



## Sistema de recomendación basado en las emociones de los usuarios en e-marketing

Recommendation system based on user emotions in e-marketing

Rivero Albarrán, Dulce 

Pontificia Universidad Católica del Ecuador

Guerra-Torrealba, Laura 

Pontificia Universidad Católica del Ecuador

### Resumen

El presente trabajo propone una estrategia para integrar un sistema de identificación de emociones en un sistema de recomendación para e-marketing. El sistema emplea la cámara del dispositivo móvil (con autorización del usuario) para capturar imágenes faciales y, mediante la red neuronal Mini-Xception, identifica las emociones. El sistema de recomendación, de tipo híbrido, combina un enfoque basado en contenido, que utiliza TF-IDF y la similitud del coseno, con otro basado en popularidad. Se aplica un esquema de decaimiento temporal para ponderar las detecciones emocionales más recientes y se normaliza el nivel emocional para facilitar la comparación entre usuarios. Los resultados muestran una precisión del 92% en la identificación de emociones. El

---

Rivero Albarrán, Dulce, <https://orcid.org/0000-0003-2736-5117>, Pontificia Universidad Católica del Ecuador-Ibarra, [dmrivero@pucesi.edu.ec](mailto:dmrivero@pucesi.edu.ec)

Guerra-Torrealba, Laura, <https://orcid.org/0000-0001-6325-943X>, Pontificia Universidad Católica del Ecuador-Ibarra, [lrguerra@pucesi.edu.ec](mailto:lrguerra@pucesi.edu.ec)

**Forma de citar este artículo:** Rivero Albarrán, D y Guerra-Torrealba, L. (2025). Sistema de recomendación basado en las emociones de los usuarios en e-marketing. *Redmarka. Revista de Marketing Aplicado*, vol 29, núm. 2, 78-94. <https://doi.org/10.17979/redma.2025.29.2.12767>

sistema de recomendación arroja una lista de productos ajustados a los niveles de emociones calculados durante la sesión.

**Palabras clave:** sistema de recomendación, redes neuronales profundas, sistema de identificación de emociones, e-marketing, aprendizaje automático

### **Abstract**

This work proposes a strategy for integrating an emotion recognition system into an e-marketing recommendation system. The system uses the mobile device's camera (with the user's authorization) to capture facial images and, via the Mini-Xception neural network, identifies emotions. The hybrid recommendation system combines a content-based approach, which uses TF-IDF and cosine similarity, with one based on popularity. A temporal decay scheme is applied to weight more recent emotional detections, and the emotional level is normalized to facilitate comparison across users. The results show 92% accuracy in emotion recognition. The recommendation system outputs a list of products tailored to the emotion levels computed during the session.

**Keywords:** recommendation systems, deep neural networks, emotion identification systems, e-marketing, machine learning

## **1. INTRODUCCIÓN**

El rápido crecimiento del comercio electrónico en los últimos años evidencia un cambio en los hábitos de compra. Asimismo, el uso masivo de teléfonos inteligentes con aplicaciones que permiten adquirir productos en cualquier horario y sin necesidad de perder tiempo para desplazarse a un centro comercial ha impulsado la expansión de este tipo de comercio (Bandyopadhyay et al., 2021). En este contexto, para ser competitivas en la era digital, aumentar su productividad y mejorar su eficiencia, las empresas necesitan incorporar sistemas de venta en línea que ayuden a los usuarios a encontrar productos acordes con sus gustos y preferencias. Integrar estas soluciones como parte de sus estrategias de mercadeo contribuye a incrementar la productividad y a mejorar la eficiencia organizacional (Fernandez y Ygnacio,, 2023).

Estos sistemas necesitan captar el interés de los usuarios; por ello, deben ofrecer un servicio adaptado a sus necesidades. Dado que las personas pueden manifestar sus emociones a través de expresiones faciales, estas aplicaciones, al igual que un buen vendedor, deberían identificar qué objetos son de preferencia para cada usuario. Como se sabe, las emociones, al ser un estado psicológico complejo, influyen en la forma en que las personas piensan, actúan, interactúan y toman decisiones. En consecuencia, una aplicación de comercio electrónico puede captar la emoción reflejada en el rostro del consumidor para predecir sus preferencias y sugerir los productos que satisfacen sus

expectativas o necesidades. Esta predicción utiliza las emociones de un conjunto de usuarios (clientes), registrados previamente.

Con los recientes avances de los algoritmos de aprendizaje profundo, se han desarrollado diversos métodos para el reconocimiento facial de emociones en tiempo real. Por lo que el reconocimiento automático de las emociones humanas ha cobrado relevancia y se aplica en múltiples ámbitos: desde el desarrollo de interfaces humano-computador hasta la educación personalizada, los sistemas de salud, la monitorización del clima organizacional y los sistemas de recomendación.

La detección de emociones humanas se evalúa cuantitativamente. Hay dos enfoques básicos utilizados para estudiar la naturaleza de las emociones, que son: el enfoque discreto y el enfoque bidimensional de valencia-excitación. El primero identifica seis emociones fundamentales: alegría, tristeza, ira, miedo, disgusto y sorpresa (Ekman, 1992), mientras que el segundo opera en los ejes duales de valencia (la positividad o negatividad de la emoción) y excitación (la intensidad de la experiencia emocional) (Bagherzadeh et al., 2024).

Los sistemas de recomendación han incorporado módulos de reconocimiento de emociones para comprender el estado afectivo del cliente en relación con los productos. La incorporación de las emociones ha allanado el camino para aprovechar estas como descriptores que explican una mayor proporción de la variación en las preferencias del usuario que los descriptores como género y edad, usados tradicionalmente. Por ejemplo, un usuario podría querer consumir un tipo de contenido diferente cuando está feliz que cuando está triste. Por lo que la lista de elementos recomendados se debe adaptar al estado de ánimo del usuario, el sistema debe ser capaz de detectar dicho estado y utilizarlo en el algoritmo como información contextual (Tkalcic et al., 2011).

Este trabajo propone un sistema recomendador para el comercio electrónico a través del reconocimiento facial de emociones.

### 1.1. Trabajos previos

La literatura ofrece varios trabajos donde se ha incorporado la detección de emociones en los sistemas de recomendación de productos. Entre ellos se encuentra el de Won et al. (2023), que presenta un sistema de recomendación de imágenes personalizadas basado en información ambiental y expresiones faciales. El sistema utiliza un algoritmo de aprendizaje profundo para clasificar las expresiones faciales, mientras que los datos ambientales se recopilan mediante sensores de teléfonos inteligentes. Además, para realizar las recomendaciones, emplea un enfoque híbrido que combina el filtrado basado en contenido y el filtrado colaborativo.

Partiendo de la premisa de que los usuarios suelen buscar música acorde con su estado emocional, Singh y Dembla (2023) proponen un sistema de recomendación que integra

el reconocimiento de emociones mediante análisis facial. Para la detección, emplean aprendizaje por transferencia con modelos como CNN, ResNet50V2, VGG16 y EfficientNet-B0, entrenados sobre FER2013 y reforzados con imágenes adicionales obtenidas de Google. La recomendación se apoya en un conjunto de datos de Spotify y utiliza K-means para agrupar las pistas por emoción, personalizando así las sugerencias según el estado emocional detectado.

Por su parte, Manimaran et al. (2025) presentan también un sistema de recomendación musical. En este trabajo, se utiliza una CNN para la detección de emociones y, para la recomendación de canciones coherentes con ese estado, se analiza un conjunto de datos que correlaciona emociones con atributos musicales como género, tempo y contenido lírico. El sistema aprende continuamente de las respuestas del usuario frente a las sugerencias, ajustando sus recomendaciones con el tiempo. Además, incorpora un enfoque de recomendación híbrido para mejorar la precisión y la satisfacción del usuario.

Recomendaciones inteligentes basadas sólo en las emociones detectadas sobre la marcha, sin depender de datos históricos o registros de compra anteriores pueden mejorar los tiempo de respuestas, tal como es el trabajo propuesto por Bandyopadhyay et al. (2021), en cual divide la captura de imágenes del cliente en frames, los convierte a escala de grises y escanea la región de interés y se pasa por el clasificador de emociones para identificar la emoción de cada frame. El sistema calcula la respuestas positivas (suma de emociones de felicidad y sorpresa) y las respuestas negativas (suma de expresiones de tristeza e ira) para finalmente, calcular el nivel de satisfacción (diferencia absoluta entre las respuestas positivas y negativas). El nivel de satisfacción se categoriza en cuatro niveles diferentes para selecciona los objetos asociados al nivel de satisfacción.

También, se ha utilizado Redes Neuronales Recurrentes, en específico redes LSTM para la identificación de emociones, Por ejemplo, el Ruan et al. (2025) se extraen características emocionales de los comentarios de los usuarios (análisis de sentimiento basado en contenido) y las combina con sus datos de comportamiento histórico (filtrado colaborativo) para realizar recomendaciones personalizadas. El sistema utiliza un modelo ATT-LSTM para el análisis de sentimientos, que mejora la precisión en el reconocimiento de emociones en los comentarios de texto. Luego, integra estas características emocionales con el comportamiento histórico del usuario para predecir qué contenido es más probable que le interese al usuario. En este trabajo, para usuarios nuevos (*cold-start users*) que carecen de datos de interacción suficientes, el modelo emplea una estrategia de recomendación basada en contenido, utilizando la similitud semántica entre los artículos y los patrones generales de preferencia del usuario derivados de las tendencias de sentimiento agregadas.

## 2. MARCO TEÓRICO

### 2.1. Sistemas de recomendación

Un sistema de recomendación es un algoritmo que muestra a los usuarios elementos de interés, ya sean películas para ver, libros para leer, productos para comprar o cualquier otro elemento dependiendo de la empresa u organización. Si se tiene una base de datos que contiene la información de estos elementos, el sistema de recomendación mostrará el número máximo de elementos de esta base de datos que se pueden recomendar al usuario (Rivero A, 2023).

Formalmente, un problema de recomendación se puede formular de la siguiente manera: Sea  $C$  el conjunto de todos los usuarios y  $S$  el conjunto de todos los elementos posibles que se pueden recomendar, tales como libros, películas o calzado, etc. Sea  $U$  una función que almacena la utilidad (valoraciones) de los elementos  $S$  para el usuario  $c \in C$ , es decir:

$$U: C \times S \rightarrow R$$

donde,  $R$  es un conjunto totalmente ordenado de valoraciones (enteros no negativos o números reales dentro de un cierto rango). El sistema de recomendación debe seleccionar para el usuario  $c$  el conjunto de elementos  $s \in S$  que maximice la función de utilidad  $U$ .

Basados en el algoritmo utilizado existen diferentes familias de sistemas de recomendación: basados en popularidad, en su contenido o colaborativos.

#### 2.1.1. Sistemas basados en popularidad

Este sistema se adapta a las tendencias actuales. Básicamente, utiliza los elementos (productos) más populares. El elemento más popular entre el público general tiene más probabilidades de ser recomendado a nuevos clientes (Sreekala, 2020). No recomienda según la preferencias de los usuarios, sino que recomienda de manera general al rubro en el que se encuentre el elemento. Este sistema se basa en la cantidad de ventas, en la valoración del producto y las promociones (Fernandez y Ygnacio, 2023)

#### 2.1.2. Sistemas basados en el contenido

Estos sistemas utilizan el historial de las preferencias del usuario (ya sea marca, color, precio y calificaciones) para predecir que producto puede ser interesante, tanto para el usuario como para la empresa, y a raíz de ello mostrar sugerencias similares a los intereses del usuario (Afoudi et al., 2021). El método tiene su origen en los sistemas de recuperación de información y en los sistemas de filtrado de información y requiere información del producto, es decir, se deben identificar los atributos del producto que lo caracterizan y que permiten identificar las preferencias de los usuarios. Estos datos son llamados etiquetas( $t$ ). En este método la función de utilidad  $u(c,s)$  del

elementos para el usuario  $c$  se calcula a partir de las valoraciones que el usuario  $c$  le asigna a los  $s'$  elementos que son “similares” al elemento  $s$  (Rivero A, 2023).

### 2.1.3. Sistemas colaborativos

Es uno de los sistemas más utilizados, su mecanismo de recomendación se basa en filtrar únicamente los gustos y preferencias de los usuarios. También se encarga de evaluar las calificaciones de los productos y predecir cuales podrían gustar al cliente (Tewari, 2020). Este método se basa en el supuesto que los usuarios tienden a seleccionar elementos similares a los que en el pasado usuarios similares a este han seleccionado. Para ello, calcula la utilidad de un elemento  $s$  de un usuario particular  $c$  usando las valoraciones que previamente han realizado otros usuarios. Este tipo de sistemas presentan el problema del arranque en frío.

Estos sistemas (métodos) presentan algunos problemas, como:

- Arranque en frío (cold start), dificultad que enfrentan los sistemas de recomendación, especialmente los colaborativos, para generar buenas recomendaciones cuando falta historial de interacciones., ya sea porque no se tiene historial del usuario, del elemento o es un nuevo sistema.
- Falta de diversidad y novedad, todos los enfoques pueden caer en burbujas de filtro
- Equidad y sesgos: pueden perjudicar a grupos/creadores minoritarios.
- Métricas desalineadas: optimizar clics no siempre maximiza satisfacción a largo plazo.
- Explicabilidad: difícil justificar por qué se recomendó algo.
- Contexto y multi espacio: preferencias cambian por tiempo, lugar, dispositivo; no siempre se modela.

Una de las posibles soluciones a estos problemas es utilizar un método híbrido. Este método propone la combinación de dos o más de los sistemas existentes, ya sea una combinación en paralelo, incorporando características de un sistema en el otro sistema o creando un método unificado (Rivero A, 2023).

## 2.2. E-marketing

El e-marketing, también llamado marketing digital o marketing electrónico, se refiere a la utilización de internet, las redes de telecomunicación y las tecnologías digitales relacionadas para conseguir los objetivos de marketing de la organización (Luque-Ortiz, 2021). Es decir, es el uso del conjunto de herramientas y estrategias digitales, como los sistemas de recomendación, con el fin de mejorar o promover la venta de productos o servicios. Por tanto, el *marketing* digital viene a ser la estrategia que tienen las organizaciones para llegar a sus clientes mediante la implementación de tecnologías de información (Fernandez y Ygnacio, 2023).

### 3. METODOLOGÍA

El trabajo propone una estrategia que integra un sistema inteligente que identifica emociones a partir del rostro de los clientes con un sistema de recomendación. Se llevó a cabo en dos fases. En la primera se desarrolló el modelo de detección de emociones, y en la segunda se construyó el sistema de recomendación integrando el modelo de detección de emociones.

#### 3.1. Sistema de identificación de emociones

El sistema emplea la cámara del dispositivo móvil para la captura de imágenes faciales, proceso que se realiza únicamente previa autorización del usuario. Con el propósito de identificar emociones a partir de dichas imágenes, se entrenó la red neuronal Mini-Xception (Arriaga et al., 2017), una versión simplificada y de baja complejidad computacional de la arquitectura Xception. Esta variante se caracteriza por un número reducido de parámetros y una menor demanda de recursos, sin comprometer el desempeño en la tarea de clasificación. Finalmente, cada emoción es detectada dentro de un intervalo temporal específico y remitida al módulo de recomendación, donde se integra como señal de entrada para la personalización del contenido.

El flujo de trabajo comienza con la activación de la cámara y captura una imagen facial. Esta imagen se procesa posteriormente para extraer los rasgos faciales requeridos por el modelo pre-entrenado, que está cargado y se encarga de detectar las emociones a partir de la imagen capturada. Finalmente, la emoción identificada es enviada a sistema de recomendación. Como se mencionó anteriormente, el modelo de red neuronal profunda empleado se basa en la arquitectura Mini-Xception, entrenada con la base de datos FER2013, que contiene 35,887 imágenes en escala de grises de 48x48 píxeles recopiladas de Internet. Las imágenes muestran rostros humanos etiquetados con siete tipos de emociones: ira, disgusto, miedo, felicidad, tristeza, sorpresa y neutralidad. El desarrollo se llevó a cabo en Python, utilizando varias bibliotecas: OpenCV para el procesamiento de imágenes, TensorFlow y Keras para construir, entrenar y validar el modelo de red neuronal profunda, y Dlib para localizar puntos clave faciales, lo que permite una alineación precisa de los rostros antes del análisis emocional.

Modelo de red neuronal para identificar emociones

##### 3.1.1. Modelo de la red neuronal para identificar emociones

Para desarrollar el modelo de reconocimiento de emociones, se siguieron los pasos clásicos de aprendizaje automático. Se utilizó el enfoque discreto para entrenar el modelo de clasificación, usando imágenes de rostro con etiquetas para las siete categorías definidas en la base de datos FER2013. A continuación, se describen los principales pasos del proceso:

### *Captura y procesamiento de imágenes*

El sistema cuenta con un sensor de presencia, por lo que cuando detecta la proximidad de una persona, activa la cámara USB conectada a la Raspberry Pi 5 y captura las imágenes faciales del profesor en formato JPG con tres canales (RGB).

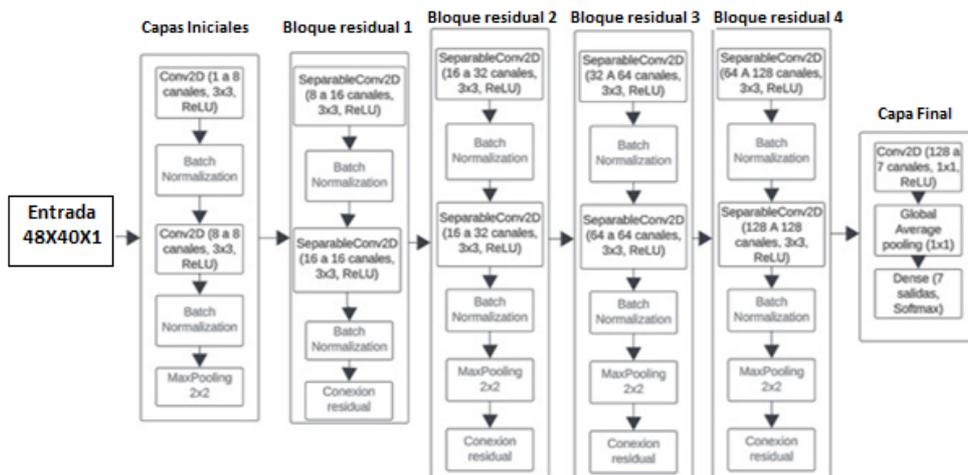
### *Preprocesamiento de imágenes*

Cada imagen capturada se procesa mediante conversión en escala de grises (para reducir la complejidad computacional y centrarse solo en las características relevantes de las expresiones faciales); Ecuilibración de histograma (para mejorar el contraste de la imagen, lo que facilita la detección de rasgos faciales en condiciones de iluminación desigual); Detección facial - Punto de referencia (la biblioteca Dlib se utilizó para identificar 68 puntos clave en la cara, como los ojos, la nariz y la boca); Alineación facial (para que todas las imágenes estuvieran en una posición estándar); Normalización de imágenes (Las imágenes se adaptaron a las dimensiones requeridas por el modelo de 48x48 píxeles); Filtrado de imágenes: Las imágenes borrosas o mal capturadas se descartaron utilizando algoritmos que evaluaron su calidad y nitidez, asegurando que el conjunto de datos procesado fuera consistente.

### *Construcción del modelo*

Se empleó una arquitectura de red neuronal profunda (DNN) basada en el modelo mini-Xception. Esta arquitectura utiliza convoluciones separables en profundidad (SeparableConv2D), junto con mecanismos de normalización por lotes (BatchNormalization) y activaciones ReLU para estabilizar y acelerar el entrenamiento. El bloque final incluye un Global Average Pooling seguido de una capa densa con activación softmax para la clasificación. En conjunto, estas decisiones de diseño favorecen la eficiencia computacional y la capacidad de generalización del modelo. Mini-Xception fue seleccionado debido a su capacidad para equilibrar la precisión y el rendimiento en dispositivos con recursos limitados y su excelente rendimiento en el conjunto de datos FER2013. A continuación, se describe la arquitectura de la red neuronal profunda Mini-Xception (Figura 1).

Figura 1. Detalle de la arquitectura Mini-Xception utilizada (fuente: elaboración propia)



Capas iniciales: Su objetivo es preparar y comprimir la información de la imagen para el resto de los bloques residuales tipo Xception. Esta capa es una combinación de:

- Primera Conv2D (1→8 canales, Relu): Extrae características básicas como bordes y texturas simples. Aumentar a 8 canales permite capturar diferentes aspectos de las imágenes faciales, La función Relu introduce no linealidad controlada.
- Segunda Conv2D (8→8 canales, Relu): profundiza en la detección de características, refinando los patrones detectados en la primera capa.
- Normalización por lotes (*Batch Normalization*): Estabiliza el proceso de aprendizaje mediante la normalización de las activaciones, reduciendo problemas como el cambio de covariables.
- MaxPooling (2x2): Disminuye la resolución (dimensión) a la mitad para que las capas posteriores trabajen con mapas de características más compactos mejorando la eficiencia computacional.

Bloques residuales (bloques Mini-Xception): Su propósito era capturar características más complejas y abstraer patrones emocionales relevantes. Cada bloque incluye convoluciones separables, conexiones residuales y técnicas de agrupación. Se consideraron cuatro bloques residuales.

- Bloque residual 1 (8→16 canales): extrae características más complejas mientras conserva la información de entrada a través de una conexión residual.
- Bloque residual 2 (16→32 canales): Detecta patrones faciales como formas de ojos y boca. El mayor número de canales permite capturar más detalles.
- Bloque residual 3 (32→64 canales): Analiza características de alto nivel, como expresiones completas, utilizando un mayor número de filtros para identificar patrones más sutiles.

- Bloque residual 4 (64→128 canales): Genera representaciones abstractas de expresiones faciales. Este bloque utiliza el máximo número de filtros, condensando la información para las etapas finales.

Capa final – Clasificación: realiza la agregación y decisión de clase, traduciendo los rasgos de alto nivel aprendidos por los bloques residuales en probabilidades de clasificación eficientes y con buena capacidad de generalización.

- Conv2D (128→7 canales): Reduce la dimensionalidad a 7, correspondiente a las emociones estudiadas para la clasificación.
- Global Average Pooling (GAP): Reemplaza las capas densas tradicionales, eliminando parámetros innecesarios y reduciendo el riesgo de sobreajuste. Este paso también mejora la solidez del modelo a las distorsiones espaciales.
- Denso + Softmax (7 salidas): Realiza la clasificación final, asignando probabilidades a cada una de las 7 emociones posibles. La activación de Softmax garantiza que los resultados se puedan interpretar como probabilidades, sumando un total de 1.

Entrenamiento, optimización y ajuste de hiperparámetros del modelo: Durante el entrenamiento, se ajustaron los parámetros del modelo. Se aplicó validación cruzada para garantizar la robustez del modelo y minimizar el riesgo de sobreajuste, asegurando así que el sistema pudiera generalizar bien en condiciones del mundo real. Se utilizó la siguiente división de datos: Conjunto de entrenamiento: ~28.709 imágenes, Conjunto de validación: ~3.589 imágenes y Conjunto de prueba: ~3.589 imágenes. Los parámetros considerados fueron:

- Tamaño del lote: 32 muestras.
- Número de épocas: 100.
- Dimensiones de la imagen: 48x48 píxeles.
- Tasa de aprendizaje inicial: 0,001.
  - Optimizador: Adam.
  - Función de pérdida: Entropía cruzada categórica.

Validación del modelo: Se utilizaron métricas como la Precisión y Sensibilidad (*Sensitivity*, *Recall*) para garantizar un reconocimiento de emociones preciso y confiable. Algunas pruebas se realizaron en entornos reales bajo diferentes condiciones ambientales, incluidos diferentes niveles de iluminación, para determinar la robustez y eficiencia del sistema.

### 3.2. Sistema de recomendación

El sistema de recomendación de productos dispone de datos básicos del usuario como género, edad y estado emocional. Los dos primeros se obtienen durante el registro, mientras que el estado emocional se infiere mediante el sistema de reconocimiento de emociones. El sistema asume que los usuarios, al encontrarse en un estado emocional determinado, tienden a preferir productos asociados con dicho estado.

Con el fin de capturar la dinámica emocional y robustecer la estimación, se aplica un esquema de decaimiento temporal que pondera con mayor peso las detecciones más recientes. Adicionalmente, se normaliza el nivel emocional respecto del total observado para facilitar la comparación entre usuarios con diferentes cantidades de observaciones. En particular, se define un índice emocional normalizado como:

$$\hat{e}(t) = \frac{\#pos(t) - \#neg(t)}{N(t)}$$

Donde  $\#pos(t)$  y  $\#neg(t)$  representan, respectivamente, las detecciones ponderadas de emociones positivas y negativas hasta el instante  $t$ , y  $N(t)$  es la suma ponderada total de detecciones. Bajo un decaimiento exponencial con semivida  $T_{1/2}$ , las ponderaciones se definen como  $w_k = \exp(-\lambda (t - t_k))$ , con  $\lambda = \ln(2)/T_{1/2}$

La interpretación operativa de  $\hat{e}(t)$  se establece mediante umbrales relativos:

- $|\hat{e}(t)| < \tau$ : estado neutral.
- $|\hat{e}(t)| \geq \tau$ : predominio de emociones positivas.
- $|\hat{e}(t)| \leq -\tau$ : predominio de emociones negativas,

donde  $\tau \in (0,1)$  es un parámetro calibrable (p. ej.,  $\tau=0.25$ ).

Para la prueba del sistema de reconocimiento se utilizó una de las tablas de datos (`wdcproducts50cc50rnd000un_valid_large.json`) proporcionados por WDC Products: A Multi-Dimensional Entity Matching Benchmark (Peeters, Der & Bizer, 2023), que contiene 4500 registros. Se leccionaron las cinco primeras columnas, se renombraron las columnas y se anexaron unas columnas adicionales con valores aleatorios, esto son número de visitas, número de likes y nivel de emoción. La Figura 2 presenta las primeras 5 filas de la tabla.

Finalmente, esta señal emocional, junto con los metadatos del usuario y su historial reciente de productos visualizado, se integra en el módulo de recomendación para personalizar el contenido.

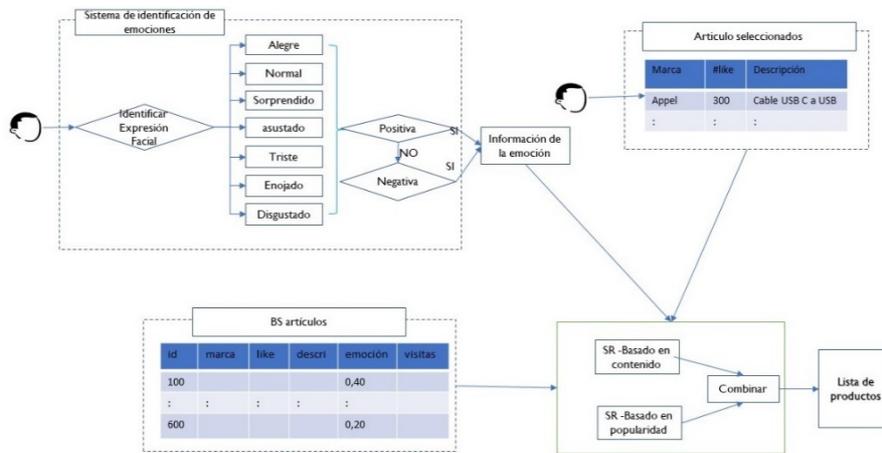
El sistema de reconocimiento es híbrido, combina un sistema de reconocimiento por contenido que utiliza las etiquetas: descripción, precio y nivel\_emoción para encontrar productos similares. Para asignar un valor numérico al atributo descripción se utilizó la frecuencia inversa TD-IDF; para encontrar los productos similares se utilizó la similitud del coseno. También se utilizó un sistema de recomendación basado en popularidad de los productos, para la selección de los artículos a recomendar se utilizaron los atributos likes, nivel\_emocional y número\_visitas seleccionando aquello que estuvieran por encima de un valor  $\mu$  del total de productos de esa marca, el valor  $\mu$  se ajusta de acuerdo al número de productos de la marca (ver Figura 3).

Figura 2. Primeros 5 registro de la BD utilizada para el Sistema de Recomendación (fuente: elaboración propia)

```
print(df_sub.head())
```

	id	marca	titulo	descripcion	precio	likes	nivel_emocion	numero_visitas
0	15228268	Epson	Genuine Epson C1379034010 Magenta High Capaci...	None	34.99	256	0	82
1	62156844	Western Digital	WD Blue 250GB PC SSD - SATA 6 Gb/s 2.5 Inch So...	With superior performance and a leader in reli...	84.99	262	1	59
2	56989328	None	Traveler's Notebook Refill - Grid	Midori grid white paper great for capturing sc...	7.00	263	1	94
3	49805702	Sony	Sony 24mm F1.4 G-Master Full Frame E-Mount Lens	Sony 24mm F1.4 G-Master Full Frame E-Mount Lens	2195	129	1	20
4	66291779	None	Switch 8 puertos 10 100 1000	Switch 8 puertos 10 - 100 - 1000 tp - link	51,73€	283	0	77

Figura 3. Estructura del Sistema de Recomendación (fuente: elaboración propia)



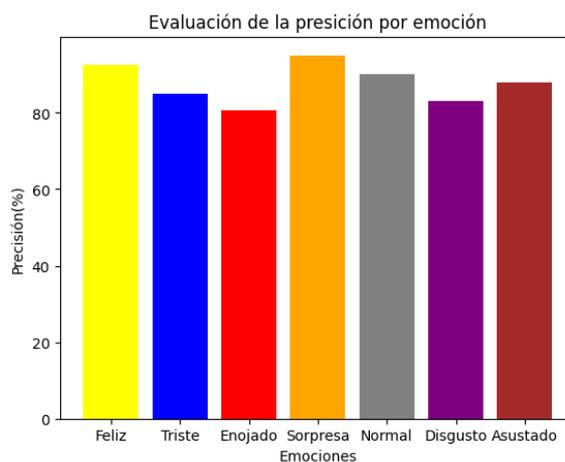
#### 4. RESULTADOS

Al evaluar el modelo de identificación de emociones con los datos de prueba de la base de datos, utilizando los indicadores de precisión, sencibilidad (recall) y F1-Score, se determinó que cumplía con los requisitos requeridos (92% de precisión). La Tabla 1 y la Figura 4 muestran el desempeño del modelo Mini-Xception en la clasificación de cada una de las siete emociones en el conjunto de datos de FER2013, observando que el indicador de precisión supera el 80% para cualquier tipo de emoción estudiada.

**Tabla 1.** Evaluación del modelo de identificación de emociones (fuente: elaboración propia)

Emotion	Accuracy (%)	Recall (%)	F1-Score (%)
Feliz	97	94	95
Neutral	93	90	91
Sorprendido	96	93	95.5
Angry	94	90	92
Sad	88	85	86
Disgusted	89	87	88
Scared	90	88	89

**Figura 4.** Diagrama de barras de la evaluación del sistema de identificación de emociones (fuente: elaboración propia)



Adicionalmente, para validar el rendimiento del modelo, se realizaron 24 pruebas secuenciales de detección de emociones, obteniendo un tiempo de procesamiento promedio de 27,92 ms ± 3,94 ms. Estos valores garantizan que el modelo pueda mantener un rendimiento constante y confiable durante la operación continua, cumpliendo y superando los requisitos de respuesta en tiempo real establecidos.

Los tiempos de latencia estuvieron dentro de los parámetros esperados, con un promedio de 150 milisegundos desde la captura de la imagen hasta la visualización de la emoción detectada. La interfaz respondió correctamente en el 98% de las interacciones evaluadas. La respuesta del sistema tuvo un tiempo promedio de 300 ms. Las pruebas de usuario se realizaron con la participación de algunos usuarios, que interactuaron con el sistema en varios escenarios de uso, lo que permitió la recopilación de datos significativos sobre el rendimiento del sistema en condiciones del mundo real. Además, se realizaron pruebas complementarias utilizando un script especializado, acumulando 178 registros de 10 participantes diferentes en condiciones óptimas de iluminación, lo que arrojó una precisión general del modelo del 93,57%. Este resultado superó significativamente las expectativas iniciales y confirmó la solidez del sistema para

su aplicación en entornos educativos. La matriz de confusión generada durante las pruebas de usuario reflejó la ira y la sorpresa como las emociones más difíciles de clasificar.

El Sistema de recomendación arrojó una lista de productos que se ajustaban a los niveles de emociones que se fue calculando durante la sesión de consulta.

La Figura 5 muestra la salida del sistema de recomendación para un usuario que ha realizado la consulta: `recs_q = recomendar_por_consulta(texto_descripcion="FeaturesPage yield25000 pagesBrand compatibilityXeroxCompatible productsPhaser 7760Colour of productBlackQuantity1Printing coloursYellowTypeTonerPackaging dataPackaging width95 mmPackaging height89 mmPackage weight680 gPackaging length368 mmPackaging typeBox", precio=216.44, tao=0.35, top_k=8)`

Donde *tao* es el nivel de emoción calculada en el marco de tiempo, *top\_K* es el número de productos a recomendar

**Figura 5.** Diagrama de barras de la evaluación del sistema de identificación de emociones.

	id	titulo	precio
3382	45493711	Switch TP-LINK TL-SF1024 19'' Rackmount 24x10/...	217.30
1932	45493711	Switch TP-LINK TL-SF1024 19'' Rackmount 24x10/...	217.30
2734	66749842	Details aboutNEW 2020 Shimano Dura Ace 11 Spee...	218.88
1174	66749842	Details aboutNEW 2020 Shimano Dura Ace 11 Spee...	218.88
2087	66749842	Details aboutNEW 2020 Shimano Dura Ace 11 Spee...	218.88
1697	66749842	Details aboutNEW 2020 Shimano Dura Ace 11 Spee...	218.88
1380	56902092	Jabra BIZ 2400 II DQ Duo NC WB Corded Headset	209.00
661	47121877	Sluŝalke Audio-Technica ATH-PRO7x	208.95

	nivel_emocion	score_similitud
3382	1	0.992278
1932	1	0.992278
2734	1	0.992278
1174	1	0.992278
2087	1	0.992278
1697	1	0.992278
1380	1	0.992278
661	1	0.992278

## 5. DISCUSIÓN Y CONCLUSIONES

La evaluación del sistema se llevó a cabo mediante pruebas controladas, utilizando tanto datos de la base FER2013 como interacciones en tiempo real. A continuación se presentan los aspectos más relevantes.

En relación a la precisión del modelo se observa que emociones de felicidad, neutralidad y sorpresa se detectan con altos niveles de exactitud, mientras que las emociones de disgusto y miedo presentan desafíos adicionales en términos de reconocimiento. En cuanto a los tiempos de respuesta del sistema de reconocimiento de emociones se alcanzó una capacidad para operar en entornos reales con tiempos de respuesta mínimos, lo que es esencial para el correcto funcionamiento de aplicaciones de comercio electrónico y sistemas interactivos de recomendación.

La incorporación del módulo de reconocimiento emocional en el sistema de recomendación permitió personalizar la oferta de productos en función del estado

afectivo del usuario. Para ello, se aplicó un algoritmo híbrido que combina técnicas basadas en contenido y popularidad. Este método permitió:

- Ajustar las recomendaciones en función del índice emocional: Los productos sugeridos se correspondieron con los niveles de emoción detectados durante la sesión, ofreciendo una experiencia dinámica y contextualizada<sup>1</sup>.
- Mejorar la relevancia de los productos recomendados: Al utilizar información en tiempo real, el sistema pudo adaptar rápidamente sus sugerencias, incrementando la satisfacción y la interacción del usuario.

En resumen, los datos experimentales respaldan la viabilidad técnica del sistema en cuanto a la detección emocional y su integración en algoritmos de recomendación. Los resultados obtenidos demuestran tanto la alta precisión como la capacidad operativa en tiempo real, lo que confirma su aplicabilidad para entornos de e-marketing. Además, este trabajo demuestra que la aplicación de técnicas de aprendizaje profundo para el reconocimiento de emociones, combinadas con métodos híbridos de recomendación, es una estrategia prometedora para transformar el e-marketing. La capacidad de adaptar dinámicamente la oferta de productos en función del estado afectivo del usuario puede traducirse en una mayor relevancia de las recomendaciones y, por ende, en una mejora de la satisfacción del cliente.

## 6. DECLARACIÓN DE AUTORÍA SEGÚN TAXONOMÍA CRediT

**Dulce Rivero Albarrán:** conceptualización; metodología; software; validación; análisis formal; investigación; recursos; curación de datos; redacción-borrador original; redacción-revisión y edición; visualización; supervisión; administración de proyectos.

**Laura Guerra-Torrealba:** conceptualización; metodología; análisis formal; investigación; recursos; redacción-borrador original; redacción-revisión y edición.

## 7. REFERENCIAS BIBLIOGRÁFICAS

Afoudi, Y., Lazaar, M., and Al Achhab, M. (2021). Hybrid recommendation system combined content-based filtering and collaborative prediction using artificial neural network. *Simulation Modelling Practice and Theory*, 113, Article 102375.

<https://doi.org/10.1016/j.simpat.2021.102375>

Arriaga, O., Valdenegro-Toro, M., and Plöger, P. (2017). Real-time convolutional neural networks for emotion and gender classification. *arXiv*. <https://doi.org/10.48550/arXiv.1710.07557>

Bagherzadeh, S., Norouzi, M. R., Hampa, S. B., Ghasri, A., Kouroshi, P. T., Hosseininasab, S., ... and Nasrabadi, A. M. (2024). A subject-independent portable emotion recognition

system using synchrosqueezing wavelet transform maps of EEG signals and ResNet-18. *Biomedical Signal Processing and Control*, 90, Article: 105875. <https://doi.org/10.1016/j.bspc.2023.105875>

Bandyopadhyay, S., Thakur, S. S., and Mandal, J. K. (2021, November). Online recommendation system using human facial expression based emotion detection: a proposed method. In En Mandal, J.K., Buyya, R. y De, D. (Eds) *Proceedings of International Conference on Advanced Computing Applications: ICACA 2021* (pp. 459-468). Springer Singapore. [https://doi.org/10.1007/978-981-16-5207-3\\_38](https://doi.org/10.1007/978-981-16-5207-3_38)

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200. <https://doi.org/10.1080/02699939208411068>

Fernandez, A. G., y Ygnacio, M. A. C. (2023). Una revisión de métodos, técnicas y algoritmos para sistemas de recomendación de productos tecnológicos. *Interfases*, (18), 255-280. <https://doi.org/10.26439/interfases2023.n018.6357>

Luque-Ortiz, S. (2021). Estrategias de marketing digital utilizadas por empresas del retail deportivo. *Revista CEA*, 7(13). <https://doi.org/10.22430/24223182.1650>

Manimaran, S., Saranya, S. R., Selvam, N., and Sherine, S. B. (2025). Emotion-Based Music Recommendation System Using Facial Expression Recognition. *American Journal of Engineering, Mechanics and Architecture (2993-2637)*, 3(3), 174-190. <https://grnjournal.us/index.php/AJEMA/article/view/7186>

Peeters, R., Der, R. C., and Bizer, C. (2023). *WDC Products: a multi-dimensional entity matching benchmark*. <https://webdatacommons.org/largescaleproductcorpus/wdc-products/>.

Rivero A., D. (2023). Sistemas de recomendación: enfoque y técnicas. En D. Rivero A., F. Rivas E., y S. Arciniegas A. (Eds). *Gestión, infraestructura y aplicaciones en sistemas de tecnología de información* (pp. 53-77). Centro de Publicaciones PUCE.

Ruan, T., Liu, Q., and Chang, Y. (2025). Digital media recommendation system design based on user behavior analysis and emotional feature extraction. *PLoS One* 20 (5), Article e0322768. <https://doi.org/10.1371/journal.pone.0322768>

Singh, K. K., and Dembla, P. (2023). A study on emotion analysis and music recommendation using transfer learning. *Journal of Computer Science*, 19 (06), 707-726. <https://doi.org/10.3844/jcssp.2023.707.726>

Sreekala, K. (2020). Popularity based recommendation system. *International Journal of Engineering and Advanced Technology (IJEAT)*, 9 (3), 1561-1566. <https://doi.org/10.35940/ijeat.B4660.029320>

Tewari, A. S. (2020). Generating items recommendations by fusing content and user-item based collaborative filtering. *Procedia Computer Science*, 167, 934-1940. <https://doi.org/10.1016/j.procs.2020.03.215>

Tkalcic, M., Kosir, A., and Tasic, J. (2011). Affective recommender systems: the role of emotions in recommender systems. *The RecSys 2011 Workshops-Decisions@ RecSys 2011 and UCERSTI-2: human decision making in recommender systems; user-centric evaluation of recommender systems and their interfaces-2*, 811, 9-13.

Won, H. M., Heo, Y. S., and Kwak, N. (2023). Image recommendation system based on environmental and human face information. *Sensors*, 23 (11), Article 5304. <https://doi.org/10.3390/s23115304>