



Sistema de recomendación difuso de productos basado en lógica difusa

Fuzzy product recommendation system based on fuzzy logic

Sistema de recomendação de produtos fuzzy baseado em lógica fuzzy

Ricardo Geovany García-Morales ^I
rgarcia0961@utm.edu.ec
<http://orcid.org/0000-0003-2131-5717>

Emanuel Guillermo Muñoz-Muñoz ^{II}
emanuel.munoz@utm.edu.ec
<http://orcid.org/0000-0002-0997-0578>

Correspondencia: rgarcia0961@utm.edu.ec

Ciencias Técnicas y Aplicadas
Artículo de Investigación

* **Recibido:** 23 de julio de 2022 * **Aceptado:** 12 de agosto de 2022 * **Publicado:** 08 de septiembre de 2022

- I. Ingeniero Comercial, Universidad Técnica De Manabí, Ecuador.
- II. Magíster Scientiae Estadística Aplicada, Unalm, Universidad Nacional Agraria La Molina, Ingeniero en Sistemas Informáticos, Universidad Técnica De Manabí, Ecuador.

Resumen

Las tecnologías de la información son ampliamente reconocidas como instrumentos estratégicos, por su capacidad para modificar la estructura y los modelos de negocios de las organizaciones, en la actualidad la gran cantidad de información que debe manejar un individuo en su día a día lo imposibilitan optimizar tiempo en la búsqueda de información deseada, los sistemas de recomendaciones son en la actualidad una herramienta que ayuda a la toma de decisiones y que poco a poco se ha ido involucrando en el mercado y dándole las oportunidades a las empresas a ser más competitivas. El objetivo principal de esta investigación es diseñar un sistema de recomendaciones difuso para poder recomendar productos de la empresa Promyvent a sus clientes de acuerdo al perfil de cada uno de ellos. La metodología utilizada en esta investigación inicio con el pre procesamiento de la base de datos de entrada, se le aplico técnicas de preparación y limpieza además se aplicó un análisis exploratorio de Datos el cual permite realizar un análisis descriptivos, transformación y limpieza de datos, se le aplico el método del codo de jambu y se estableció que el número de clústers adecuado eran tres, luego en la fase de modelado se aplicó el algoritmo KNN para clasificar cada uno de los datos según las compras de los clientes con menor distancia calculada, luego para aumentar el nivel de precisión en la clasificación de datos se aplicó el método C Means, de esta manera se obtiene una partición difusa optima otorgando un grado de membresía a cada punto de datos. Por ultimo como resultado la recomendación se genera a partir del perfil de cada cliente, la precisión del modelo parte en base a los registros históricos de compras de los clientes para poder identificar el grado de membresía que tiene cada cliente de acuerdo a su perfil y partiendo de estos datos el modelo generó el filtrado de un top de siete productos en las recomendaciones de principal interés. El modelo implementado beneficia a la empresa no solo a recomendar un top siete de productos de interés de acuerdo al perfil del cliente si no también identificar los productos que tienen mayor y menor rotación de inventario.

Palabras clave: Productos; Sistemas de recomendación; Lógica difusa; Filtrado colaborativo; Ventas; KNN; C Means.

Abstract

Information technologies are widely recognized as strategic instruments, due to their ability to modify the structure and business models of organizations, currently the large amount of information that an individual must handle in their day to day makes it impossible to optimize time. In the search for desired information, recommendation systems are currently a tool that helps decision-making and that little by little has been involved in the market and giving companies opportunities to be more competitive. The main objective of this research is to design a diffuse recommendation system to be able to recommend products from the company Promyvent to its clients according to the profile of each one of them. The methodology used in this research began with the pre-processing of the input database, preparation and cleaning techniques were applied, and an exploratory data analysis was applied, which allows a descriptive analysis, transformation and cleaning of data. I applied the jambu elbow method and it was established that the appropriate number of clusters was three, then in the modeling phase the KNN algorithm was applied to classify each of the data according to the purchases of the customers with the least calculated distance, then To increase the level of precision in data classification, the C Means method was applied, in this way an optimal fuzzy partition is obtained, granting a degree of membership to each data point. Finally, as a result, the recommendation is generated from the profile of each client, the accuracy of the model is based on the historical records of customer purchases in order to identify the degree of membership that each client has according to their profile and starting From this data, the model generated the filtering of a top seven products in the recommendations of main interest. The implemented model benefits the company not only by recommending a top seven products of interest according to the customer's profile, but also by identifying the products that have the highest and lowest inventory turnover.

Keywords: Products; recommendation systems; Diffuse logic; Collaborative filtering; Sales; KNN; C Means.

Resumo

As tecnologias da informação são amplamente reconhecidas como instrumentos estratégicos, devido a sua capacidade de modificar a estrutura e os modelos de negócios das organizações, atualmente a grande quantidade de informações que um indivíduo deve manipular no seu dia a

dia impossibilita a otimização do tempo na busca pelo informações, os sistemas de recomendação são atualmente uma ferramenta que auxilia a tomada de decisão e que aos poucos vem se envolvendo no mercado e dando às empresas oportunidades de serem mais competitivas. O objetivo principal desta pesquisa é projetar um sistema de recomendação difuso para poder recomendar produtos da empresa Promyvent aos seus clientes de acordo com o perfil de cada um deles. A metodologia utilizada nesta pesquisa começou com o pré-processamento do banco de dados de entrada, foram aplicadas técnicas de preparo e limpeza, e foi aplicada uma análise exploratória dos dados, que permite uma análise descritiva, transformação e limpeza dos dados. e foi estabelecido que o número adequado de clusters era três, então na fase de modelagem foi aplicado o algoritmo KNN para classificar cada um dos dados de acordo com as compras dos clientes com a menor distância calculada, então para aumentar o nível de precisão em Na classificação dos dados, foi aplicado o método C Means, desta forma é obtida uma partição fuzzy ótima, concedendo um grau de pertinência a cada ponto de dados. Por fim, como resultado, a recomendação é gerada a partir do perfil de cada cliente, a acurácia do modelo é baseada nos registros históricos de compras dos clientes a fim de identificar o grau de adesão que cada cliente possui de acordo com seu perfil e a partir de Com esses dados, o modelo gerou a filtragem de um dos sete principais produtos nas recomendações de interesse principal. O modelo implementado beneficia a empresa não só por recomendar os sete principais produtos de interesse de acordo com o perfil do cliente, mas também por identificar os produtos que apresentam maior e menor giro de estoque.

Palavras-chave: Produtos; sistemas de recomendação; Lógica difusa; Filtragem colaborativa; Vendas; KNN; C significa.

Introducción

De Las profundas transformaciones experimentadas en el entorno empresarial desde finales de los sesenta han introducido dificultades adicionales en los procesos decisionales de las empresas, en los que las unidades económicas aplicaban métodos y modelos de gestión explorativos se han pasado a encontrar estructuras de encuadramiento institucionales, sociales, económicas, tecnológicas y socioculturales en constante transformación (Esteban, 2017), el uso de las tecnologías a nivel mundial ha ido creciendo a través de los años gracias a la constante

investigación científica y tecnológica, debido a que existen muchas maneras de usar las tecnologías en esta era de transformación digital y muchos países han ido adoptando formas en las cuales pueden realizar eficientemente trabajos, cometiendo menos errores (Kishnani García, 2020).

Ante este nuevo panorama las empresas necesitan conocer métodos de obtención de información y técnicas para tratarla por lo cual es indispensable buscar ventajas competitivas y por ende un desarrollo económico a largo plazo (Esteban, 2017), en la actualidad existen multitudinarios cambios en los mercados, organizaciones, tecnologías, sociedades y culturas, razón por la cual se considera poco pertinente seguir maniobrando bajo el mismo enfoque tradicional, para lograr ser competitivo dentro de este entorno tan cargado de dinamismo y turbulencia (Pita, 2018), de manera que es cada vez mayor la cantidad de organizaciones empresariales que confían en sistemas de información para llevar a cabo el manejo de sus operaciones, interactuar con proveedores y clientes y competir en el mercado en el que se desenvuelven (Proaño, 2018).

Las tecnologías basadas en la Inteligencia Artificial (IA) ya están siendo utilizadas para ayudar a los humanos a beneficiarse de mejoras significativas y disfrutar de una mayor eficiencia en casi todos los ámbitos de la vida por lo que se puede aplicar en casi todas las situaciones (Rouhiainen, 2018), por esta razón los sistemas de recomendación son actualmente soluciones exitosas para facilitar el acceso de los usuarios a la información que se ajusta a sus preferencias y necesidades en espacios de búsqueda sobrecargados, además se han desarrollado varias metodologías para mejorar su desempeño (Yera, 2017), en consecuencia un Sistema de Recomendación puede brindar información valiosa para asistir en el proceso de toma de decisiones del consumidor con el objetivo de proveer una recomendación con exactitud y de manera precisa, ya que las recomendaciones que se producen ayudan a los seres humanos a satisfacer sus gustos personales y descubrir nuevos elementos, con menos esfuerzo, que si realizaran la actividad de manera manual (Pérez, 2018), también permiten agrupar productos o servicios similares que sean de interés para el consumidor y aumentar su fidelización. La fidelización del cliente se logra mostrando que la empresa tiene en cuenta las distintas necesidades y preferencias del consumidor y ayudándolo a orientar su búsqueda hacia los productos que más le interesan, en razón de que el sistema ha “aprendido” sus gustos y necesidades, lo que naturalmente aumenta la satisfacción del cliente (Chesñevar, 2018).

Este artículo tiene como objetivo desarrollar un modelo de sistema de recomendación difusa de productos, los mismos que proyectan determinar los factores que permitan identificar el perfil de los clientes de la empresa y definir los escenarios para poder dar las recomendaciones productos que sean de interés del cliente para facilitar la labor de venta y la evaluación del modelo a través de los experimentos realizados con nuevos clientes para optimizar el proceso y la recomendación sea más eficiente. Los resultados en el proceso se garantizan mediante la utilización de algoritmos de agrupación knn (vecino más cercano) y c means los cuales establecen la precisión del modelo en base a los registros históricos de compras de los clientes para poder identificar el grado de membresía que tiene cada cliente de acuerdo a su perfil y poderlo clasificar a un determinado cluster, cada cliente se encasilla dependiendo del tipo de clase que sean los vecinos más próximos a él y que generalmente su similitud es medida mediante una función de distancia euclidiana o manhattan.

De esta manera poder recomendar un top siete de productos de acuerdo a el perfil de cada cliente, recomendaciones que podrán ser utilizadas mediante el uso de aplicaciones móviles o páginas web, las cuales serán herramientas que ayudaran a mejorar el desempeño de los vendedores en el proceso de ventas, cabe destacar que las técnicas y modelos aplicados en esta investigación son dinámicos, es decir que el desempeño de la recomendación varía dependiendo de los registros históricos de cada cliente, pues este puede ir evolucionando con el tiempo y las nuevas tendencias. Además este modelo beneficia a la empresa no solo a recomendar un top siete de productos de interés de acuerdo al perfil del cliente si no también identificar los productos que tienen mayor y menor rotación de inventario, también se lo puede aplicar en otras áreas de la empresa como en departamentos crédito, marketing, Inventarios, entre otros para impulsar su desarrollo.

Base teórica

Evolución e innovación tecnológica empresarial

La evolución digital, apresurada por la situación actual de la pandemia del covid-19, nos ha inclinado a rediseñar nuestra forma de trabajar, influyendo en la forma que vivimos dando como efecto la apertura a nuevos modelos de negocios, y cambios en el mercado laboral (Chuya, 2021),

En la actualidad existen multitudinarios cambios en los mercados, organizaciones, tecnologías, sociedades y culturas, razón por la cual se considera poco pertinente seguir maniobrando bajo el mismo enfoque tradicional, para lograr ser competitivo dentro de este entorno tan cargado de dinamismo y turbulencia (Pita, 2018), en el contexto de una empresa, los sistemas de información cumplen los objetivos de cualquier otro sistema en general, tales como: el procesamiento de entradas, el almacenamiento de datos relacionados con la entidad y la producción de reportes y otro tipo de instrumentos de resumen de datos (Proaño, 2018).

De modo que en un mercado globalizado es necesario que las empresas busquen soluciones basadas en las nuevas tecnologías, de forma que les permita proponer cambios trascendentales en su modelo de negocio, las tecnologías de información y comunicación hoy por hoy se consideran parte significativa en la estrategia competitiva de la empresa, la operatividad de toda organización se respalda en el uso adecuado de tecnologías, a través de las cuales es posible bajar costos, mejorando la calidad en la producción con el uso eficiente de recursos en todas las áreas de la empresa (Costa, 2018).

En los últimos años hemos visto una explosión de nuevas tecnologías que prometen cambiar la forma en que vivimos y definen a la innovación como el motor de desarrollo socioeconómico y al emprendimiento como la herramienta para la consolidación de estos procesos (Camino, 2017), así mismo la tecnología, a través de los años, ha adquirido un papel importante en la manera como la humanidad realiza sus tareas cotidianas, en otras palabras, si se analiza, en los últimos cien años todas las sociedades desarrolladas y en vía de desarrollo han tenido que hacer frente a una revolución tecnológica en los diferentes sectores primarios, secundarios y terciarios. El origen de esta revolución es el aprovechamiento y el deseo de conseguir totalmente ventajas sobre los competidores de cada sector (PRIETO, 2018).

En este contexto estos sistemas son muy atractivos en situaciones donde la cantidad de información que se ofrece al usuario supera ampliamente cualquier capacidad individual de exploración (Morales, 2018), la organización o institución, está creada y formada por una persona o conjunto de personas, que se dedica a una actividad económica concreta para conseguir unos objetivos que luego se conviertan en beneficios (Mero-Vélez, 2018)

De este modo las organizaciones en orden con el desarrollo científico y tecnológico, han modificado sus estructuras en función a los cambios requeridos, entre ellos han centrado la atención al diseño de controles internos, que promuevan la eficiencia en la operaciones, la

salvaguarda de activos y la reducción de los riesgos, garantizando fiabilidad de los estados financieros en observancia al cumplimiento de las leyes y normas vigentes (Jara, 2017)

Es por ello que la problemática que presenta la empresa Promyvent, el crecimiento tecnológico y las exigencias actuales en el sector comercial, establece la importancia de desarrollar un sistema de recomendaciones que le permitan alcanzar un desempeño óptimo y así evolucionar para lograr satisfacer las necesidades de sus clientes y garantizar su competitividad y permanencia en un mercado; en constante cambio. Es por ello que la innovación es entendida como la concepción y establecimiento de cambios significativos en el producto, el proceso, el marketing o la organización de la empresa, con el propósito de mejorar sus resultados (Altamirano, 2020), es el pequeño o gran cambio que determinará el beneficio, el interés y la acumulación de riqueza va en conjunto con el talento humano, la experiencia y aprendizaje, con lo que coexistirá siempre y cuando, en su conjunto, la demanda del bien a producir sea alta y en base a estudios de comportamiento del consumidor (González, 2018)

Sistema de recomendaciones

Debido al crecimiento exponencial de los datos en la era de la información moderna, existe una gran sobre carga de información y se estima que para el 2025, se creen 463 exabytes de datos a nivel mundial todos los días. Aunque las empresas cobran estos datos generados, más de la mitad de estos datos no se utilizan debido a la falta de herramientas útiles y de las habilidades (Choenyi, 2021), por lo cual los Sistemas de Recomendación son ampliamente utilizados en la recomendación de productos en diferentes plataformas, los algoritmos que trabajan las recomendaciones son sofisticados y pueden requerir esfuerzos computacionales muy altos cuando trabajan en entornos de muchos usuarios y datos (Holguera, 2021)

Es una subclase de un sistema de filtrado, el cual pretende predecir el grado de aceptación de un usuario hacia determinado objeto de esa manera poder realizarle recomendaciones a dicho usuario sobre objetos que a este le puedan interesar (Cuenca Mogro, 2019), de esta manera la función principal de los sistemas recomendados depende de los objetivos que hayan sido planteados al momento de su definición e implementación; pero en términos generales suelen utilizar tres tipos de objetos 1 Recopilación de información 2 Extracción del conocimiento 3 La Recomendación (Gómez, 2019), por esta razón la tecnología de recomendación es una parte

importante debido a que puede brindar un mejor servicio a los usuarios y ayudarlos a obtener información en cualquier momento y en cualquier lugar, el sistema de recomendaciones debe proporcionar a los usuarios información rápida y precisa (Cui, 2020) .

Por tanto la importancia de los sistemas recomendaciones ha incrementado con el pasar de los años. Los beneficios obtenidos con la implementación de los mismos han sido exitosos debido a sus implementaciones en el campo de la ciencia, la educación, la ingeniería, la medicina, etc. Como se puede evidenciar en esta última década la tecnología ha cambiado de forma positiva diferentes ámbitos académicos, sociales y culturales de las personas (Charnelli, 2017), es por ello que la innovación es considerada como la variable clave para el éxito de las empresas, ya que estas se desenvuelven en un ambiente dinámico que aumenta constantemente sus exigencias y demanda tanto en mercados nacionales como internacionales (Polo Otero, 2017)

El sistema de recomendaciones se puede realizar mediante diversas técnicas como filtrado basado en contenido, el cual se utilizan filtros basados en colaboración, esta técnica tiene en cuenta la perfil que se construye en base a su anteriores calificaciones las mismas calificaciones que determinan su inclinación e intereses, formando la base para recomendar un nuevo artículo, una calificación más alta denota una mayor probabilidad del usuario para visitar artículos similares por lo tanto, se recomienda un nuevo artículo de acuerdo con el número máximo de calificaciones dadas por el usuario en un género (Keerthana, 2021), por tanto el filtrado colaborativo es un método para hacer predicciones sobre los intereses de un usuario basándose en información sobre las preferencias de otros usuarios. La suposición subyacente es que si la persona A tiene la misma opinión que la persona B sobre algunos temas, entonces es más probable que A al tener opiniones similares a las de B sobre diferentes temas (Kaushik, 2018).

Los sistemas de filtrado de información están dirigidos a personas que carecen de la experiencia, la capacidad o el tiempo necesarios para evaluar la inmensa cantidad de tópicos que generalmente están a su disposición en un sitio Web y se basan en la premisa de que, habitualmente, la gente tiene en cuenta las recomendaciones provistas por otros para tomar decisiones rutinarias (Pacheco Pazmiño, 2018)

Fuzzy C-Means

El ruido y los valores atípicos contribuyen al procedimiento computacional de los algoritmos, así como a los puntos de datos reales que conducen a centros de clúster inexactos y fuera de lugar, El

algoritmo Fuzzy C-Means (FCM) que adolece de estos problemas es el algoritmo de agrupamiento difuso más popular y ha sido objeto de numerosas investigaciones y desarrollos (Askari, 2021), el cual permite extender la clasificación nítida en una noción de clasificación difusa usando la medida de la suma de todos los errores cuadráticos ponderados y la distancia entre el centro se maximiza.

Podemos asignar membresía a los diversos puntos de datos en cada conjunto difuso, el resultado del algoritmo depende de la partición inicial lo cual nos puede llevar a un mínimo de la función objetivo, los resultado otorgan valores de pertenencia son números menor o igual a 1. Por lo contrario, como los valores de agrupamiento se vuelven duros, es decir 0 o 1, es decir, controla el grado de participación compartida entre grupos difusos (Flores Macalupú, 2022)

KNN

El KNN o vecino más cercano es uno de los clasificadores basados en vecindad más populares en el aprendizaje automático, dado su simplicidad y eficiencia para detectar y clasificar elementos en categorías, el parámetro k en KNN hace referencia al número de vecinos con el cual se define la pertenencia a una categoría, este parámetro usualmente se determina empíricamente, dependiendo del problema se prueba con diferentes valores de K , eligiendo el parámetro con el mejor desempeño en precisión (De-La-Hoz, 2019).

Es por ello que este algoritmo es catalogado como clasificador basado en instancias. Para clasificar, compara las instancias no vistas con aquellas etiquetadas del conjunto de entrenamiento utilizando una función de similitud y generalmente la similitud es medida mediante una función de distancia Euclidiana o Manhattan (Maillo, 2018)

Análisis de componentes principales

El análisis de componentes principales (ACP) es una técnica de análisis multivariante que permite la reducción de la dimensionalidad de la base de datos y mejor interpretación de la misma, con una mínima pérdida de información. Desde el punto de vista matemático los componentes principales son combinaciones lineales de las variables originales de manera que sean ortogonales entre sí. La búsqueda de los componentes principales se reduce a la solución de un problema de auto vectores y auto valores (de la Maza, 2019).

El método análisis por componentes principales consiste en la representación de las medidas numéricas de varias variables en un sitio de espacio reducido donde se puedan percibir las relaciones que de otra forma, permanecerían no visibles en dimensiones superiores tiene como objetivo transformar un conjunto de variables originales, en un nuevo conjunto de variables las cuales se caracterizan por estar correlacionadas entre sí (Gallego, 2019)

El codo de Jambu

El método del codo es un método que se utiliza para producir información para determinar el mejor número de grupos al observar el porcentaje de la comparación entre el número de grupos que formarán un codo en un punto. Este método proporciona ideas/ideas seleccionando valores de conglomerados y luego agregando el valor del conglomerado que se usará como modelo de datos para determinar el mejor conglomerado (Nainggolan, 2019).

En este contexto es un método visual para probar la consistencia del mejor número de clústeres, por ende la idea es determinar el número de conglomerados, luego sumar los conglomerados, calcular el error cuadrático de la suma por conglomerado hasta el número máximo de conglomerados que se haya determinado, luego comparando la diferencia del error cuadrático de la suma de cada conglomerado, la diferencia más extrema forma el ángulo del codo muestra el mejor número de grupos (Umargono, 2020)

Nube de Palabras

Una nube de palabras es una representación gráfica vívida de las palabras delimitadas en una sección de texto generada por herramientas orientadas a la web y es un método de visualización de texto sencillo y visualmente atractivo. Tiene una utilidad versátil para proporcionar una visión general fraccionando el texto a aquellas palabras con mayor frecuencia (Padmanandam, 2021), son una tecnología emergente para el análisis de datos, se recomienda utilizar nubes de palabras como tableros porque brindan un análisis fácil, rápido y significativo de datos cualitativos además las nubes de palabras resumen o analizan datos de texto y brindan interpretaciones significativas a través del tamaño y el color del texto. (Sellars, 2018)

Metodología

La perspectiva de la recomendación parte por la necesidad de la innovación y automatización empresarial con el fin de satisfacer las demandas de los clientes. Se refiere al cliente y a la empresa, donde el cliente tiene la necesidad de adquirir productos que sean de su interés y la empresa trata de cubrir esas necesidades. De esta manera, el conjunto de datos será recopilado en la base de datos de la empresa, información proporcionada por los agentes vendedores que visitan al cliente de manera directa en distintos cantones de la provincia de Manabí, teniendo en cuenta que cada cantón presenta características distintas por población, economía y ubicación geográfica siendo en la ciudad de Portoviejo donde presenta un mayor número de clientes como se detalla en la (Fig. 1), el cual se describirá el comportamiento y características del perfil de los clientes ya que debido a la cantidad de productos que distribuye la empresa esta posee clientes con distintos tipos de negocios lo que hace que cada uno tenga necesidades distintas teniendo en cuenta que la mayor cantidad de clientes que posee son Comerciales como se puede observar en la (Fig. 2).

Figura 1: Base de datos Promyvent Clientes por Cantón

	ciudad	Record Count ▾
1.	PORTOVIEJO	431
2.	CALDERON	239
3.	CHONE	224
4.	JUNIN	203
5.	SANTA ANA	198
6.	CALCETA	186
7.	CANUTO	144
8.	SAN VICENTE	87
9.	LEONIDAS PLAZA	86
10.	CHARAPOTO	81
11.	ROCAFUERTE	62
12.	BAHIA	49
13.	COLON	43
14.	RIO CHICO	23
15.	SAN CLEMENTE	21
16.	SAN JACINTO	17
17.	PUEBLO NUEVO	11
18.	PRTO	1
19.	MANTA	1

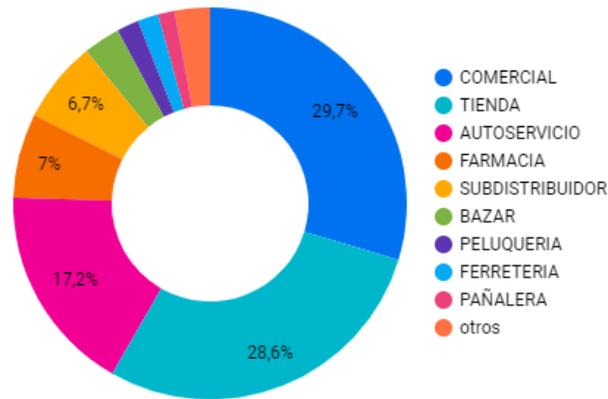


Figura 2: Distribución de Tipo de Negocio

Diseño del modelo

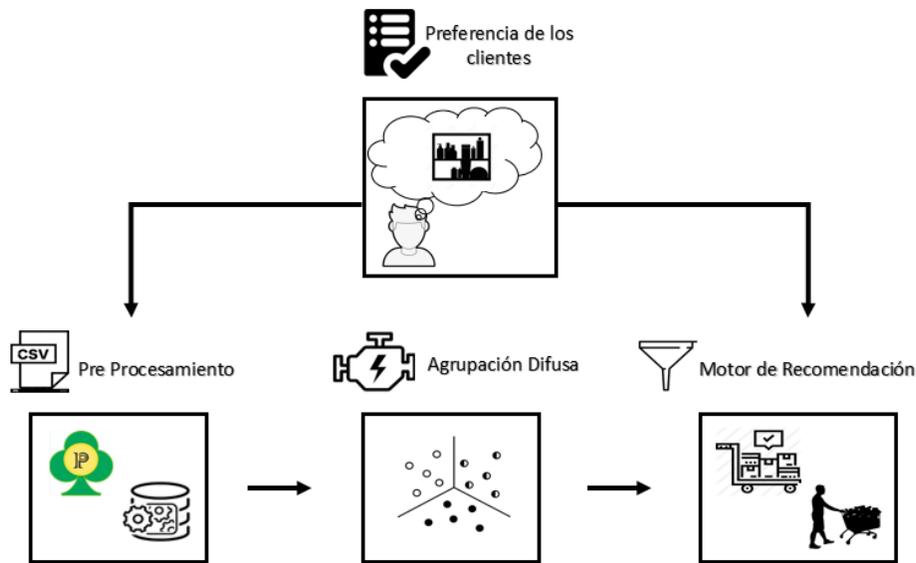


Figura 3: Diseño del Modelo

El diseño en la Figura 3 describe el sistema de recomendación de productos que realiza el análisis descriptivo, predictivo y prescriptivo. Consta de varias fases que son esenciales para el desarrollo del modelo; la primera fase es la recopilación de las necesidades del cliente de acuerdo a su tipo de negocio Figura 2, la segunda fase se realiza el pre procesamiento de la información, donde se aplica técnicas de transformación y limpieza de datos mediante el uso de Python. La tercera fase de modelado, es la agrupación de vecinos aplicando técnicas de aprendizaje automático como

Media C Difusa, KNN (Vecinos más cercanos), análisis de componentes principales, función Test. La cuarta fase comprende a las recomendaciones de productos mediante las configuraciones de las especificaciones difusas en el perfil del cliente que podrán ser utilizadas por aplicaciones móviles o páginas web. En este caso utilizamos la información de la Base de Datos de Promyvent Figura 1.

Pre procesamiento

En esta fase comienza un pre procesamiento de datos de entrada. Se le aplica técnicas de preparación y limpieza de datos mediante el uso de Phyton, además se aplica un análisis Exploratorio de Datos el cual permite realizar un análisis descriptivos, transformación y limpieza de datos, (en este proceso se eliminaron los registros que no tienen datos los cuales serían las transacciones anuladas las cuales presentaban valores en cero y no serían de utilidad en esta investigación) y de esta forma obtener datos más precisos y confiables, Luego de la limpieza, estos datos están preparados para ser modelar el sistema de recomendación difuso, además mediante la utilización del método del codo de jambu, el cual utiliza los valores de la inercia obtenidos tras aplicar el K-means a diferente número de Clusters, siendo la inercia la suma de las distancias al cuadrado de cada objeto del Cluster a su centroide como se puede observar en la Figura 4 a todos los datos determina que número adecuado de clústeres a generar serian tres en el cual encasilla a cada uno de los clientes de acuerdo a su perfil. En la tabla 1 podemos apreciar las variables y los datos que servirán de entradas para todos los procesos de recomendación de productos los cuales serán filtrados para obtener el final la recomendación.

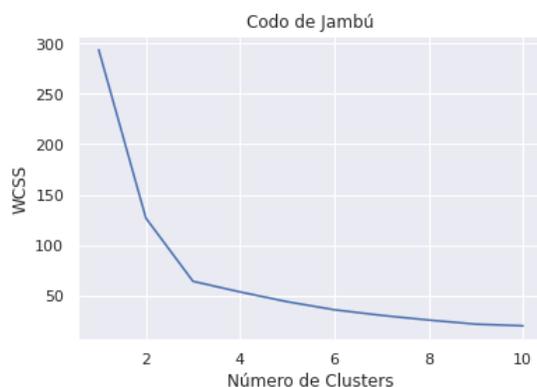


Figura 4: Número Adecuado de Clústeres

Figura 5: Grupo de Variables de entrada

	cliente	ciudad	valor	tipo.cliente	forma de pago	días de 1	descuentos especiales
0	10042	8	50.25	13	1	2	2
1	10081	11	184.98	4	1	3	1
2	10099	8	181.63	4	1	3	1
3	10104	2	103.63	7	1	3	1
4	10108	5	53.44	12	1	2	2

$$\text{Minimizar} \left\{ \sum_{k=1}^k \frac{1}{|c_k|} \sum_{i,i \in c_k} \sum_{k,j=L}^P (x_{ijp} - x_i'j)^2 \right\} \quad (1)$$

La Ecuación 1 se emplea para determinar los grupos creados de acuerdo al registro de las características del grupo de variables, este proceso se lleva a cabo asignando centroides aleatorios, los cuales se actualizan a los centros de cada grupo, con sus distancias hasta que la asignación de registros queda fija.

$$w_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (2)$$

La ecuación 2 es la función de pertenencia, que utiliza los centroides que se encuentran en la ecuación 1 para calcular el grado de relevancia de cada registro para el grupo. Se obtiene una muestra restringida filtrando los registros, ajustando un umbral del grado de pertenencia y normalizándolo.

$$\text{Euclidiano} \sqrt{\sum_{i=1}^M (y_i - x_i)^2} \quad (3)$$

Modelado

Es un algoritmo de aprendizaje difuso no supervisado que asigna grados de membresía a cada uno de los puntos de datos. La estructura del algoritmo agrupa a los clientes por sus preferencias, gustos y necesidades relacionadas expectativas de los productos de interés. En el entrenamiento de este modelo se recomiendan productos según las compras de los clientes con menor distancia

calculada por el método knn que se aplica en la Ecuación 3, creando tres clusters Figura 3 que posteriormente para aumentar la precisión se aplica el método C Means que caracteriza la participación de cada muestra en todos los grupos utilizando funciones de pertenencia que toman valores entre cero y uno, lo cual indicara que los valores cercanos a uno indican una mayor similitud, mientras que los cercanos a cero indican una menor similitud, de esta manera se obtiene una partición difusa optima otorgando un grado de membresía a cada punto de datos.

Motor de Recomendación

En esta fase las preferencias, gustos y necesidades del cliente sobre los productos son considerados mediante la aplicación que usa un algoritmo difuso que asocia las características del historial de compra de los clientes. De esta forma estas características mediante reglas de asociación se guardan en la memoria del motor de Recomendación. Allí estas características sufren una modificación previa según las preferencias en compras anteriores, en caso de no hallar similitud con compras anteriores se dirige al modelo borroso para agruparlo al grupo que potencialmente pertenece con un grado de membresía que indican el grado en que cada punto de datos pertenece a cada grupo. Supongamos que en la búsqueda de los grados de membresía tenemos estos valores 0.40, 0.74, 0.32, claramente podemos determinar que el grado de pertenencia se encuentra en el clusters 2. Una vez identificado el grado de pertenencia del perfil de los clientes se filtran las recomendaciones aplicando la función test y aplicando el algoritmo KNN calculado en la Ecuación 3. El valor k se fija en siete, y el sistema nos recomienda un top siete de productos que podrán ser consumidos por aplicaciones móviles o páginas web.

Resultados

La investigación se realizó con la información de los clientes de la empresa Promyvent la cual se dedica a distribuir una gran variedad de productos en distintos cantones de la provincia de Manabí Figura 1, teniendo así una extensa cartera de clientes, cada uno con preferencias gustos y necesidades distintas, se utilizó el 100% de datos de los clientes de la base de datos de la empresa y los registros de compra de todos los clientes como se observa en la Tabla 2, estos datos se los trasladó y almaceno en una hoja de cálculo de Excel extensión CVS para posteriormente estos datos sean manipulados por el software Phytion con mayor facilidad.

Como resultado de esta investigación se pudo recomendar un top siete de productos de acuerdo a el perfil de cada cliente, la recomendación se podrá utilizar mediante el uso de aplicaciones móviles o páginas web que ayudaran a mejorar el desempeño de los vendedores en el proceso de ventas, cabe destacar que las técnicas y modelos aplicados son dinámicos, es decir que el desempeño de la recomendación varía dependiendo de los registros históricos de cada cliente, pues este puede ir evolucionando con el tiempo y las nuevas tendencias.

Los resultados en el proceso mediante la utilización de algoritmos de Agrupación KNN (vecino más cercano) y C Means establecen la precisión del modelo en base a los registros históricos de compras de los clientes para poder identificar el grado de membresía que tiene cada cliente de acuerdo a su perfil y poderlo clasificar a un determinado cluster, cada cliente es encasillado dependiendo del tipo de clase que sean los vecinos más próximos a él y que generalmente su similitud es medida mediante una función de distancia Euclidiana o Manhattan como podemos observar en la Figura 5 que un cliente de acuerdo a sus características fue encasillado en el clusters 2, luego esta información se filtra de acuerdo al perfil de cada cliente recomendando un top siete de productos de interés el cual lo hemos representado en una nube de palabras la cual brinda un análisis fácil, rápido y comprensivo de los datos en el cual los nombres de los productos recomendados aparecen más grandes y de distinto color como se aprecia en la Figura 6. También se estableció en base a la frecuencia del historial de compras de cada cliente un top diez de productos más vendidos y con mayor aceptación permitiendo de esta manera que la empresa pueda conocer cuáles son los productos con mayor rotación en el mercado y así poder tomar una mejor decisión al ofrecer el producto a su cliente como se ilustra en la Figura 7. Con el uso de 6 variables y 8000 registros, los datos fueron pre procesados mediante el uso algoritmo de aprendizaje difuso no supervisado.

Figura 6: Registro de Compras de Clientes

	codigo	descripcion	uni_med	cantidad	valor	codigo_cliente	cuidad
0	1 0000003	LECHE CREMA NO LACTEA FANS 1 KG.	UNI	42	253,5	10042	JUNIN
1	1 0010128	ACOND. SAVITAL BIOTINA Y SABILA 530ML.	UNI	1	4,09	10042	JUNIN
2	1 0010137	TRATAMIENTO NUTRIBELA REPOLARIZACION 300ML	UNI	1	4,91	10042	JUNIN
3	1 0010150	SH. SAVITAL 550ML. BIOTINA	UNI	1	4,09	10042	JUNIN
4	1 0010151	SH. SAVITAL 550ML. KERATINA	UNI	1	4,09	10042	JUNIN
5	1 0010152	SH. SAVITAL 550ML. MULTIVITAMINA	UNI	7	27,53	10042	JUNIN
6	1 0010165	SH. SAVITAL COLAGENO Y SABILA 550ML.	UNI	1	4,09	10042	JUNIN
7	1 0010168	ACOND. SAVITAL MULTIVITAMINAS 530ML.	UNI	1	4,09	10042	JUNIN
8	1 0010173	TRATAMIENTO NUTRIBELA BIOKERATINA 300ML	UNI	1	4,91	10042	JUNIN
9	1 0011030	ACEITE ROJO PARA MUEBLE MOEBEL 240ML.	FCO	10	19,2	10042	JUNIN
10	1 0041006	PILA SONY CARBON X 1 AAA	UNI	40	8	10042	JUNIN
11	1 0120120	HOJA DE AFEITAR GUILLETE 1X5	UNI	40	24	10042	JUNIN
12	1 0160050	ACEITE JOHNSON BABY 50 ML.	UNI	7	8,68	10042	JUNIN
13	1 0160140	COTONETES MADERA CARLITOS 3 x100	UNI	26	10,32	10042	JUNIN
14	1 0222056	COLONIA PERSONI 100ML CON ATOMIZADOR	UNI	4	8,18	10042	JUNIN

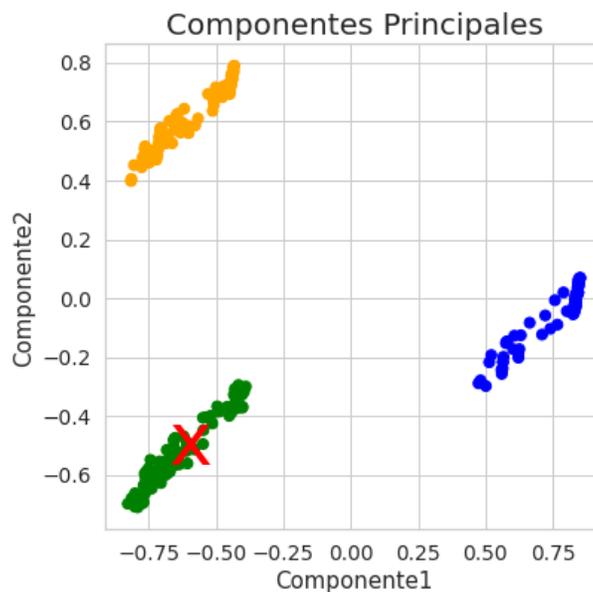


Figura 7: C Medias Grado de Membresía

importante debido a que puede brindar un mejor servicio a los usuarios y ayudarlos a obtener información en cualquier momento y en cualquier lugar.

La fase del pre procesamiento de datos es esencial en todo proceso, debido a que en esta fase se identifica valores atípicos que pueden alterar resultados y propiedades de los datos.

Uno de los inconvenientes que tuvimos en el desarrollo de esta investigación fue que las frecuencias absolutas arrojaban muchos ítems de productos con pocos valores lo cual dificultaba la presentación de los valores de la recomendación, por lo cual utilizamos una nube de palabras para hacer un gráfico con una mejor presentación.

El modelo implementado beneficia a la empresa no solo a recomendar un top siete de productos de interés de acuerdo al perfil del cliente si no también identificar los productos que tienen mayor y menor rotación de inventario, esto permite a la empresa lanzar campañas de marketing con los productos que menos acogida tienen entre los clientes. En esta investigación se implementó un sistema de recomendación difuso el cual fue probado y ejecutado a la base de datos de clientes de la empresa Promyvent.

En trabajos futuros este modelo de recomendaciones planteado puede ser más óptimo y obtener mejores resultados si la empresa implementa un sistema que permita a los clientes valorar los productos lo que permitirá optimizar el algoritmo y obtener resultados más precisos.

Este modelo de recomendaciones también se lo puede aplicar en otras áreas de la empresa como en departamentos crédito, marketing, Inventarios, entre otros para impulsar su desarrollo.

Referencias

1. Altamirano, K. L. (2020). Innovación en empresas de Cuenca-Ecuador: Empleo de modelística inteligente en el sector textil. *Revista de ciencias sociales*, 26(3), 148-162.
2. Askari, S. (2021). Algoritmo de agrupamiento Fuzzy C-Means para datos con tamaños de clúster desiguales y contaminados con ruido y valores atípicos: revisión y desarrollo. *Sistemas expertos con aplicaciones*, 165 , 113856.
3. Camino, R. C. (2017). Emprendimiento e innovación en Ecuador, análisis de ecosistemas empresariales para la consolidación de pequeñas y medianas empresas. *INNOVA Research Journal*, 2(9.1), 73-87.

4. Charnelli, M. E. (2017). Sistema recomendador basado en tópicos latentes. *In Congreso Argentino de Ciencias de la Computación*, Vol 23.
5. Chesñear, C. I. (2018). El comercio electrónico en la era de los bots. *Universidad Nacional del Sur y CONICET*, 126-135.
6. Choenyi, T. T. (2021). A review on the filtering techniques used in the restaurant recommendation system. *In t. J. Computation. Science Crowd. Compute*, 10 (4), 113-117.
7. Chuya, J. C. (2021). Economía digital, herramienta para mejorar la competitividad y productividad en las PYMES caso: Machala-Ecuador. *593 Digital Publisher CEIT*, 6(3), 76-86.
8. Costa, M. P. (2018). Inversión en TICS en las empresas del Ecuador para el fortalecimiento de la gestión empresarial Periodo de análisis. *Revista Espacios*, 39(47).
9. Cuenca Mogro, E. M. (2019). *Investigación y análisis de sistemas de recomendaciones basados en reglas y Fuzzy para el control de la diabetes sobre artículos científicos y sistemas existentes*. Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas., Guayaquil.
10. Cui, Z. X. (2020). Sistema de recomendación personalizado basado en filtrado colaborativo para escenarios IoT. *IEEE Transactions on Services Computing* , 13 (4), 685-695.
11. de la Maza, L. L. (2019). Análisis de componentes principales aplicado a la fermentación alcohólica. *Revista Científica de la UCSA*, 6(2), 11-19.
12. De-La-Hoz, E. J.-L.-H. (2019). Metodología de aprendizaje automático para la clasificación y predicción de usuarios en ambientes virtuales de educación. *Información tecnológica*, 30(1), 247-254.
13. Esteban, I. G. (2017). *Fundamentos y técnicas de investigación comercial*. Esic Editorial.
14. Flores Macalupú, L. E. (2022). Una introducción a la clasificación difusa. *Una introducción a la clasificación difusa*. Universidad Nacional de Piura, Piura.
15. Gallego, L. &. (2019). Variables de influencia en la capacidad de aprendizaje. un análisis por conglomerados y componentes principales. *Información tecnológica*, 30(2), 257-264.
16. Gómez, P. G. (2019). Sistemas de Recomendación: un enfoque a las técnicas de filtrado. *Revista Ibérica de Sistemas e Tecnologias de Informação*, 286-293.

17. González, J. &. (03 de 2018). Innovación tecnológica en las empresas. *Observatorio de la Economía Latinoamericana*.
18. Holguera, F. P. (2021). *Sistemas de recomendación basados en filtrado colaborativo:: aceleración mediante computación reconfigurable y aplicaciones predictivas sensoriales*. Universidad de Extremadura.
19. Jara, A. A. (2017). Evaluación comparativa del sistema de control interno del sector comercial y del sector público del Cantón Morona. *Killkana sociales: Revista de Investigación Científica*, 1(1), 31-38.
20. Kaushik, A. G. (2018). Una recomendación de película. Sistema utilizando Redes Neuronales. *Revista internacional de investigación avanzada, ideas e innovaciones en tecnología*, 4 (2), 425-430.
21. Keerthana, T. B. (2021). Sistema de recomendación de productos Flipkart. *Journal of Engineering Sciences*, 33 , 34.
22. Kishnani García, S. F. (2020). *Uso de TIC y el desempeño laboral en empresas de comercio en Lima Metropolitana*. Universidad San Ignacio de Loyola, Lima.
23. Maillo, J. L. (2018). Un enfoque aproximado para acelerar el algoritmo de clasificación Fuzzy kNN para Big Data. *Asociación Española para la Inteligencia Artificial (AEPIA)*, (pp. 1143-1148).
24. Mero-Vélez, J. M. (2018). Empresa, administración y proceso administrativo. *Revista Científica FIPCAEC (Fomento de la investigación y publicación en Ciencias Administrativas, Económicas y Contables)*. ISSN: 2588-090X. *Polo de Capacitación, Investigación y Publicación (POCAIP)*, 3(8), 84-102.
25. Morales, C. A. (2018). Algoritmo SVD aplicado a los sistemas de recomendación en el comercio. *Tecnología Investigación y Academia*, 6(1), 18-27.
26. Nainggolan, R. P.-a. (2019). Se mejoró el rendimiento del clúster de K-means utilizando la suma de errores cuadráticos (SSE) optimizada mediante el uso del método Elbow. *En Journal of Physics: Conference Series*, (Vol. 1361, No. 1, p. 012015).
27. Pacheco Pazmiño, H. W. (2018). *Estudio de algoritmos de filtrado basado en contenidos para sistemas recomendadores de información*. Universidad Técnica de Cotopaxi: Facultad de Ciencias de la Ingeniería y Aplicadas), Latacunga.

28. Padmanandam, K. B. (2021). Una visualización de nube de palabras dinámica con reconocimiento de voz para resúmenes de texto. *En 2021 6th International Conference on Inventive Computation Technologies (ICICT) IEEE.*, (pp. 609-613).
29. Pérez, S. C. (2018). Análisis, diseño y desarrollo de un sistema de recomendación basado en datos restauranteros de TripAdvisor y Foursquare. *Research in Computing Science*, 147, 209-224.
30. Pita, G. E. (2018). Las TICs en las empresas: evolución de la tecnología y cambio estructural en las organizaciones. *Dominio de las Ciencias*, 4(1), 499-510.
31. Polo Otero, J. L. (2017). Relación entre I+ D, actividades innovadoras y resultados empresariales: un análisis para el sector de alimentos y bebidas en Colombia. *Dimensión Empresarial*, 15(1), 175-197.
32. PRIETO, J. E. (2018). Los beneficios de la inteligencia artificial en el sector empresarial. *Los beneficios de la inteligencia artificial en el sector empresarial*. UNIVERSIDAD SANTO TOMAS, Bogota.
33. Proaño, M. F. (2018). Los sistemas de información y su importancia en la transformación digital de la empresa actual. *Revista Espacios*, 39(45).
34. Rouhiainen, L. (2018). *Inteligencia artificial*. Madrid: Alienta Editorial.
35. Sellars, B. S.-A. (2018). Uso de nubes de palabras para analizar datos cualitativos en entornos clínicos. *Gestión de enfermería*, 49 (10), 51-53.
36. Umargono, E. S. (2020). Optimización de agrupamiento de K-medias utilizando el método del codo y la determinación temprana del centroide en función de la media y la mediana. *En Actas de las Conferencias Internacionales sobre Sistemas y Tecnologías de la Información-CONRIST*, (pp. 234-240).
37. Yera, R. y. (2017). Herramientas difusas en sistemas de recomendación: una encuesta. *Revista Internacional de Sistemas de Inteligencia Computacional*, 10 (1), 776-803.