,60% de las transacciones con más de un artículo incluyen más de un departamento, así mismo los participantes del focus group afirmaron que aprovechan cuando el cliente se encuentra en el probador para aplicar estrategias de cross selling, sugiriendo

prendas complementarias. A partir de esto surge la necesidad de generar recomendaciones guiadas a complementar estilos y contribuir a la generación de compras en distintos departamentos.

Se diseñó una matriz que evidencia la relación que existe entre departamentos y colores según las transacciones realizadas históricamente, para esto se tomaron todas las facturas en las cuales se compraba más de un departamento y con un *confidence* del 0.2 se escogieron las principales combinaciones que contenían distintos departamentos. El objetivo del diseño de esta matriz es que sirva como herramienta para el vendedor en el momento en que el cliente está en el probador.

6.2. Medición del impacto.

Durante el desarrollo del proyecto se logró identificar que la implementación del aplicativo generaría un impacto positivo en la percepción de los clientes sobre la experiencia de compra en el retail, se reducirían costos operacionales por acumulación de inventario, lograría impulsar productos mediante cross-selling y permitiría crear relaciones duraderas con los clientes. Adicional, la estrategia que generaría mayor impacto es la implementación de productos en fila, ya que la marca en la actualidad no contempla este escenario.

No obstante, el aplicativo podría generar un mayor impacto en el mercado, si también estuviera disponible para los clientes en diferentes plataformas digitales, ahorrando el tiempo que estos permanecen en el retail y que para muchos representa un costo de oportunidad significativo.

Con el fin de medir el impacto económico que conllevaría implementar alguna de las estrategias propuestas en el retail, se tomaron las ventas realizadas para el local V27 en el mes de junio del 2019 y con ellas se midió el impacto que generaría proponer artículos en el punto de pago a través de las reglas de asociación. Como las reglas acertaron en un 54,29% de los casos, se tomó este valor como la probabilidad de que un cliente adquiera el producto consecuente a los adquiridos cuando se le sugiere este en caso de existir una regla de asociación. Se seleccionaron aleatoriamente transacciones que cumplirían con esta probabilidad y se simularon en el aplicativo, finalmente se calculó el ingreso monetario adicional si se hubieran implementado las reglas de asociación en estas ventas y se obtuvo que de ser implementada la estrategia en el retail, las ventas para este local tendrían un aumento del 18,44%.

7. Conclusiones y recomendaciones.

El presente trabajo propone el diseño de un sistema de recomendación de prendas de vestir para las tiendas de una marca local, aplicable en cada uno de los momentos por los que pasa un cliente potencial durante su permanencia en la tienda. Se diseñaron con base en estos momentos, diferentes estrategias enfocadas en aumentar las ventas, mejorar la experiencia de compra del cliente y lograr que el retail consiga responder las necesidades del mercado.

Durante el desarrollo del proyecto se intentó dar respuesta a la pregunta planteada inicialmente: *¿Cómo lograr identificar las preferencias de los usuarios de esta marca para llegar a cumplir las expectativas de sus clientes y permitir que la organización cumpla con sus objetivos estratégicos?* Realizar un buen análisis estadístico inicial de los datos fue clave para identificar patrones de comportamiento en los clientes y con ello explorar la posibilidad de implementar diferentes tipos de recomendaciones que van más allá de las que se podrían obtener con un modelo netamente teórico, logrando así crear un aplicativo que ejecute un sistema de recomendación alineado a las necesidades de la marca y el cliente.

A pesar de que no fue posible encontrar reglas de asociación por referencias en la marca local, debido a la gran variedad existente y a su corta duración en exhibición, se encontraron 76 reglas de asociación que al ser probadas fueron acertadas en un 54,29% de los casos. Lo anterior refleja un buen resultado en cuanto al

funcionamiento de las reglas de asociación, asegurando que se recomienden artículos que respondan a los patrones de comportamiento de los clientes.

Teniendo en cuenta que el principal objetivo de los sistemas de recomendación es ayudar a los usuarios a encontrar contenido de su interés entre una gran cantidad de opciones existentes, se implementaron estrategias que permitieron involucrar métricas como serendipia y diversidad en las recomendaciones para evitar recomendar los artículos más populares y que el usuario probablemente ya conoce.

Considerando que aplicar herramientas de análisis cualitativo en el mercado permite realizar predicciones acerca del impacto que tendrá el aplicativo sobre los usuarios finales y considerar oportunidades de mejora para garantizar su éxito, se realizó un focus group con cinco de los vendedores de la marca para evaluar la aceptación del usuario respecto al aplicativo y sus recomendaciones. De este se detectaron oportunidades prontas de mejora y por ello se realizaron cambios en la fuente y los colores del aplicativo para hacerlo visualmente más atractivo. También, se diseñó un instructivo que explica cada una de las funciones del aplicativo, para garantizar el entendimiento de su funcionamiento y dar una idea de los pasos que se deben seguir para implementar las estrategias desarrolladas en este trabajo.

Durante el desarrollo del proyecto y la evaluación del aplicativo, se detectaron algunas recomendaciones que de ser consideradas en trabajos futuros permitirían mejorar la calidad de las recomendaciones, aumentar impacto del aplicativo y perfeccionar su funcionamiento. Así mismo, se detectaron oportunidades estratégicas para la marca, que les permitirían mejorar sus ventas.

En la aplicación no se utilizaron fotos de las referencias, pero se reconoce la importancia que estas tienen visualmente para el usuario y para el vendedor a nivel práctico, por ello es necesario integrarlas en el aplicativo si se quiere que este tenga un mejor nivel de aceptación por ambas partes.

Para el desarrollo del proyecto no se contaba con información sobre el stock y el porcentaje de descuento de cada referencia, para implementar el aplicativo en la marca local se recomienda reemplazar esta información por la real para obtener recomendaciones más ajustadas a la oferta. Adicional, lo ideal sería que el aplicativo tuviera acceso a la información en tiempo real sobre el stock, con el fin de que sea más eficiente y preciso.

Además de las estrategias de recomendación de productos existen distintos métodos para aumentar las ventas, por un lado, se recomienda a la marca hacer uso de la matriz relación departamentos-colores para hacer cross selling durante la permanencia del cliente en el probador. Por el otro, existe una estrategia conocida como *Visual Merchandasing* mediante el cual se pueden atraer compradores a la tienda para luego alentarlos a permanecer en esta, comprar y tener una experiencia positiva (Morgan, 2011). El diseño de la tienda es uno de los factores más importantes en este método, una buena distribución conduce al cliente a verse obligado a atravesar la tienda para encontrar la salida, mientras lo hace, es conducido por distintos departamentos de la tienda con el propósito de incentivar la compra. La estrategia que se propone implementar en el retail consiste en situar los productos más vendidos en el centro de tienda o en la parte posterior y de esta forma garantizar que durante la búsqueda de estos productos los compradores probablemente agreguen otros productos que no tenían pensado comprar.

Referencias

Adamopoulos, P., & Tuzhilin, A. (2014, October). On over-specialization and concentration bias of recommendations: Probabilistic neighborhood selection in collaborative filtering systems. Paper presented at the *Proceedings of the 8th ACM Conference on Recommender Systems*, 153-160.

- Ahmadon, M. A. B., & Yamaguchi, S. (2018, October). User workflow preference analysis based on confidence and lift value of association rule. Paper presented at the *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*, 610-613.
- Akcura, M. T., & Srinivasan, K. (2005). Research note: Customer intimacy and cross-selling strategy. *Management Science*, *51*(6), 1007-1012. doi:10.1287/mnsc.1050.0390
- Bernard, F., & Pawel, W. (2010). Cross- and up-selling techniques in E-commerce activities. *Journal of Internet Banking and Commerce*, *15*(3), 1.
- Blattberg, R. C., & Neslin, S. A. (2010). *Sales promotion: Concepts, methods, and strategies*. Englewood Cliffs, NJ: Prentice Hall.
- Bodaghi, A., & Homayounvala, E. (2018, April). Personalization of interactive recommender systems for expert users. Paper presented at the *2018 4th International Conference on Web Research (ICWR)*, 58-62.
- Bolívar, M., & Giraldo, J. E. (2017). Sistema de recomendación de recursos digitales como apoyo a las actividades de trabajo independiente en cursos de programación de computadores. Paper presented at the *Desarrollo E Innovación En Ingeniería*, 63-78.
- Burke, R. (2007). Hybrid web recommender systems. *The adaptive web* (pp. 377-408). Berlin, Heidelberg: Springer.
- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, *12*(4), 331-370.
- Chan, C. C. H. (2008). Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer. *Expert Systems with Applications*, *34*(4), 2754-2762. doi:10.1016/j.eswa.2007.05.043
- Demirezen, E. M., & Kumar, S. (2016). Optimization of recommender systems based on inventory. *Production and Operations Management*, *25*(4), 593-608. doi:10.1111/poms.12540
- Departamento administrativo nacional de estadística. (2018). Censo nacional de de población y vivienda 2018. Retrieved from <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivenda-2018>
- Devi, P., & Rajagopalan, S. (2012). Analysis of customer behavior using clustering and association rules. Paper presented at the *International Journal of Computer Applications*, *43*(23), 19-27.
- Faggella, D. (2017). The ROI of recommendation engines for marketing. Retrieved from <https://martechtoday.com/roi-recommendation-engines-marketing-205787>
- Fry, E. (2018). *Zalando*. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=126930500&site=eds-live>
- Golbeck, J. (2006). Generating predictive movie recommendations from trust in social networks. *Trust management* (pp. 93-104). Berlin, Heidelberg: Springer.
- Gower, S. (2014). *Netflix prize and and SVD*
- Gutta, S., Kurapati, K., Lee, K., Martino, J., Milanski, J., Schaffer, D., & Zimmerman, J. (2000). Tv content recommender system. *Aaai/Iaai, 2000*.

- Hao, K. (2019). DeepMind is asking how AI helped turn the internet into an echo chamber. Retrieved from <https://www.technologyreview.com/s/613083/deepmind-is-asking-how-google-helped-turn-the-internet-into-an-echo-chamber/>
- Hargreaves, S. (2015). Your clothes are killing us. *CNNMoney*, Retrieved from <https://money.cnn.com/2015/05/22/news/economy/true-cost-clothing/index.html>
- He, M., Zhang, S., & Meng, Q. (2019). Learning to style-aware bayesian personalized ranking for visual recommendation. *IEEE Access*, 7(1), 14198-14205. doi:10.1109/ACCESS.2019.2892984
- Hernández, R., Carrasco, J. A., Martínez, J. F., & Hernández, J. (2010, November). Classifying using specific rules with high confidence. Paper presented at the *2010 Ninth Mexican International Conference on Artificial Intelligence*, 75-80.
- Hwangbo, H., Kim, Y. S., & Cha, K. J. (2018). Recommendation system development for fashion retail e-commerce. *Electronic Commerce Research and Applications*, 28(1), 94-101. doi:10.1016/j.elerap.2018.01.012
- Iaquinta, L., de Gemmis, M., Lops, P., Semeraro, G., Filannino, M., & Molino, P. (2008, September). Introducing serendipity in a content-based recommender system. Paper presented at the *2008 Eighth International Conference on Hybrid Intelligent Systems*, 168-173.
- IBM. (2017). *Documentación de IBM SPSS statistics subscription* Retrieved from https://www.ibm.com/support/knowledgecenter/es/SSLVMB_sub/statistics_mainhelp_ddita/spss/rfm/rfm_intro.xml.html
- ISO 25000. (2019). Iso/iec 25012. Retrieved from <https://iso25000.com/index.php/normas-iso-25000/iso-25012>
- Ja-Hwung, S., Wei-Yi, C., & Tseng, V. (2017). Effective social content-based collaborative filtering for music recommendation. *Intelligent Data Analysis*, 21(1), 195-216. doi:10.3233/IDA-170878
- Jariha, P., & Jain, S. K. (2018, April). A state-of-the-art recommender systems: An overview on concepts, methodology and challenges. Paper presented at the *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 1769-1774.
- Johnson, J. S., & Friend, S. B. (2015). Contingent cross-selling and up-selling relationships with performance and job satisfaction: An MOA-theoretic examination. *Journal of Personal Selling & Sales Management*, 35(1), 51-71. doi:10.1080/08853134.2014.940962
- Johnson, J. S., & Friend, S. B. (2015). Contingent cross-selling and up-selling relationships with performance and job satisfaction: An MOA-theoretic examination. *Journal of Personal Selling & Sales Management*, 35(1), 51-71. doi:10.1080/08853134.2014.940962
- Kabir, M. F., Ludwig, S. A., & Abdullah, A. S. (2018, December). Rule discovery from breast cancer risk factors using association rule mining. Paper presented at the *2018 IEEE International Conference on Big Data (Big Data)*, 2433-2441.
- Kamakura, W. (2008). Cross-selling. *Journal of Relationship Marketing*, 6(1), 41-58. doi:10.1300/J366v06n03_03
- Kotkov, D., Wang, S., & Veijalainen, J. (2016). A survey of serendipity in recommender systems. *Knowledge-Based Systems*, 111(1), 180-192. doi:10.1016/j.knosys.2016.08.014

- Kotsiantis, S., & Kanellopoulos, D. (2013). Association rules mining: Recent overviews. *International Journal of Advanced Research in Computer Science*, 4(8), 71-82.
- Kottage, G. N., Keshawa Jayathilake, D., Chankuma, K. C., Upeksha Ganegoda, G., & Sandanayake, T. (2018, January). Preference based recommendation system for apparel E-commerce sites. Paper presented at the *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, 122-127.
- Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *Ieee*, 7(1), 76-80. doi:10.1109/MIC.2003.1167344
- Liu, Y., Gao, Y., Bian, L., Wang, W., & Li, Z. (2018). How to wear beautifully? clothing pair recommendation. *Journal of Computer Science and Technology*, 33(3), 522-530. doi:10.1007/s11390-018-1836-1
- Lopes, A., Pinho, R., Paulovich, F., & Minghim, R. (2007). Visual text mining using association rules. *Computers & Graphics*, 31(3), 316-326. doi:10.1016/j.cag.2007.01.023
- Ma, M., & Uchyigit, G. (2008). *Personalization techniques and recommender systems*. Hackensack, NJ: World Scientific.
- Magne, S., Oivind, E., & Leif, H. (2004). Can advertisements for brand extensions revitalise flagship products? an experiment. *International Journal of Advertising*, 23(2), 173-196.
- Manca, M., Manca, M., Boratto, L., Boratto, L., Carta, S., & Carta, S. (2018). Behavioral data mining to produce novel and serendipitous friend recommendations in a social bookmarking system. *Information Systems Frontiers*, 20(4), 825-839. doi:10.1007/s10796-015-9600-3
- Mohamed, M. H., Khafagy, M. H., & Ibrahim, M. H. (2019, February). Recommender systems challenges and solutions survey. Paper presented at the *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, 149-155.
- Morgan, T. (2011). *Visual merchandising* (2nd ed.). London: King. Retrieved from http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=025129838&sequence=000002&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA
- Moro, S., Laureano, R., & Cortez, P. (2011). *Using data mining for bank direct marketing: An application of the CRISP-DM methodology* EUROSIS-ETI. Retrieved from https://www.openaire.eu/search/publication?articleId=od_307::af1168082b29a5bd464f056b598fbab3
- Muchardie, B. G., Gunawan, A., & Aditya, B. (2019, August). E-commerce market segmentation based on the antecedents of customer satisfaction and customer retention. Paper presented at the *2019 International Conference on Information Management and Technology (ICIMTech)*, 103-108.
- Mukherjee, S. (2016). *F# for machine learning essentials* (1st ed.). Birmingham, England: Packt Publishing. Retrieved from <http://unr-ra.scholarvox.com/book/88843057>
- Nahm, U., & Mooney, R. (2000, July). A mutually beneficial integration of data mining and information extraction. Paper presented at the *Seventeenth National Conference on Artificial Intelligence*, 627-632.
- Neamah, A. A., & El-Ameer, A. S. (2018, October). Design and evaluation of a course recommender system using content-based approach. Paper presented at the *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 1-6.

- Observatorio eCommerce. (2019). Medición de indicadores de consumo del observatorio eCommerce. Retrieved from <https://www.observatorioecommerce.com.co/el-observatorio-ecommerce-publica-su-primer-estudio-de-consumo-del-comercio-electronico-en-colombia/>
- Omiecinski, E. R. (2003). Alternative interest measures for mining associations in databases. *IEEE Transactions on Knowledge and Data Engineering*, 15(1), 57-69. doi:10.1109/TKDE.2003.1161582
- Omran, A. (2018, January). A novel recommender system for websites. Paper presented at the *100th the IRES International Conference*, 126-130.
- Orbis. (2018). Ventas reales en colombia por marca durante 2018. Retrieved from <https://neworbis-bvdinfo-com.ezproxy.javeriana.edu.co/version-20191219/orbis/1/Companies/Search>
- Osadchiy, T., Poliakov, I., Olivier, P., Rowland, M., & Foster, E. (2019). Recommender system based on pairwise association rules. *Expert Systems with Applications*, 115, 535-542. doi:10.1016/j.eswa.2018.07.077
- Potafolio. (2017). Los hogares colombianos gastan muy poco en ropa. Retrieved from <http://www.portafolio.co/negocios/cuanto-gasta-los-colombianos-en-ropa-506386>
- Ramesh, N., & Moh, T. (2018). Outfit recommender system. Paper presented at the *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 903-910. doi:10.1109/ASONAM.2018.8508656
- Reagan, C. (2016). A \$260 billion 'ticking time bomb': The costly business of retail returns. Retrieved from <https://www.cnbc.com/2016/12/16/a-260-billion-ticking-time-bomb-the-costly-business-of-retail-returns.html>
- Rebholz, J. (2019). *Alibaba does AI: THE CHINESE RETAIL TITAN LAUNCHES FASHIONAI, AN IN STORE PERSONALIZED RECOMMENDATION SYSTEM* Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=134726744&site=eds-live>
- Rodo, P. (2017). Normalización estadística. *Economipedia*, Retrieved from <https://economipedia.com/definiciones/normalizacion-estadistica.html>
- Rodríguez, P., Duque, N., Ovalle, D., Peluffo, D., & Salazar, O. (2014, October). Análisis de características del perfil de usuario para un sistema de recomendación de objetos de aprendizaje. Paper presented at the *Laclo 2014*, 487-493.
- Ruchika, Singh, A. V., & Sharma, M. (2017, December). Building an effective recommender system using machine learning based framework. Paper presented at the *2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS)*, 215-219.
- Semana. (2017). Cómo compran los colombianos. Retrieved from <https://www.semana.com/contenidos-editoriales/moda-asi-lo-hacemos/articulo/comportamiento-de-compra-de-ropa-en-los-colombianos/533736>
- Sharma, L., & Gera, A. (2013). A survey of recommendation system: Research challenges. Paper presented at the *Engineering Trends and Technology (IJETT)*, 4(5), 1989-1992.

- Siswanto, B., & Thariqa, P. (2018). Association rules mining for identifying popular ingredients on YouTube cooking recipes videos. Paper presented at the *2018 Indonesian Association for Pattern Recognition International Conference (INAPR)*, 95-98. doi:10.1109/INAPR.2018.8627002
- Tamturk, V. (2017). The ROI of recommendation engines. Retrieved from <http://www.cms-connected.com/News-Archive/January-2017/The-ROI-of-Recommendation-Engines>
- Terry, Y., Estrada, V., & Arteaga, Y. (2016). REP: Sistema para recomendación de patrones de diseño de recursos educativos abiertos. *Ciencias De La Información*, 47(1), 3-8.
- Vaidya, N., & Khachane, A. (2017). Recommender systems-the need of the ecommerce ERA. Paper presented at the *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 100-104. doi:10.1109/ICCMC.2017.8282616
- Wakita, Y., Oku, K., Huang, H., & Kawagoe, K. (2015). A fashion-brand recommender system using brand association rules and features. Paper presented at the *2015 IIAI 4th International Congress on Advanced Applied Informatics*, 719-720. doi:10.1109/IIAI-AAI.2015.230
- Walid Ghobar, E. (2018). *Un sistema de recomendación basado en perfiles generados por agrupamiento y asociaciones*
- Walter, F., Battiston, S., Yildirim, M., & Schweitzer, F. (2012). Moving recommender systems from on-line commerce to retail stores. *Information Systems and E-Business Management*, 10(3), 367-393. doi:10.1007/s10257-011-0170-8
- Wang, K., Chen, F., Zhaowen, W., & McAuley, J. (2017, November). Visually-aware fashion recommendation and design with generative image models. Paper presented at the *2017 IEEE International Conference on Data Mining (ICDM)*, 207-216.
- Wirth, R. (2000). *Crisp-dm: Towards a standard process model for data mining* Retrieved from <https://www.semanticscholar.org/search?q=CRISP-DM%3A%20Towards%20a%20Standard%20Process%20Model%20for%20Data%20Mining&sort=relevance>
- Xing, X., Jianxun, L., Zheng, L., Xiting, W., Fangzhao, W., Hongwei, W. & Zhongxia, C. (2018). Personalized recommendation systems: Five hot research topics you must know. Retrieved from <https://www.microsoft.com/en-us/research/lab/microsoft-research-asia/articles/personalized-recommendation-systems>
- Zalando: Fashion forward. (2016). *The Economist*, 412(901) Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85020240678&site=eds-live>
- Zeng, X., Koehl, L., Wang, L., & Chen, Y. (2013, January). An intelligent recommender system for personalized fashion design. Paper presented at the *2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*, 760-765.
- Zheng, Q., Chan, C., & Ip, H. H. S. (2015). An unexpectedness-augmented utility model for making serendipitous recommendation. Paper presented at the *Advances in Data Mining: Applications and Theoretical*, 216-230.