

Sistema de recomendación de artículos de línea blanca basado en el algoritmo KNN

White line article recommendation system based on the KNN algorithm

 Guevara-Fernandez, Alexander^{1*}

 Coral-Ygnacio, Marco A.²

¹Universidad Católica Sedes Sapientiae, Nueva Cajamarca, Perú

²Universidad Católica Sedes Sapientiae, Lima, Perú

Recibido: 04 May. 2023 | **Aceptado:** 05 Jul. 2023 | **Publicado:** 10 Jul. 2023

Autor de correspondencia*: 2017101042@ucss.pe

Como citar este artículo: Guevara-Fernandez, A. & Coral-Ygnacio, M. A. (2023). Sistema de recomendación de artículos de línea blanca basado en el algoritmo KNN. *Revista Científica de Sistemas e Informática*, 3(2), e557. <https://doi.org/10.51252/rcsi.v3i2.557>

RESUMEN

En la presente investigación se busca mejorar el proceso del marketing digital para temas del comercio electrónico, su principal objetivo es implementar y poner en funcionamiento un sistema de recomendación que permita recomendar correctamente un producto a un cliente ahorrándole tiempo en su proceso de búsqueda y decisión. Se utiliza el algoritmo K vecinos más cercanos junto a su fórmula de la distancia euclidiana que permite mejorar la precisión al momento de brindar resultados. Para el presente caso se trabajó con las preferencias de un usuario y una cantidad de más de 100 productos de distintos modelos y funcionalidades que son identificadas mediante variables de identificación como, color, marca, modelo, precio que nos sirven para poder realizar el cálculo de la distancia y generar "N" recomendaciones más cercanas a los gustos del cliente, los resultados muestran que el algoritmo propuesto es eficiente en cuanto a la recomendación de productos logrando generar recomendaciones de manera eficiente en relación a las preferencias de los clientes.

Palabras clave: comercio electrónico; distancia euclidiana; K-vecinos; marketing digital

ABSTRACT

This research seeks to improve the digital marketing process for e-commerce issues, its main objective is to implement and operate a recommendation system that allows to correctly recommend a product to a customer saving time in their search and decision process. The K nearest neighbors' algorithm and its Euclidean distance formula are used to improve the accuracy of the results. For this case we worked with the preferences of a user and a quantity of more than 100 products of different models and functionalities that are identified by identification variables such as color, brand, model, price, which are used to calculate the distance and generate "N" recommendations closer to the customer's tastes, the results show that the proposed algorithm is efficient in terms of product recommendation, generating recommendations efficiently in relation to customer preferences.

Keywords: e-commerce; Euclidean distance; K-neighbors; digital marketing; digital marketing

1. INTRODUCCIÓN

Desde hace años los electrodomésticos han ayudado de manera muy directa a los seres humanos facilitándoles realizar de manera más rápida algunas tareas domésticas (Martínez Rodríguez & Alarcón Martínez, 2020). Hoy en día ante la implementación de tecnología inteligente en los productos de línea blanca, estos artículos han generado grandes ganancias en ventas a las empresas de este rubro (Sánchez, 2019),

Asimismo, los clientes ante la gran variedad de productos que se le ofrecen presentan algunas dificultades de decisión por la elección de sus productos ideales, generando de esta manera confusiones en los clientes y trayendo consigo la pérdida de interés de los compradores, afectando las ventas en las empresas y perdiendo potenciales clientes, de esta manera despertando el interés de las organizaciones por conocer a sus usuarios e investigar sobre los gustos y preferencias de estos mismos. A partir de sistemas de recomendación y algoritmos inteligentes que ayuden a filtrar las preferencias de los usuarios, se busca determinar alternativas para capturar la atención de los clientes y aumentar las ventas en las organizaciones.

Los sistemas de recomendación desde mediados de los años 90 se han convertido en una rama muy importante de la investigación, con el fin de mejorar la relación entre el cliente y la empresa mediante la aplicación de métodos, técnicas y algoritmos inteligentes para ayudar en la toma de decisiones, basándose en las recomendaciones a partir de los gustos y preferencias de los usuarios (Franco Zapata, 2021). En el año 2007 los autores Cao & Li (2007), propusieron un sistema de recomendación basado en fuzzy inteligente para productos electrónicos de consumo, utilizando la técnica de minería de datos y lógica difusa, el sistema se compone de cuatro módulos, por lo que en el primero identifica las necesidades de los clientes, en el segundo módulo se analiza las características de los productos, el tercero se basa en los requisitos de las características ideales y por último, el cuarto es el módulo de las recomendaciones.

Los sistemas de recomendación de productos de línea blanca involucran el uso de modelos matemáticos para poder generar recomendaciones a partir de la utilización de métodos, técnicas y algoritmos que se adapten a las necesidades de la organización, de esta manera poder analizar la gran cantidad de información que estas manejan en relación a los clientes y sus preferencias (Franco Zapata, 2021), por ello, se requiere conocer los métodos técnicas y algoritmos más eficientes para la implementación de estos sistemas.

Diversas investigaciones hacen énfasis que para la implementación de sistemas de recomendación de línea blanca existen técnicas y algoritmos de aprendizaje supervisado debido a su eficiencia al momento de generar recomendaciones (Criado González, 2018), asimismo, Castro et al. (2012), hacen énfasis a que la implementación de este tipo de sistemas de recomendación mejora las ventas en las organizaciones e incrementa las ganancias significativamente, mediante la personalización de productos, contenido e interfaces que son más amigables con los usuarios. Por otra parte, algunos sistemas de recomendación de productos de línea blanca presentan ineficiencias al momento de generar recomendaciones debido a una mala implementación de los algoritmos o un mal manejo de la técnica de implementación (Guevara Albán et al., 2018).

Debido a lo anteriormente mencionado podemos evidenciar que, así como existen diversos métodos de implementación eficientes, también hay sistemas que presentan deficiencias al generar recomendaciones, esto debido a un mal manejo del algoritmo o técnica al momento de ser implementados, lo que nos motiva a investigar los sistemas de recomendación existentes más eficientes para la venta de artículos en línea.

Por ello, se pretende conocer las diversas técnicas y algoritmos de recomendación que se utilizan en la implementación de estos sistemas en relación con la recomendación de productos de línea blanca, así

mismo se pretende conocer los diversos factores que influyen en la implementación de este tipo de sistemas.

En el presente trabajo se pretende construir un sistema de recomendación de productos de línea blanca utilizando el algoritmo de aprendizaje supervisado KNN de esta manera incrementar las ventas en las organizaciones, recomendar a los usuarios productos a sus gustos y preferencias mejorando de esta manera la relación entre la empresa y sus clientes.

El sistema propuesto utilizará una lógica basada en 5 pasos, los cuales ayudarán en el proceso de recomendación de productos, en el primer paso se realizara el ingreso de variables de las preferencias de los usuarios con una cantidad de doce variables, como segundo paso se hace el cargado de datos de la base de datos para hacer el cálculo de las distancias entre los productos de la base de datos y las preferencias del usuario, para luego realizar el cálculo de la distancia en relación a las variables entrantes con la de los productos, como paso cuatro se almacenarán y ordenarán las distancias, para luego en el paso cinco poder recomendar los “N” productos más cercanos a las preferencias del usuario.

Finalmente, se ha organizado el trabajo de la siguiente manera, el punto 1 detalla la introducción, el punto 2 muestra el estado del arte del tema, seguido del algoritmo de recomendación en el capítulo 3, así mismo en el capítulo 4, se presenta la propuesta de la solución y el software de recomendación y para finalizar los resultados, la discusión y las conclusiones y recomendaciones.

2. MATERIALES Y MÉTODOS

2.1. Sistemas de recomendación

Los sistemas de recomendación son una herramienta que ayudan al usuario proporcionando una lista de recursos o etiquetas que podrían ser de su agrado (Mlika & Karoui, 2020), desde sus inicios cuando empezaron a surgir los sistemas de recomendación como un área de investigación lo que se hacía era que la mayoría de investigadores aplicaban diversos patrones de valoración en el comportamiento de los usuario (Jain et al., 2020), lo que se consideraba para un sistema de recomendación básicamente era la precisión de sus recomendaciones, sin embargo, es importante definir aspectos muchos más allá ya que se presentan una diversidad de recomendaciones ubicadas por categorías (Cao et al., 2018).

2.2. Tipos de sistemas de recomendación

Los sistemas de recomendación se clasifican según el tipo por el cuál serán desarrollados, es decir, con la finalidad de la tarea que estos van a cumplir, para ello, detallamos a continuación algunos de los tipos de sistemas más utilizados en los sistemas de recomendación:

Basados en filtrado colaborativo

El filtrado colaborativo ha obtenido un gran éxito dentro del ámbito académico como en el industrial debido a que este modelo se basa principalmente en la valoración que se agrega cierta información cuando el usuario interactúa con el producto (Liu et al., 2021). Este modelo consiste en tener que encontrar dos o más usuarios similares (Li et al., 2020). El problema es que un sistema de filtrado colaborativo en muchos casos es debido a los escasos datos, ya que cada vez se venden más productos dentro de un sitio web (Zhang et al., 2020).

Basado en contenido

El filtrado basado en el contenido tiene por objetivo basarse de las preferencias del usuario y agrupar todos aquellos productos con atributos similares que se encuentran dentro de un conjunto de datos ya definido (Afoudi et al., 2021).

Basado en redes neuronales

La red neuronal por lo general está diseñada por varias capas siendo de entrada y de salida, cada una contiene información, así como también están conectadas entre ellas (Ravnik et al., 2021), son un conjunto de algoritmos los cuales se manejan por diversos patrones tomando el diseño del cerebro humano (Afoudi et al., 2021).

Sistema de recomendación híbrido

Los sistemas de recomendación híbrido son aquellas que combinan dos características es decir los perfiles y la membresía (Afoudi et al., 2021), pero diseñar un sistema de recomendación que sea seguro en un reto, basándonos en el modelo híbrido se ha buscado realizar un sistema de recomendación seguro de manera que busque cuidar los derechos de privacidad de información de los usuarios (Ni et al., 2018).

2.3. Algoritmos de recomendación

Un algoritmo es un conjunto de instrucciones que se ejecutan de manera secuencial es decir uno detrás de otro, son utilizados para realizar ciertas acciones con el fin de generar soluciones óptimas (Korus et al., 2021), para ello detallamos algunos algoritmos más utilizados en sistemas de recomendación.

El algoritmo K-means, este algoritmo es utilizado como estrategia de clustering para reducir el conjunto de entrenamiento y para reducir el tiempo de ejecución del algoritmo, así como también el coste computacional.

Los resultados experimentales muestran que el algoritmo presentado puede producir una serie de recomendaciones objetivas, cobertura sobre la base de la precisión y la diversidad (Cai et al., 2020), otro de los algoritmos más usados en sistemas de recomendación es el “algoritmo genético”, el uso de este algoritmo permite automatizar la búsqueda de cursos adaptados al perfil del usuario. En efecto asignamos diferentes rutas de aprendizaje a los usuarios que pertenecen a la misma clase, utilizando algoritmos genéticos para buscar un camino óptimo.

En conclusión, el uso de algoritmos genéticos nos permitirá automatizar la búsqueda de contenidos adaptados al perfil de un usuario (Hssina & Erritali, 2019), así mismo el “algoritmo Probs tradicional” es usado para generar la lista de recomendaciones final consideró una variedad de condiciones con diferente número de clústeres y niveles de dispersión, puede mejorar significativamente la puntuación del ranking y la cobertura de los sistemas de recomendación.

En resumen, el algoritmo Probs es capaz de mejorar la calidad de las recomendaciones (Zhang et al., 2020), a diferencia del “algoritmo KNN” que mediante la aplicación de la distancia euclidiana o la similitud del coseno entre los clústeres de entrenamiento y los de prueba, este algoritmo utiliza los vecinos más cercanos donde “K” esta diferenciado por la cantidad de vecinos que este tiene (Adeniyi et al., 2016)

Por último, “algoritmo Slope One” es muy utilizado con la técnica de filtrado colaborativo debido a la alta precisión que este genera en tiempo real, gracias a la separación de matrices de valoración que los usuarios generan, separando los productos que interesan al usuario con los que no lo son de interés y así generar recomendaciones más exactas debido a que es simple en cálculo y alto en rendimiento (Ye & Zhao, 2018).

2.4. Marketing digital

El marketing digital comprende la utilización de sistemas de información para promover la venta de productos o servicios, en concreto se basa en la utilización de internet, redes de telecomunicación, en este sentido el marketing digital viene a ser como la estrategia que tienen las organizaciones para poder llegar a sus clientes mediante la implementación de tecnologías de información (Luque-Ortiz, 2021).

Para lograr los objetivos del marketing y mejorar los procesos de ventas en las empresas se utilizan los diseños de páginas web, una tienda virtual y tecnologías digitales como sistemas de recomendación mediante la aplicación de inteligencia artificial para vender en línea (Uribe & Sabogal Neira, 2021).

Tecnología para marketing digital

Machine Learning y Big Data analítica, son modelos de aprendizaje profundo expertos en el reconocimiento de patrones y en la toma de decisiones, para las empresas que buscan recopilar una gran cantidad de datos tomados de las redes sociales y entornos IOT (Luque-Ortiz, 2021), estas técnicas ayudan a aumentar las visitas de los clientes categorizando las diferentes respuestas de clics que los usuarios dan a productos y filtrando solo la información de importancia (Ullal et al., 2021).

Los sistemas de recomendación y análisis de sentimientos son herramientas esenciales en el campo del marketing digital aplicando algoritmos inteligentes, lógica difusa y percepción por análisis de valores, aumentan la productividad en las empresas mejorando la eficiencia en las ventas ayudado a los usuarios a encontrar productos de su preferencia y acorde a sus gustos (Marín López & López Trujillo, 2020). Esta lógica de análisis se basa en las preferencias, calificaciones sobre un producto y busca a los usuarios que han tomado decisiones parecidas para poder realizar las recomendaciones (Guevara Albán et al., 2018).

2.5. Algoritmo KNN como método de recomendación

El algoritmo K vecinos más cercanos (KNN) (Sen et al., 2016) está basado en instancias y es de tipo supervisado de machine learning y se puede usar para clasificar nuevas muestras o para predecir, este algoritmo es fácil de elaborar y es utilizado en la resolución de muchos problemas, como sistemas de recomendación, búsquedas semánticas y también detección de algunas anomalías (KNN) (Li et al., 2020). Se utiliza ampliamente entre los algoritmos de clasificación, cuyo concepto se deriva del principio de clasificación del vecino más cercano. Para ello se presenta los pasos que sigue el algoritmo KNN para generar una recomendación de manera general.

Paso 1: Calcular la distancia euclidiana

- Expresión matemática de manera natural

$$\text{dist}(x_{i,x_j}) = \sqrt{(x_{i1}-x_{t1})^2 + (x_{i2}-x_{t2})^2 \dots (x_{ip}-x_{tp})^2}$$

Dónde:

x_i : Es un clúster de entrada con características p ($x_{i1}, x_{i2}, \dots, x_{ip}$)

n : El número total de tuplas de entrada ($i = 1, 2, \dots, n$)

p : El número total de características ($j = 1, 2, \dots, p$)

La distancia euclidiana entre tuplas X_i y X_t ($t = 1, 2, \dots, n$)

Paso 2: Almacenamiento de distancias en un "array"

arrayResult=[""] → Variable donde se almacenará los resultados.

Paso 3: Ordenamiento de distancias

Luego de haber obtenido las distancias se orden los resultados de menor a mayor, para ello utilizaremos la siguiente función.

ArrayResult= {'P1', dist}[1], {'P2', dist}[2], ..., {'Pn', dist}[N]}

Paso 4: Filtrar “n” productos más cercanos

Sea el valor inicial de $k=0$, $\rightarrow k=10$

```

i = 0,
para (i < k) {
    print (arrayResult[i]);
    i++;}

```

El bucle permitirá mostrar los 10 productos con las distancias más cercanas y recomendarlos.

2.6. Propuesta de la solución

La empresa comercial Marcimex presenta problemas en el proceso de marketing, como el bajo crecimiento en su cartera de clientes, lo cual no le permite abarcar nuevos clientes, generando esto un decrecimiento en las ventas, también, la atención al cliente es tardía, por los gustos y preferencias variados de los clientes y que la atención no se puede realizar en paralelo. En este trabajo, se propone desarrollar un sistema de recomendación basado en el algoritmo KNN, como sistema de apoyo al proceso de marketing, este sistema se encargará de recomendar artículos de línea blanca a los clientes de acuerdo a los criterios de los usuarios, los cuales se establecieron de acuerdo al experto del área de ventas, ya que para realizar el proceso de ventas, se utiliza el método de encuesta, el cual consiste en la elaboración de preguntas y respuesta al cliente, estas preguntas contienen las variables de rango de precio, color, marca del producto ideal para el usuario, esto ayudará al experto a tener una idea clara de qué artículo se ajustará a sus preferencias al cliente que está siendo atendido en ese instante, para un mejor entendimiento se elaboró un cuadro pictográfico del funcionamiento del algoritmo en el sistema tal y como se muestra en la Figura 1, en la que se expresa de manera gráfica el funcionamiento general del sistema.

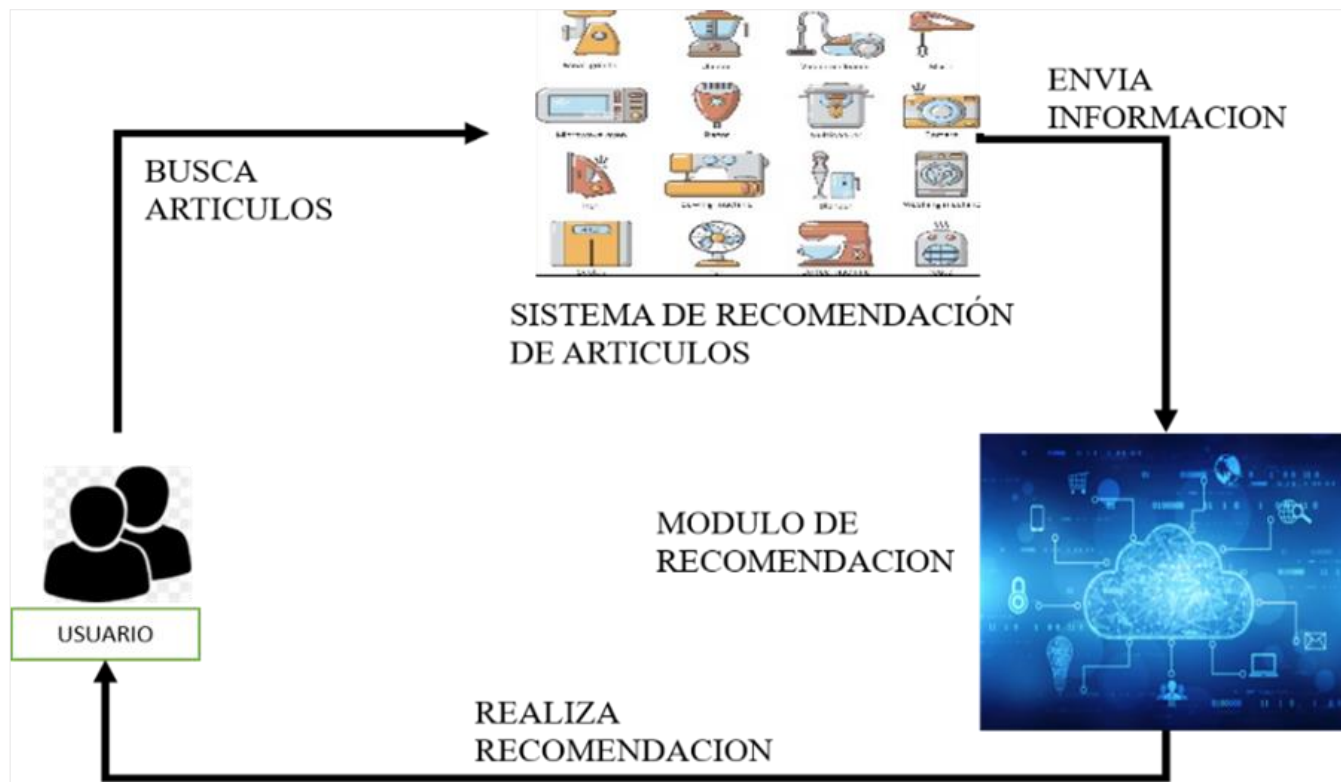


Figura 1. Cuadro pictográfico

En el cuadro pictográfico, podemos visualizar que el proceso inicia cuando el cliente busca un artículo, llevándolo a una interface de usuario donde se le muestran los productos de sus preferencias, este módulo

envía información al módulo de recomendación donde se aplica la lógica del algoritmo KNN y se le recomiendan al usuario productos que estén acorde a sus preferencias.

2.7. Especificación del algoritmo

El algoritmo tuvo 12 variables de entrada, que fueron necesarias para realizar el cálculo de las distancias y así realizar la recomendación a los usuarios, cabe recalcar que dichas variables fueron obtenidas a partir de los criterios de los productos en base a las variables con más peso que vendrían a ser las preferencias del usuario, para un mejor entendimiento se elaboró la Tabla 1 donde se detalla las variables, los nombres y los acrónimos de estas variables a utilizar.

Tabla 1.

Tipos de Variables

Nº variables	Nombre de Variables	Acrónimo
V1	Marca	M
V2	Peso	P
V3	Precio	PR
V4	Color	C
V5	Tipo consumo energético	TC
V6	Función	F
V7	Tipo Producto	TP
V8	Tipo Material	TM
V9	Garantía	G
V10	Cantidad de Likes	CL
V11	Cantidad de vistas	CV
V12	Capacidad	CAP

2.8. Lógica del funcionamiento del algoritmo

Luego de haber realizado la definición de las variables que usarán en el algoritmo, se procede a realizar la elaboración de la lógica de funcionamiento del algoritmo de KNN, adaptándolo a nuestro caso de estudio, el cual inicia con la lectura de la de las 12 variables.

2.9. Corrida del algoritmo

Al momento de generar recomendaciones, se podrá utilizar dos tipos de recomendaciones, se recomendará de manera específica con productos específicos que el usuario prefiera y recomendaciones generales con recomendaciones de productos que podrían interesar al cliente, por consiguiente para la corrida del algoritmo, se utilizará la recomendación específica con datos específicos que el usuario detalle, por lo que se elaboró la Tabla 2 donde se muestran los datos de un cliente con gustos específicos por un producto y se agregó un "0" a los datos que vendrían a ser variables secundarias que son dependientes del algoritmo y datos en el sistema.

Tabla 2.

Especificaciones de cliente

Descripciones del producto solicitado	
Marca	Samsung
Peso	0
Precio	s/.2000
Color	Plomo
Tipo Consumo energético	0
Función	0
Tipo Producto	Refrigeradora
Tipo Material	0

Garantía	0
Cantidad de Likes	0
Cantidad de vistas	0
Capacidad	0

En cuanto a la base de conocimiento, se utilizó un total de 10 productos, estos artículos pasaron por un proceso de conversión de datos cualitativos a datos cuantitativos, para que la ejecución del algoritmo sea fácil, el único que no se convirtió a cuantitativos fue el nombre del producto, ya que este nos permitirá diferenciar o saber qué producto se nos está recomendando, para ello se procedió a igualar cada variable que describe a un producto ya sea marca, color, garantía y más variables, a un número que inicia en uno y asciende en relación a la cantidad de datos cualitativos de los productos, asimismo se tomó los acrónimos de las variables de la Tabla 1, a continuación, se presenta de manera más detallada mediante la Tabla 3 el proceso de conversión de datos que simula el proceso del sistema.

Tabla 3.*Tabla de conversión de datos*

Producto (P)	Marca (M)	Color™	T. Consumo (TC)	Función (F)	Garantía (G)	Tipo material™
Lavadora	1= LG	1= Blanco	1=A	1= Lavado y Secado	1= Tres meses	1= Lamina lisa
Aspiradora	2= Samsung	2= Gris	2=B	2= Refrigeración	2= Seis meses	2= Aluminio
Cocina	3= Electrolux	3= Negro	3=C	3= Cocinas	3= Ocho meses	3= Plástico
Hornos	4= Coldex	4= Amarillo	4=D	4= Climatización	4= Un año	4= Metálico
Microondas	5= Philips	5= verde	5=E	5= Ayudantes de cocina	5= Año y medio	
Licuada	6= Panasonic	6= Blanco/Negro	6=F	6= Ayudantes de hogar		
Cafetera	7= Pioneer	7= Gris/Negro	7=G			
Nevera inteligente	8= Beko	8= Blanco/Gris				
Congelador	9= Mabe	9= Rojo				
Refrigerador	10=Indurama					

Luego de haber igualado los datos de cualitativos a cuantitativos pasamos a elaborar la tabla de productos la cual consta sólo de datos cuantitativos, para poder generar la corrida, expresados en la Tabla 4, para ello se tomó valores cuantitativos de la Tabla 3 en relación a características específicas de diez artículos seleccionados para poder realizar el cálculo de la distancia en base a la preferencia específica del usuario detallado en la Tabla 3, asimismo, se tomaron datos reales fijos extraídos de otras fuentes para generar una mejor precisión del algoritmo como son el precio, peso, marca, tipo de consumo, material, garantía y capacidad. Para finalizar se remplazaron esos datos en la Tabla 3 para obtener un valor cuantitativo así poder obtener un valor numérico en relación con el cálculo de la distancia para generar recomendaciones en base a la preferencia específica de un cliente.

Tabla 4.*Tabla productos*

N°	Artículo	Marca	Peso (kg)	Precio	Color	T. Consumo	Función	T. Producto	T. Material	Garantía	Cant. Likes	Cant. Vistas	Capacidad(lt)
1	Lavadora	1	10	988	5	2	4	5	6	2	0	0	13
2	Aspiradora	1	4	385	2	2	2	1	2	1	0	0	5
3	Cocina	4	13	790	6	2	1	1	2	1	0	0	72.2
4	Hornos	4	8	964.5	1	2	3	2	4	2	0	0	23

5	Microondas	1	4	499	6	2	2	4	1	1	0	0	34
6	Licuada	4	2	290	4	2	3	1	2	1	0	0	2.3
7	Cafetera	1	6	549	3	2	5	2	1	1	0	0	1.5
8	Nevera inteligente	5	1	2050	7	1	3	3	4	1	0	0	368
9	Congelador	2	46	2100	2	2	5	1	6	2	0	0	380
10	Refrigerador	2	53	2000	3	1	2	4	1	2	0	0	370

Para la corrida del algoritmo y el cálculo de distancias se distribuyeron cinco pasos fundamentales que permitan realizar el cálculo de variables y ejecución del algoritmo, como se detalla a continuación.

Paso 1: Lectura de la información en variables

En el primer paso se hace la lectura de la información y se almacena en las variables reemplazando datos de la Tabla 2, en la que ingresan 12 variables V1, V2, V3, V4, V5, V6, V7, V8, V9, V10, V11, V12 donde V1=Samsung, V2=0 V3=2000, V4=Plomo, V5=0, V6=0, V7=Refrigeradora, V8=0, V9=0, V10=0, V11=, V12=0 Litros.

Luego de haber ingresado y almacenado, pasamos al paso dos, donde se cargarán los datos almacenados en la tabla productos tal y como se detalla a continuación.

Paso 2: Cargado de datos

En este paso se realiza el cargado de datos de la tabla productos expresada en la Tabla 4, el cual será almacenado en un "Array" el que nos permitirá convertir a datos cuantitativos toda la información de los productos, se expresa de la siguiente manera.

ArrayBD_LB = [Tabla_Productos]

Luego de haber cargado la información de los productos, se procede a hacer el cálculo de la distancia entre las variables de los gustos del cliente y de los productos.

Paso 3: Cálculo de la distancia Euclidiana

Luego de haber realizado el paso 1 y 2, que son la lectura de variables y el cargado de datos de la tabla productos, se procesa a realizar el cálculo de la distancia, para ello utilizaremos la fórmula expuesta del algoritmo KNN, la cual se detalla a continuación:

Expresión matemática de manera natural

$$\text{dist}(x_{i,x_j}) = \sqrt{(x_{i1}-x_{t1})^2 + (x_{i2}-x_{t2})^2 \dots (x_{ip}-x_{tp})^2}$$

Expresado de manera general

$$\text{dist}=(x_1,x_2)=\sqrt{\sum_{i=1}^n (x_{1i}-x_{2i})^2}$$

Para realizar el cálculo de distancia de los productos y las variables, utilizaremos el bucle "foreach" para calcular las variables uno a uno con los productos de la base de datos.

```
Foreach ($arrProductos as $key => $Crt) {
```

```
    $arrDis[$key] = [pow (($V1 - $Crt['c1']),2) + pow (($V2 - $Crt['c2']),2)]
```

```
+ pow (($V3 - $Crt ['c3']),2) + pow (($V4 - $Crt ['c4']),2)
+ pow (($V5 - $Crt ['c5']),2) + pow (($V6 - $Crt ['c6']),2)
+ pow (($V7 - $Crt ['c7']),2) + pow (($V8 - $Crt ['c8']),2)
+ pow (($V9 - $Crt ['c9']),2) + pow (($V10 - $Crt ['c10']),2)
+ pow (($V11 - $Crt ['c11']),2) + pow (($V12 - $Crt ['c12']),2)];
```

arrayResult = [""] → Variable donde se almacenará los resultados.

Para el cálculo de la distancia entre las variables entrantes y los artículos de la tabla productos, se elaboró una tabla como se muestra a continuación.

Tabla 5.
Cálculo de distancias

N°	Modelo	Cálculo de distancia	Distancia
1	Lavadora	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1073,90
2	Aspiradora	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1656,37
3	Cocina	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1246,67
4	Hornos	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1092,91
5	Microondas	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1538,85
6	Licuadora	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1749,75
7	Cafetera	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	1497,72
8	Nevera inteligente	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	70,35
9	Congelador	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	100,82
10	Refrigerador	$distancia1 = \sqrt{(V[1] - crt[c1])^2 + \dots + (V[12] - crt[c12])^2}$	5,20

Luego de haber realizado el cálculo de las distancias con el bucle for, se procede a mostrar los resultados en la siguiente tabla:

Tabla 6.
Resultados de distancia

N°	Producto	Distancia
1	Lavadora	1073,90
2	Aspiradora	1656,37
3	Cocina	1246,67
4	Hornos	1092,91
5	Microondas	1538,85
6	Licuadora	1749,75
7	Cafetera	1497,72
8	Nevera inteligente	70,35
9	Congelador	100,82
10	Refrigerador	5,20

Luego de realizar el cálculo de distancias de las preferencias del usuario con la tabla productos se procede a ordenar los resultados, para luego encontrar los “N” productos más cercanos a las preferencias.

Paso 4: Ordenar resultados

Para el ordenamiento de los resultados de las distancias, se define un “arrayResult” para el almacenamiento de las distancias ordenadas en forma ascendente.

```
ArrayResult= {['P1', dist][1], ['P2', dist][2], ..., ['Pn', dist][N]}
```

De esta manera se expresan todas las distancias en orden ascendente, desde la menor distancias hasta la distancia del “N” producto más lejano.

Tabla 7.
Ordenamiento de distancias de menor a mayor

N°	Producto	Distancia
10	Refrigerador	5,20
8	Nevera inteligente	70,35
9	Congelador	100,82
1	Lavadora	1073,90
4	Hornos	1092,91
3	Cocina	1246,67
7	Cafetera	1497,72
5	Microondas	1538,85
2	Aspiradora	1656,37
6	Licuada	1749,75

Paso 5: Definimos las “N” distancias más cercanas a mostrar

Para ello ejecutamos el siguiente bucle, que nos permitirá encontrar las distancias más cercanas en relaciona las variables entrantes de las preferencias del cliente.

Sea el valor inicial de $k=0$, $\rightarrow k=3$

```

i=0
para (i<k) {
    print (arrayResult[i]);
    i ++;
}

```

Para el cálculo de bucle expresado se trabajará con las 3 distancias más cercanas a las preferencias del cliente, entonces si $k=3$, nos muestra las 3 distancias más cercanas a las preferencias del usuario. Obteniendo como resultado los productos expresados en la siguiente tabla.

Tabla 8.
K productos más cercanos

N°	Producto	Distancia
10	Refrigerador	5,20
8	Nevera inteligente	70,35
9	Congelador	100,82

De esta manera obtenemos que mediante el uso del algoritmo “KNN” se recomendaron los tres productos más cercanos a las necesidades del usuario, por lo que se recomendaron los productos tales como refrigerador, nevera inteligente y congelador.

3. RESULTADOS

La ejecución del algoritmo produce los tres productos principales que se acercan más a las preferencias del usuario, clasificados de menor a mayor, es decir, la distancia desde el más corto hasta el más alejado de las preferencias del cliente. Como producto más cercano se recomienda el refrigerador, donde las variables con mayor peso son el precio, la marca y el color, ajustados al gusto del cliente. El siguiente producto más cercano es el refrigerador inteligente, que pertenece a la categoría de productos de refrigeración. A la hora de recomendar, se toman como variables más ponderadas las variables precio y color relacionadas con las preferencias del usuario, y finalmente se recomienda un congelador, cuya variable dependiente recomendada es el precio, lo que demuestra que el algoritmo propuesto funciona exactamente de acuerdo con los gustos y preferencias del usuario.

3.1. El software de recomendación de productos

La implementación de este sistema de recomendación se realizó a través del editor de código de programación Visual Studio Code, con el patrón de arquitectura MVC (Modelo, Vista Controlador) aplicando el lenguaje de programación PHP junto al sistema de gestión de base de datos MySQL. También se utilizó el entorno de modelado RSA (Rational Software Architect) para la elaboración de los diagramas de caso de uso, de clases de análisis de secuencia entre otros, para ello se realizó la Figura 2 en la que se expresa de manera más estructurada el funcionamiento del sistema.

En la Figura 2 se muestra los módulos con los que el sistema de recomendación cuenta, entre ellos están el módulo de acceso al sistema, interacción con el usuario, interacción con los productos y el módulo de recomendación.

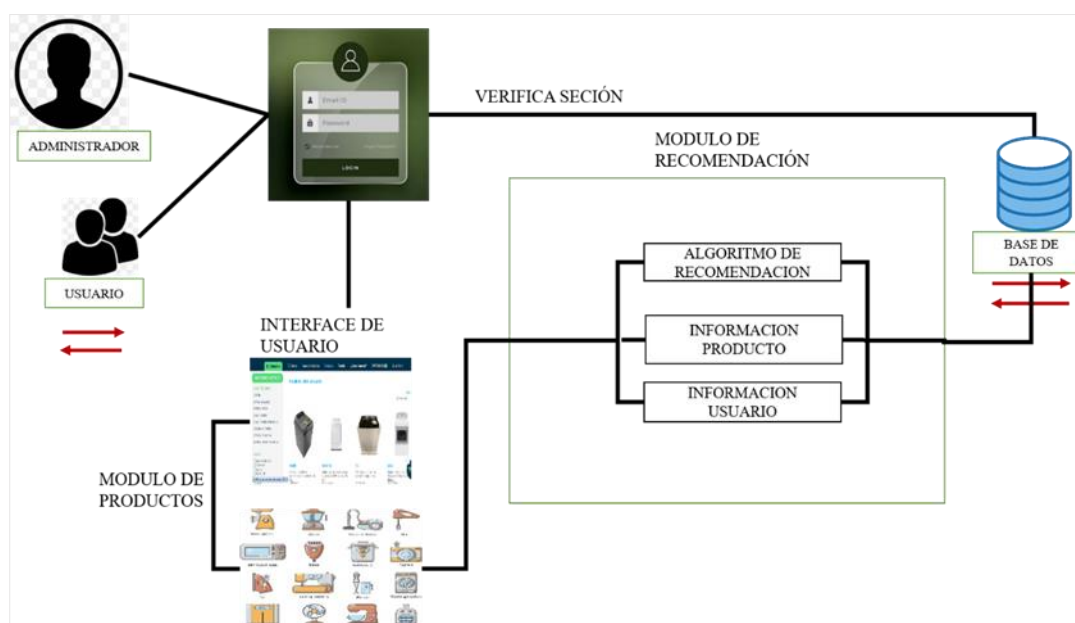


Figura 2. Modelo del sistema de recomendación

Para un mejor entendimiento, en la Figura 3 se presenta el caso de uso del sistema en la que tenemos como actores al administrador del sistema, cliente y trabajador, el cliente podrá registrarse, así mismo hacer registro de sus preferencias al entrar al sistema, podrá visualizar los productos y valorar con likes al producto, por otra parte tenemos al administrador que estará encargado del control total del sistema y podrá agregar trabajadores que a su vez estos mismos puede registrar productos y administrar los procesos de registro, modificación y eliminación de productos.

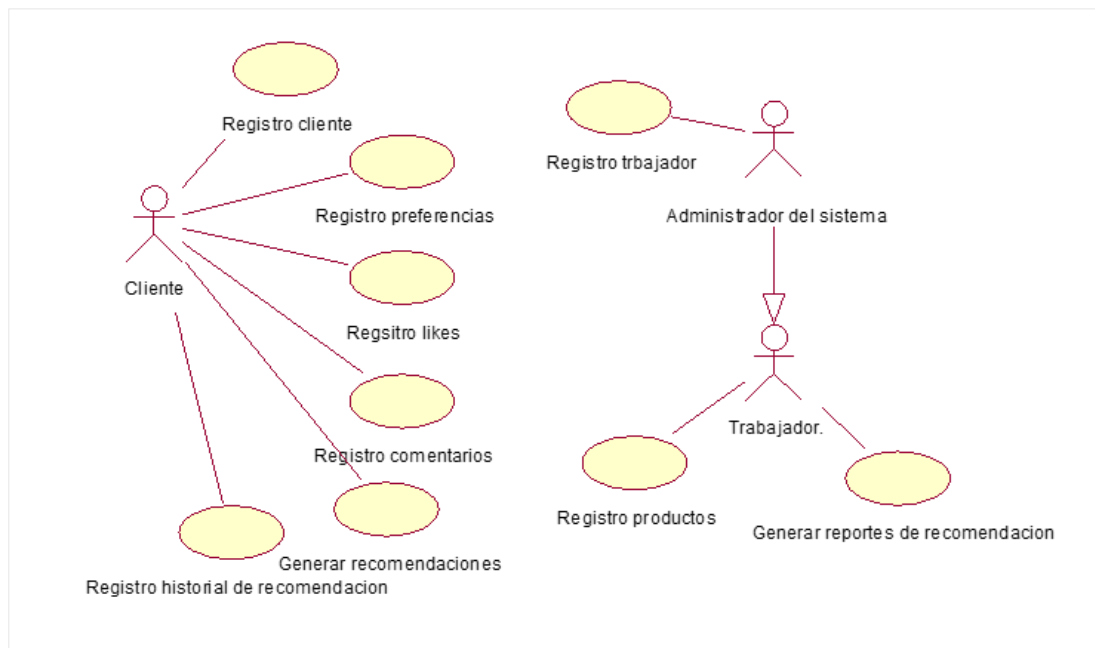


Figura 3. Casos de uso del sistema

3.2. Proceso de recomendación

Para el funcionamiento del sistema propuesto se toma en cuenta el primer inicio de sesión de un cliente mostrando un cuestionario para recolectar las 3 variables que tienen más peso al momento de trabajar con el algoritmo y generar la recomendación tal y como se muestra en la Figura 4.

Para continuar, por favor brindanos tu información

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sapiente dolore dolores a excepturi et! Repellendus quidem facere officis minima est!

¿Cual es su presupuesto?

¿Cual es su color favorito?

Seleccione una opción ▼

¿Que marca de producto buscas?

Seleccione una opción ▼

Continuar →

Figura 4. Cuestionario de recolección de datos

En la Figura 5 se muestra la parte de recomendaciones que el sistema brinda a los clientes al haber calculado y comparado la información entrante mediante el cuestionario de inicio de sesión por primera vez con la base de datos con toda la información de los productos registrados.

Una vez que el cliente realizó su inicio de sesión por primera vez el sistema ha calculado y mostrado recomendaciones basada en las variables que más peso tienen definidas en las preguntas del cuestionario mostrado en la Figura 4, para la recomendación personalizada el cliente pasa a interactuar con el sistema visualizando más productos. El sistema le permitirá visualizar los productos que el cliente elija cargando sus características dentro de las tres variables más sumadas a las del inicio de sesión haciendo un total de doce variables registrándolas como valores cuantitativos para poder ser calculadas mediante el algoritmo K vecinos más cercanos y su fórmula de la distancia euclidiana.

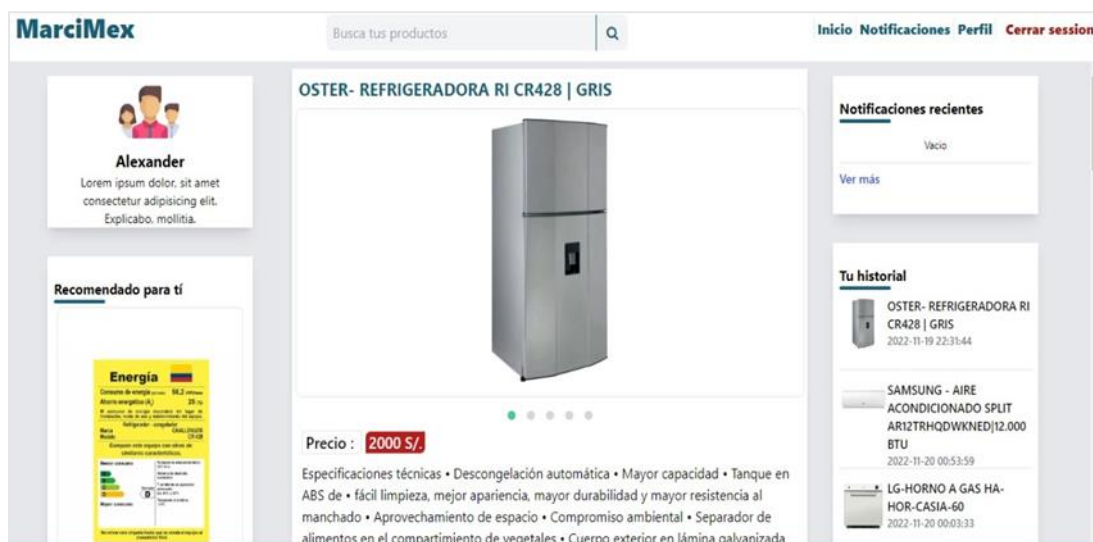


Figura 5. Recomendación de productos

El cliente también podrá comentar y calificar, esta calificación como una variable más registrada dentro de la base de datos del sistema para que se una al cálculo de la distancia a través del algoritmo permitiendo al sistema actualizar su lista de recomendaciones de manera constante tal y como se muestra en la Figura 6 y 7.

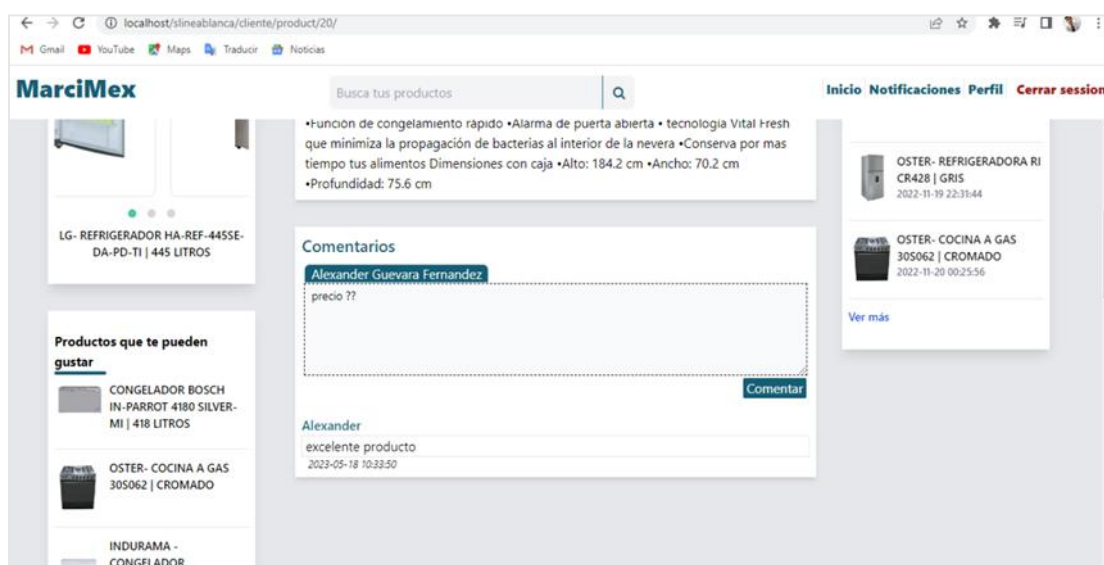


Figura 6. Registro de comentarios

Asimismo, el sistema también permitirá el registro y validación de cada like y comentario que el cliente haga para poder utilizarlo como dato cuantitativo al momento de realizar recomendaciones, así como se muestra en la Figura 7.

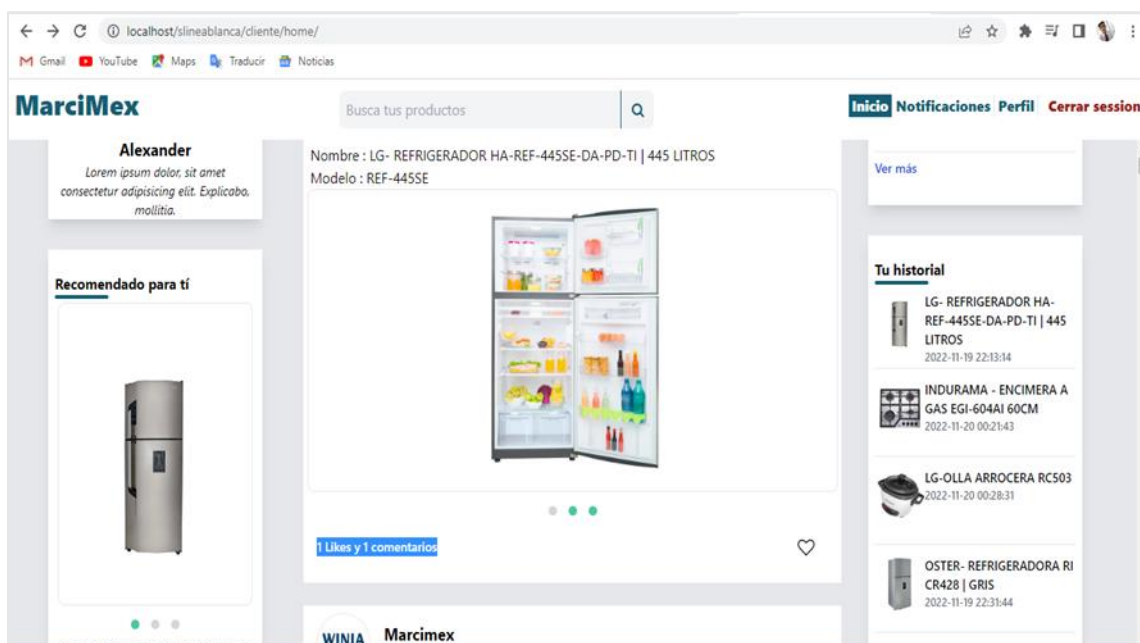


Figura 7. Registro de Likes y comentarios

4. DISCUSIÓN

La distancia euclidiana es eficiente en relación con recomendación de productos donde exista la distancia entre atributos que tengan características comunes a diferencia de otros métodos para el cálculo de medidas de similitud (Bag et al., 2019), el algoritmo KNN es el algoritmo que más se adapta al cálculo de distancia generando mayor eficiencia, precisión y adaptabilidad al momento de generar recomendaciones mediante el cálculo de distancias tal y como se demuestra en la Figura 8.

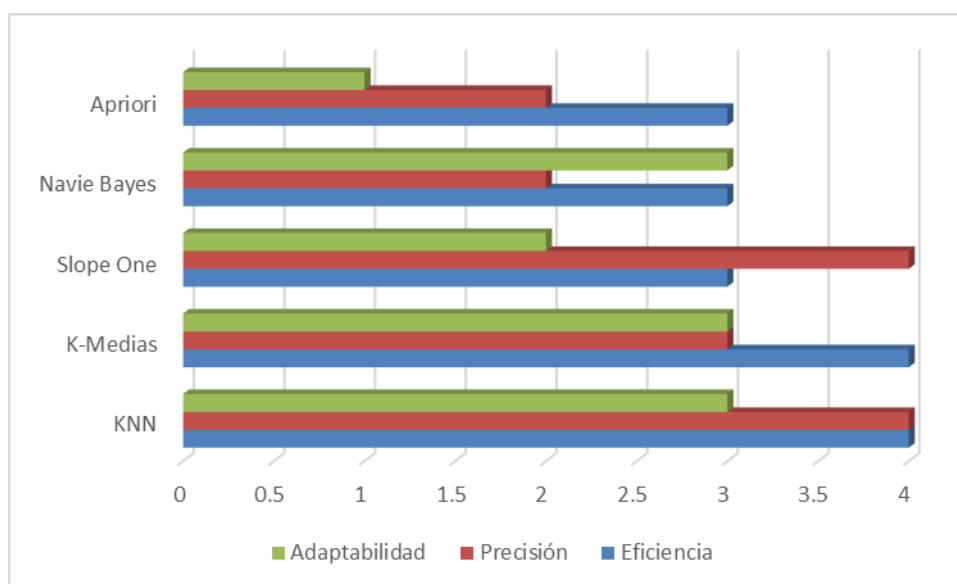


Figura 8. Comparación de algoritmos

A diferencia de otros algoritmos, mediante el cálculo de distancia el algoritmo KNN logra realizar recomendaciones precisas en relación a variables que son calculadas mediante la navegación del usuario por el sistema, a diferencia de otras investigaciones, el algoritmo KNN logra recomendar productos a los clientes de forma eficiente además de ello, se pretende implementar una técnica basada en lógica difusa que incremente el tiempo de recomendación y recomiende sólo productos más relevantes para el usuario,

(Karthik & Ganapathy, 2021), en relación a las preferencias obtenidas mediante la navegación que este realiza en el sistema.

Para la solución de arranque frío del sistema es procedente realizar el pedido de indicios de lo que busca el cliente para ello, se presenta un formulario con variables primarias donde el cliente ingrese datos estáticos y a partir de ello generar recomendaciones en relación a los puntos más cercanos a las preferencias del cliente tal y como se detalla en la Figura 4, también la implementación de la técnica de procesamiento de lenguaje natural que ayuda a predecir productos similares o productos recientes que son sacados a la venta mediante un vector de funciones que nace a raíz de imágenes (Kumar Sharma et al., 2023).

CONCLUSIÓN

En la literatura, los sistemas de recomendación son herramientas tecnológicas que juegan un papel muy importante al momento de realizar recomendaciones, ya que en la actualidad se utilizan para dar soluciones a distintos problemas donde la información que existe es abrumadora, filtrando así toda la información en paquetes más pequeños y de interés para los usuarios.

Se concluye que con la realización de este sistema de recomendación bajo sus interfases amigables se logró implementar una solución automatizada mediante el uso del algoritmo K vecinos más cercanos apoyando a los procesos de marketing digital mediante el comercio electrónico y a ello se suma el aumento significativo de las ventas de productos electrónicos.

Se concluye que el algoritmo K vecinos más cercanos al momento de analizar los valores cuantitativos de cada variable y compararlas con cada producto seleccionado por el cliente es capaz de brindar recomendaciones precisas y muy aceptables.

Se recomienda fortalecer la capacidad de los sistemas de recomendación a través del algoritmo K vecinos más cercanos mejorando los procesos de venta para que el cliente obtenga opciones precisas de electrodomésticos.

Se recomienda promover la implementación de tecnologías como los sistemas de recomendación en las empresas nacionales e internacionales que actualmente hacen uso del comercio electrónico, para mejorar diversos procesos de marketing digital.

FINANCIAMIENTO

Ninguno

CONFLICTO DE INTERESES

No existe ningún tipo de conflicto de interés relacionado con la materia del trabajo.

CONTRIBUCIÓN DE LOS AUTORES

Conceptualización: Guevara-Fernandez, A. & Coral-Ygnacio, M.

Curación de datos: Coral-Ygnacio, M.

Análisis formal: Guevara-Fernandez, A.

Metodología: Coral-Ygnacio, M.

Software: Guevara-Fernandez, A.

Visualización: Guevara-Fernandez, A.

Redacción - borrador original: Coral-Ygnacio, M.

Redacción - revisión y edición: Guevara-Fernandez, A. & Coral-Ygnacio, M.

REFERENCIAS BIBLIOGRÁFICAS

- Adeniyi, D. A., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. *Applied Computing and Informatics*, 12(1), 90–108. <https://doi.org/10.1016/j.aci.2014.10.001>
- Afoudi, Y., Lazaar, M., & Al Achhab, M. (2021). Hybrid recommendation system combined content-based filtering and collaborative prediction using artificial neural network. *Simulation Modelling Practice and Theory*, 113, 102375. <https://doi.org/10.1016/j.simpat.2021.102375>
- Bag, S., Ghadge, A., & Tiwari, M. K. (2019). An integrated recommender system for improved accuracy and aggregate diversity. *Computers & Industrial Engineering*, 130, 187–197. <https://doi.org/10.1016/j.cie.2019.02.028>
- Cai, X., Hu, Z., Zhao, P., Zhang, W. S., & Chen, J. (2020). A hybrid recommendation system with many-objective evolutionary algorithm. *Expert Systems with Applications*, 159, 113648. <https://doi.org/10.1016/j.eswa.2020.113648>
- Cao, B., Zhao, J., Liu, X., Kang, X., Yang, S., Kang, K., & Yu, M. (2018). Multiobjective recommendation optimization via utilizing distributed parallel algorithm. *Future Generation Computer Systems*, 86, 1259–1268. <https://doi.org/10.1016/j.future.2017.09.005>
- Cao, Y., & Li, Y. (2007). An intelligent fuzzy-based recommendation system for consumer electronic products. *Expert Systems with Applications*, 33(1), 230–240. <https://doi.org/10.1016/j.eswa.2006.04.012>
- Castro Gallardo, J. (2012). *Un nuevo modelo ponderado para sistemas de recomendación basados en contenido con medidas de contingencia y entropía* [Universidad de Jaén]. https://sinbad2.ujaen.es/sites/default/files/publications/TTII_JorgeCastro.pdf
- Criado González, M. (2018). Análisis e implementación de un sistema de recomendación para la lista de la compra [Universidad Carlos III de Madrid]. <http://hdl.handle.net/10016/2943>
- Franco Zapata, A. (2021). *Sistemas de recomendación contextual* [Universidad EAFIT]. <http://hdl.handle.net/10784/3141>
- Guevara Albán, G. P., Guevara Albán, C., & Valverde, I. (2018). Sistemas de Recomendaciones: Una herramienta para mejorar la gestión de la información en las PYMES. *Journal of Science and Research: Revista Ciencia e Investigación*, 3(CITT2017), 121–127. <https://doi.org/10.26910/issn.2528-8083vol3isscitt2017.2018pp121-127>
- Hssina, B., & Erritali, M. (2019). A personalized pedagogical objectives based on a genetic algorithm in an adaptive learning system. *Procedia Computer Science*, 151(2018), 1152–1157. <https://doi.org/10.1016/j.procs.2019.04.164>
- Jain, A., Nagar, S., Singh, P. K., & Dhar, J. (2020). EMUCF: Enhanced multistage user-based collaborative filtering through non-linear similarity for recommendation systems. *Expert Systems with Applications*, 161, 113724. <https://doi.org/10.1016/j.eswa.2020.113724>
- Karthik, R. V., & Ganapathy, S. (2021). A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. *Applied Soft Computing*, 108, 107396. <https://doi.org/10.1016/j.asoc.2021.107396>
- Korus, K., Salamak, M., & Jasiński, M. (2021). Optimization of geometric parameters of arch bridges using visual programming FEM components and genetic algorithm. *Engineering Structures*, 241, 112465. <https://doi.org/10.1016/j.engstruct.2021.112465>

- Kumar Sharma, A., Bajpai, B., Adhvaryu, R., Dhruvi Pankajkumar, S., Parthkumar Gordhanbhai, P., & Kumar, A. (2023). An Efficient Approach of Product Recommendation System using NLP Technique. *Materials Today: Proceedings*, 80, 3730–3743. <https://doi.org/10.1016/j.matpr.2021.07.371>
- Li, M., Li, Y., Lou, W., & Chen, L. (2020). A hybrid recommendation system for Q&A documents. *Expert Systems with Applications*, 144, 113088. <https://doi.org/10.1016/j.eswa.2019.113088>
- Liu, H., Zhao, J., Li, P., Zhao, P., & Wu, X. (2021). Shared-view and specific-view information extraction for recommendation. *Expert Systems with Applications*, 186, 115752. <https://doi.org/10.1016/j.eswa.2021.115752>
- Luque-Ortiz, S. (2021). Estrategias de marketing digital utilizadas por empresas del retail deportivo. *Revista CEA*, 7(13), 0–22. <https://doi.org/10.22430/24223182.1650>
- Marín López, J. C., & López Trujillo, M. (2020). Análisis de datos para el marketing digital emprendedor: Caso de estudio Parque de Innovación Empresarial - Universidad Nacional sede Manizales. *Revista Universidad y Empresa*, 22(38), 65. <https://doi.org/10.12804/revistas.urosario.edu.co/empresa/a.7135>
- Martinez Rodriguez, J. R., & Alarcón Martínez, G. J. (2020). Análisis de la flexibilidad del proveedor y la participación de abastecimientos en el desempeño de manufactura para el sector de electrodomésticos (Analysis of supplier flexibility and purchasing participation in the manufacturing performance for applianc. *Revista Innovaciones de Negocios*, 17(33), 98–127. <https://doi.org/10.29105/rinn17.33-6>
- Mlika, F., & Karoui, W. (2020). Proposed Model to Intelligent Recommendation System based on Markov Chains and Grouping of Genres. *Procedia Computer Science*, 176, 868–877. <https://doi.org/10.1016/j.procs.2020.09.082>
- Ni, L., Lin, H., Zhang, M., & Zhang, J. (2018). Hybrid Filtrations Recommendation System based on Privacy Preserving in Edge Computing. *Procedia Computer Science*, 129, 407–409. <https://doi.org/10.1016/j.procs.2018.03.016>
- Ravnik, J., Jovanovac, J., Trupej, A., Vištica, N., & Hriberšek, M. (2021). A sigmoid regression and artificial neural network models for day-ahead natural gas usage forecasting. *Cleaner and Responsible Consumption*, 3, 100040. <https://doi.org/10.1016/j.clrc.2021.100040>
- Sánchez, N. (2019). *Línea Blanca en auge, de la mano de la innovación y la eficiencia*. Electromarket. https://www.electromarket.com/uploads/2019/08/linea_blanca_auge_19618_20190801014454.pdf
- Ullal, M. S., Hawaldar, I. T., Soni, R., & Nadeem, M. (2021). The Role of Machine Learning in Digital Marketing. *SAGE Open*, 11(4), 215824402110503. <https://doi.org/10.1177/21582440211050394>
- Uribe, C. I., & Sabogal Neira, D. F. (2021). Marketing digital en micro y pequeñas empresas de publicidad de Bogotá. *Revista Universidad y Empresa*, 23(40). <https://doi.org/10.12804/revistas.urosario.edu.co/empresa/a.8730>
- Ye, G., & Zhao, X. (2018). Improved SVD algorithm based on Slope One. *Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018*, 1, 1002–1006. <https://doi.org/10.1109/CCDC.2018.8407276>
- Zhang, F., Qi, S., Liu, Q., Mao, M., & Zeng, A. (2020). Alleviating the data sparsity problem of recommender systems by clustering nodes in bipartite networks. *Expert Systems with Applications*, 149, 113346. <https://doi.org/10.1016/j.eswa.2020.113346>