

# Sistema de recomendación de películas basado en análisis emocional con chatbot, AngellIA.

Rubén Aponte Núñez, Edison Giraldo Pérez, Joel Orellana Martínez

Universidad Internacional de la Rioja, Logroño (España)

20 de septiembre de 2023

## RESUMEN

La presente investigación se enfoca en el diseño de un sistema de recomendación de películas basado en el análisis de las emociones, implementando un sistema de conversación interactivo llamado AngellIA. Este tiene como principal objetivo optimizar la experiencia del usuario, proponiendo películas que sean acordes a sus estados emocionales predominantes. La metodología adoptada para este proyecto se estructura en distintas etapas: la recopilación de datos cinematográficos, el desarrollo y entrenamiento de un modelo especializado para el análisis de las emociones y la configuración y despliegue del sistema de recomendaciones. Los resultados muestran una capacidad notable de clasificar con precisión las emociones, lo que resulta en recomendaciones cinematográficas pertinentes. Por su parte, las evaluaciones ponen de manifiesto la habilidad del sistema para entregar sugerencias que estén acordes con el perfil emocional del usuario. Es evidente que la aplicación del análisis de las emociones en este contexto potencia de manera significativa la personalización de las recomendaciones, lo que a su vez conduce a un aumento en la satisfacción del usuario.

## PALABRAS CLAVE

Análisis emocional, Chatbot, Sistemas de recomendación, Películas, Procesamiento del lenguaje natural

## I. INTRODUCCIÓN

En este artículo, se propone un sistema de recomendación de películas basado en análisis emocional integrado en un chatbot [1], llamado "AngellIA", la cual simula mantener una conversación con una persona al proveer recomendaciones automáticas, utilizando técnicas avanzadas de Procesamiento del Lenguaje Natural (NLP, por sus siglas en inglés), como lo son los actualmente muy utilizados transformers [2]. Además, se considera el análisis emocional para ofrecer recomendaciones de películas más precisas y personalizadas.

La introducción de la arquitectura de transformers en el sistema de recomendación permite una mayor precisión y adaptabilidad en la captura de las preferencias y emociones de los usuarios en tiempo real, dado que estos utilizan mecanismos de autoatención para capturar dependencias a largo plazo en secuencias de texto y han demostrado ser eficientes en una amplia gama de tareas de NLP, incluida la comprensión y generación de texto [2]. Los transformers han logrado el entrenamiento masivo de corpus de texto, los Grandes Modelos de Lenguaje (Large Language Models o LLM) capaces de poder realizar una amplia gama de tareas de NLP como la generación de texto, por ejemplo, entrenar un chatbot a través de una instrucción (Prompt en inglés) en lenguaje natural [3].

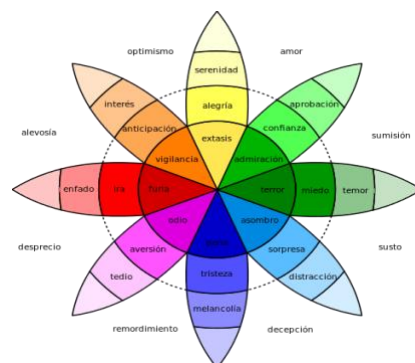
## II. ESTADO DEL ARTE

### 2.1. Avances en el estudio de emociones

En la década de 1980, la investigación de las emociones tomó un giro significativo con la formulación de la Teoría Psicoevolucionista de las Emociones [4]. Esta teoría,

basada en la idea que las emociones desempeñan un papel adaptativo en la evolución humana, postula que sirven como mecanismos adaptativos en respuesta a los desafíos del entorno. Uno de sus logros más notables fue la identificación de ocho emociones primarias organizadas en pares opuestos, como alegría versus tristeza, confianza versus aversión, sorpresa versus anticipación, y enojo versus temor.

Otra contribución importante fue la creación de la "rueda de las emociones", una representación gráfica que ilustra la estructura y jerarquía de las emociones [4]. Esta rueda no solo destaca las emociones primarias y sus opuestos, sino que también permite visualizar las combinaciones y grados de intensidad de cada emoción, como se muestra en la **Figura 1**.



**Figura 1.** Rueda de las emociones.

### 2.2. Avances en la detección de sentimientos y emociones en texto mediante IA

Los primeros esfuerzos en la detección de emociones en texto se centraron en el análisis de sentimiento, buscando la capacidad de identificar si un fragmento de texto tiene una carga emocional positiva, negativa o neutra. Uno de los hitos tempranos fue la creación del Diccionario de Sentimiento AFINN [5], que asigna puntuaciones de sentimiento a palabras en inglés.

Otro hito significativo en esta área tuvo lugar en 2013 con la presentación del enfoque basado en Redes Neuronales Recurrentes (RNN, por sus siglas en inglés). Este enfoque contribuyó de manera fundamental a una comprensión más profunda y contextual de las emociones presentes en el texto [6].

Posteriormente la incorporación de modelos de atención y contexto en la detección de emociones marcó el estándar de la industria. Estos modelos consideran no solo las palabras individuales en el texto, sino también las relaciones y conexiones entre ellas. Un ejemplo es el *Transformer*, un modelo de atención múltiple, que ha demostrado ser efectivo en el análisis de texto emocional [2], la evolución de este modelo a otros como BERT [7] y GPT [8] hicieron una revolución en la detección de sentimientos y emociones. Estos modelos, entrenados en vastas cantidades de texto, pueden transferir su conocimiento previo a tareas específicas de análisis de emociones, mejorando la precisión y la capacidad de adaptación a diferentes dominios.

### 2.3. Enfoques en la extracción de emociones

**Léxico:** utilizan diccionarios y listas de palabras previamente construidas y asociadas con sentimientos o emociones específicas [9], [10]. Estos enfoques suelen ser simples y rápidos, pero pueden ser limitados en términos de cobertura y flexibilidad.

**Machine learning:** emplean técnicas de clasificación supervisada, como máquinas de vectores de soporte (SVM, por sus siglas en inglés), árboles de decisión y clasificadores Bayesianos, para predecir sentimientos y emociones a partir de características extraídas del texto [11][12]. Aunque estos métodos pueden ofrecer un rendimiento mejor que los enfoques basados en léxicos, requieren datos etiquetados y pueden ser sensibles a la selección de características y la elección del algoritmo.

**Deep learning:** las redes neuronales convolucionales [13] (CNN, por sus siglas en inglés) y las RNN [14], han demostrado ser efectivas en el análisis de sentimientos y la detección de emociones debido a su capacidad para modelar representaciones de texto de alto nivel y capturar información contextual [6], [15]. Sin embargo, estos enfoques pueden requerir grandes cantidades de datos y recursos computacionales para entrenamiento y ajuste de modelos. [16].

### 2.4. Sistemas de recomendación

Los sistemas de recomendación son herramientas esenciales en diversas aplicaciones, como comercio electrónico, entretenimiento y redes sociales, para ayudar a los usuarios a descubrir contenido relevante y personalizado [17]. En su esencia, es un sistema que emplea algoritmos para sugerir elementos o productos a usuarios cuando visitan un sitio web o exploran un producto en línea. Estas sugerencias pueden abarcar diversas categorías, desde productos que visualizar y adquirir hasta películas que ver, anuncios para visualizar e

información para leer.

La taxonomía propuesta en *Practical Recommender Systems* [18], describe un sistema en función de varias dimensiones como dominio, propósito, contexto, nivel de personalización, opiniones, privacidad y confiabilidad, interfaces y algoritmos mediante los cuales puede describirse por completo un sistema de recomendación.

### 2.5. Chatbots en sistemas de recomendación

La incorporación de *chatbots* en sistemas de recomendación ha ganado popularidad en los últimos años debido a su capacidad para interactuar con los usuarios de manera natural y en tiempo real [19]. Los *chatbots* pueden mejorar la experiencia de usuario y facilitar la exploración y el descubrimiento de contenido, al tiempo que proporcionan información adicional y responden a preguntas relacionadas con el contenido recomendado [20].

**Chatbots basados en reglas:** los *chatbots*, desde su concepción inicial, han evolucionado notablemente en su complejidad y aplicaciones. Uno de los primeros enfoques en la construcción de *chatbots* fue el modelo basado en reglas. Estos *chatbots* funcionan mediante un conjunto predefinido de reglas que dictan cómo el *chatbot* debe responder ante entradas específicas de los usuarios [21]. Dicho de otro modo, para cada entrada específica del usuario, existe una respuesta predefinida por parte del *chatbot*. El *chatbot* ELIZA, creado por Joseph Weizenbaum en el *Massachusetts Institute of Technology* (MIT), es uno de los ejemplos más emblemáticos de este tipo de sistema, simulando una terapeuta empática y proporcionando respuestas basadas en scripts preprogramados.

**Chatbots basados en técnicas de DL:** el uso de técnicas avanzadas de NLP, como los *transformers* en *chatbots* ha sido un área de investigación activa [2]. Los *transformers* han demostrado ser efectivos en una amplia gama de tareas de NLP, incluido el análisis de sentimientos, la respuesta a preguntas y la generación de texto, gracias a su capacidad para capturar información contextual a largo plazo y modelar representaciones semánticas de alto nivel [7] [22].

Recientemente, los LLM con billones de parámetros pre entrenados en *corpus* gigantescos de texto (GPT, Llama, PALM, etc.) [3] han demostrado habilidades para generar respuestas coherentes y contextuales, mejorando significativamente la calidad y la naturalidad de las interacciones con los usuarios.

## III. OBJETIVOS Y METODOLOGÍA

### A. Objetivo general

Desarrollar un *chatbot* para recomendación de películas personalizadas que utilice enfoques avanzados de NLP, que logre una comprensión y análisis de las emociones y preferencias del usuario, proporcionando recomendaciones altamente relevantes y precisas.

### B. Objetivos específicos

1. Investigar y analizar las técnicas y algoritmos de NLP más adecuados para el análisis de emociones.
2. Revisar y comparar diferentes enfoques y arquitecturas de sistemas de recomendación de películas, incluidos los métodos basados en contenido, filtrado colaborativo y enfoques híbridos.

3. Desarrollar e implementar un modelo de clasificación de emociones para películas, basado en la información textual y las opiniones de los usuarios.
4. Diseñar e implementar un *chatbot* que utilice técnicas de NLP y el modelo de clasificación de emociones para generar recomendaciones de películas basadas en el estado de ánimo y los gustos del usuario.
5. Evaluar la efectividad y precisión del *chatbot* para recomendar películas mediante pruebas con usuarios y análisis de sus respuestas y satisfacción en función de las recomendaciones recibidas.

### C. Metodología del trabajo

A continuación, se describe la metodología del trabajo:

**Revisión bibliográfica:** en esta etapa se llevó a cabo una revisión exhaustiva de la literatura científica y académica relacionada con NLP, el análisis de sentimientos, los sistemas de recomendación de películas y los *chatbots*.

**Creación del modelo de clasificación de emociones:** en esta etapa, se desarrolló e implementó un modelo de clasificación de emociones para películas, utilizando la información textual y las opiniones de los usuarios.

**Diseño e implementación del *chatbot*:** con el modelo de clasificación de emociones desarrollado en la etapa anterior, se diseñó e implementó un *chatbot*.

**Evaluación del *chatbot* para recomendar películas:** en esta etapa, se llevará a cabo la evaluación de la efectividad y precisión del *chatbot* para recomendar películas.

**Análisis de resultados y conclusiones:** por último, se analizarán los resultados obtenidos.

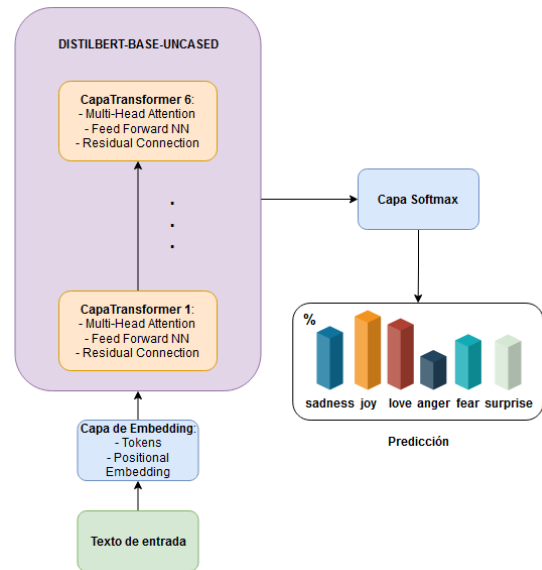
## IV. CONTRIBUCIÓN

La adecuada preparación de un conjunto de datos constituyó un paso esencial en el proceso de construcción de un modelo clasificatorio destinado a identificar las emociones percibidas en películas por consumidores previos. Para el desarrollo del presente estudio se seleccionó el *conjunto de datos IMDB – IEEE Movie reviews* [23]: que incluye casi 1 millón de reseñas provenientes de 1.150 películas distintas registradas en IMDb y 17 géneros distintos.

### 4.1. Creación del modelo para el análisis de emociones

Para abordar la tarea de clasificación emocional en las respuestas de los usuarios y las reseñas de películas, se optó por aprovechar la potencia de los modelos pre entrenados en NLP. Uno de estos modelos, en particular, el "*distilbert-base-uncased*", mostró ser una elección adecuada debido a su eficiencia y rendimiento en tareas de clasificación textual [24].

Al emplear el modelo "*distilbert-base-uncased*", previamente entrenado en una amplia gama de textos, se aprovechó su capacidad para comprender el contexto y la semántica del lenguaje. Este modelo tiene la capacidad de captar matices emocionales sutiles presentes en el lenguaje humano, lo que fue crucial para el éxito de la tarea de clasificación emocional en este proyecto (ver *Figura 2*).



**Figura 2.** Modelo "*distilbert-base-uncased*" con *fine tuning* para clasificación de emociones.

### 4.1.1. Entrenamiento del modelo

Para el entrenamiento del modelo se utilizó el *pipeline* de *Hugging Face* y *PyTorch* y se configuraron los parámetros de entrenamiento, como el tamaño del lote (*batch\_size*), la frecuencia de registro de métricas (*logging\_steps*), el número de epochs de entrenamiento (*num\_train\_epochs*), la tasa de aprendizaje (*learning\_rate*), entre otros. Los parámetros más relevantes de configuración del modelo se presentan en la *Tabla 1*.

**Tabla 1.** Parámetros de entrenamiento del modelo de clasificación de emociones.

| Parámetro               | Valor   |
|-------------------------|---|
| <i>model_name</i>       | <i>distilbert-base-uncased-finetuned-417k</i> |
| <i>batch_size</i>       | 64  |
| <i>num_train_epochs</i> | 2   |
| <i>learning_rate</i>    | $2.5 \times 10^{-5}$                          |
| <i>weight_decay</i>     | 0.01  |

El proceso de entrenamiento tuvo una duración de 44 minutos en 2 *epochs* y se obtuvieron las métricas presentes en la *Tabla 2*.

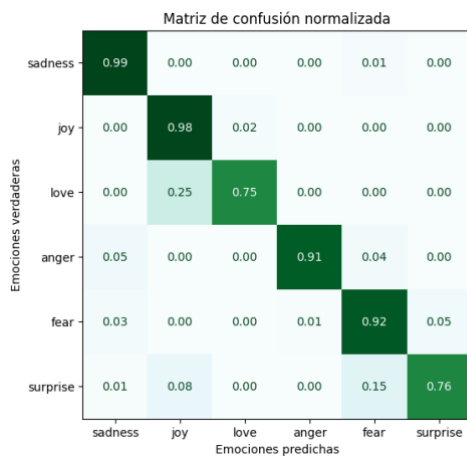
**Tabla 2.** Resultado de Métricas del modelo clasificador de emociones.

| Dataset | Training Loss | Test Loss | Accuracy | F1 Score |
|---------|---------------|-----------|----------|----------|
| Train   | 0.0802        | -         | 0.9448   | 0.9426   |
| Test    | -             | 0.8843    | 0.9392   | 0.9385   |

Para la evaluación de modelo de clasificación, se utilizó la matriz de confusión que permite visualizar el desempeño del modelo para cada clase específica.

En la **Figura 3**, se muestra la matriz de confusión del modelo. Las filas de la matriz representan las emociones reales, mientras que las columnas reflejan las predicciones del modelo. Se puede observar que la matriz de confusión muestra que el modelo es más preciso a la hora de detectar tristeza (99%), seguida de alegría (98%). El modelo es menos preciso a la hora de detectar la ira, el miedo, la sorpresa y el amor. También hay algunos falsos positivos y falsos negativos en la matriz de confusión.

Estos resultados sugieren que el modelo está funcionando adecuadamente en la clasificación de emociones, pero todavía hay margen de mejora, especialmente en la detección de emociones menos comunes como la sorpresa y el amor. Trabajos futuros podrían centrarse en mejorar la precisión del modelo a la hora de detectar estas emociones.



**Figura 3.** Matriz de confusión para evaluar la precisión del modelo de clasificación de emociones.

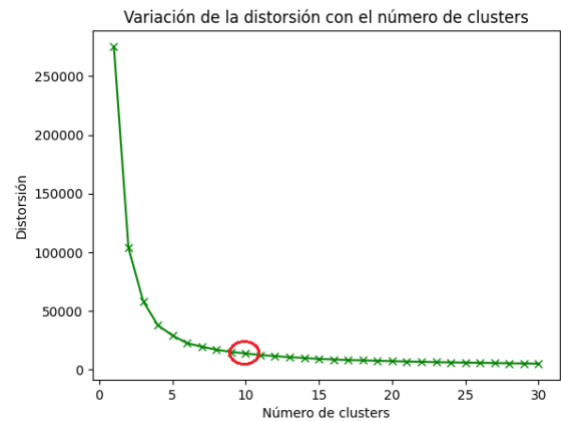
El modelo fue subido a un repositorio de Hugging Face creado para este trabajo llamado: *TFMUNIR/distilbert-base-uncased-finetuned-417k*.

#### 4.1.2. Segmentación del conjunto de datos de películas

Se realizó una segmentación del conjunto de datos debido a la necesidad de incrementar la variabilidad y distribución en las recomendaciones, garantizando un enfoque más dirigido y personalizado. Al dividir el conjunto de datos en grupos homogéneos basados en criterios como género, año o emociones, se posibilitó una personalización más afinada de las recomendaciones, alineándolas con las emociones del usuario en un momento específico.

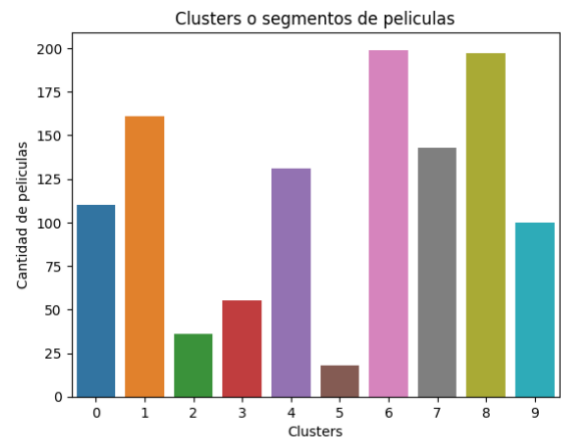
Se estableció la necesidad de analizar un espectro de segmentos que abarca desde 1 hasta 30 segmentos tal como se observa en la **Figura 4**. A partir de esta exploración, se observó que un punto óptimo que proporciona una segmentación diversificada y representativa de las películas se sitúa alrededor de los 10 segmentos. Esta elección se alinea con un equilibrio entre la especificidad y la generalidad: se distancia del valor identificado por el método del codo, que sugiere  $k = 5$ , evitando una posible subsegmentación, pero a su vez, no se inclina hacia valores excesivamente elevados, superiores a  $k = 10$ , que podrían conllevar una sobresegmentación y, por ende, reducir la robustez y

utilidad práctica de los segmentos definidos.



**Figura 4.** Variación de la distorsión con el número de clusters.

Para  $k = 10$ , la segmentación de las películas se puede observar en la **Figura 5**.



**Figura 5.** Variación de la distorsión con el número de clusters.

#### 4.1.3. Reajuste emocional para recomendación de películas

Con base en la rueda de las emociones presentada en la Sección 1, se desarrolló un modelo que reajusta emociones negativas detectadas en las interacciones del usuario con el *chatbot*, orientándolas hacia sus contrapartes positivas. El objetivo de este ajuste es que, al identificar emociones tales como el miedo, la tristeza o el enojo, estas sean procesadas por un algoritmo que, sin eliminar la emoción negativa original, pondere hacia su opuesto positivo.

Esta metodología tiene como finalidad ofrecer recomendaciones de películas que sean más amenas y placenteras, con el propósito de contribuir a la mejora del estado anímico del usuario. Esta estrategia se alinea con la idea de proporcionar una experiencia de usuario optimizada y empática, reconociendo y respondiendo adecuadamente a sus estados emocionales. La propuesta de balance de emociones se muestra en la **Tabla 3**.

**Tabla 3.** Contrapartes de emociones implementadas en el *chatbot*.

| Emoción del usuario | Contraparte emocional del modelo |
|---------------------|----------------------------------|
| Tristeza (sadness)  | Alegría (joy)                    |
| Alegría (joy)       | Alegría (joy)                    |
| Amor (love)         | Amor (love)                      |
| Rabia (anger)       | Amor (love)                      |
| Temor (fear)        | Alegría (joy)                    |
| Sorpres (surprise)  | Alegría (joy)                    |

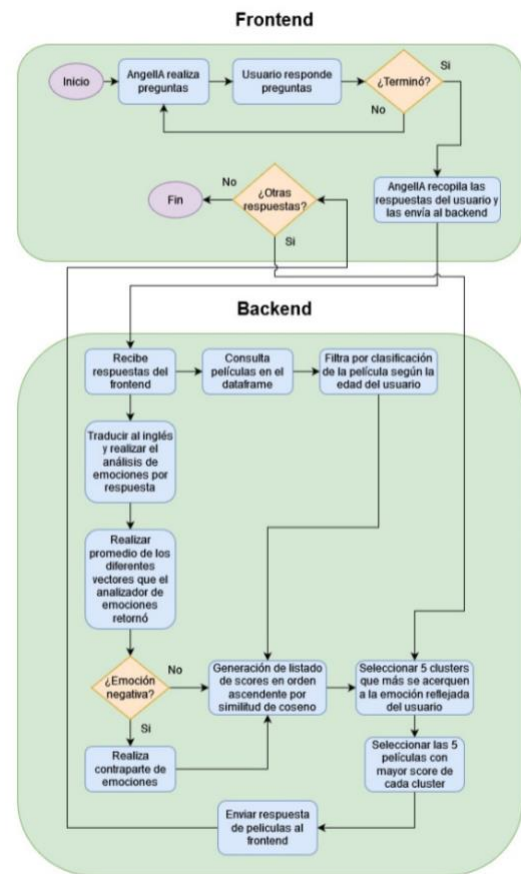
En la implementación del modelo, específicamente si una emoción como la tristeza, excede un umbral definido (en este caso 0.2), el modelo incrementa su contraparte positiva (alegría) mientras reduce proporcionalmente la emoción original. Esta operación se realiza mediante un factor de ajuste, que en este contexto se ha establecido en 0.5. Posteriormente, para garantizar una respuesta coherente, el vector resultante es normalizado, de modo que la suma total de sus componentes emocionales o probabilidades sea 1. De esta manera, el sistema proporciona una respuesta emocionalmente equilibrada y adaptada al contexto del usuario.

#### 4.2. Desarrollo del sistema de recomendación integrado en el *chatbot*

##### 4.2.1. *Chatbot*: AngellIA

"AngellIA" es un sistema de IA diseñado para brindar a los usuarios recomendaciones de películas personalizadas a través de interacciones en tiempo real. Para lograr esta funcionalidad, el sistema se apoya en un *backend* desarrollado en el *framework Django*, que comprende una variedad de funciones y vistas para llevar a cabo tareas esenciales como procesamiento de texto, análisis de emociones, recomendación de películas e interacción con el usuario. Además, el sistema se completa con un *frontend* desarrollado en HTML, JavaScript y CSS, que brinda la interfaz con la que los usuarios interactúan.

El diagrama de la **Figura 6** presenta el funcionamiento simplificado de AngellIA.



**Figura 6.** Funcionamiento simplificado de AngellIA.

##### 4.2.2. *Backend*

El backend coordina las operaciones para ofrecer recomendaciones cinematográficas personalizadas. Su estructura incluye un análisis de emociones, donde un modelo pre entrenado identifica y clasifica las emociones en los mensajes de los usuarios. En el procesamiento textual, también se emplean técnicas de NLP para analizar el contenido lingüístico del mensaje.

##### 4.2.3. *Frontend*

La página web presenta un diseño sencillo, incluyendo un encabezado y un área donde se desarrolla la interacción entre el usuario y el chatbot, AngellIA. Inicialmente, el chatbot pregunta por la edad del usuario, dato esencial para filtrar recomendaciones. Posteriormente, AngellIA plantea interrogantes sobre el día del usuario, su estado anímico y eventos recientes, con el propósito de extraer emociones y sentimientos. Las preguntas buscan capturar una visión general del estado emocional, solicitar una representación narrativa de los sentimientos y obtener relatos de experiencias recientes. Una vez recabadas las respuestas, estas se almacenan en el backend para su posterior análisis y generación de recomendaciones cinematográficas. Al final, el usuario tiene la opción de solicitar más recomendaciones o finalizar la conversación. La interfaz de usuario puede verse en la **Figura 7**.





**Figura 7.** Frontend del chatbot AngellIA.

Cuando AngellIA recopila la información de los usuarios, genera recomendaciones de películas basadas en esas respuestas.

El usuario puede calificar las recomendaciones dando "Me gusta" o "No me gusta". Estas calificaciones se guardan en una base de datos *Postgres* para posterior análisis.

## V. EVALUACIÓN Y RESULTADOS

### 5.1. Pruebas del sistema

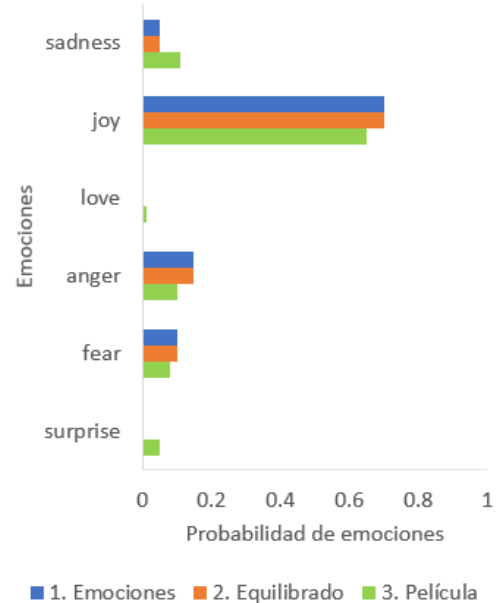
Las pruebas realizadas buscaron evaluar la pertinencia de las recomendaciones del sistema en distintos contextos emocionales. Las recomendaciones dirigidas a emociones negativas se inclinaban hacia géneros como drama, romance, aventura, suspenso y ciencia ficción, que tratan temas de superación y reafirmación de vida. Las películas sugeridas presentan tramas que permiten al usuario confrontar y procesar sus emociones. Por otro lado, en respuesta a emociones positivas, el sistema prioriza géneros como comedia, aventura, animación y acción, aunque mantiene una diversidad en sus sugerencias, permitiendo la exploración de nuevas temáticas por parte del usuario.

### 5.2. Pruebas realizadas por terceros

En el marco del presente estudio, se llevaron a cabo pruebas promocionando el sitio web AngellIA a través de diversas plataformas de redes sociales. Durante un período comprendido entre el 20 y el 28 de agosto de 2023, se extendió una invitación a potenciales usuarios para que procedieran a registrarse y explorar el funcionamiento del sistema. Es importante mencionar que, dentro de la base de datos, únicamente quedaron registrados aquellos usuarios que efectuaron evaluaciones sobre las recomendaciones presentadas. Como resultado de esta fase, se contabilizaron 183 calificaciones aportadas por un

conjunto de 68 usuarios diferentes.

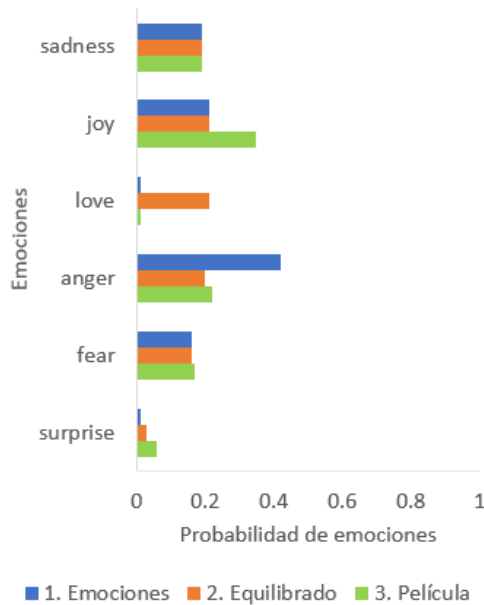
La **Figura 8** muestra que ante una emoción predominante positiva no realiza un reajuste del vector de emociones del usuario y pasa a la etapa de identificación de la película por medio de la distancia de coseno más cercana, que en este caso arroja un valor de similitud de 0.99. El usuario manifestó que la recomendación fue de su agrado.



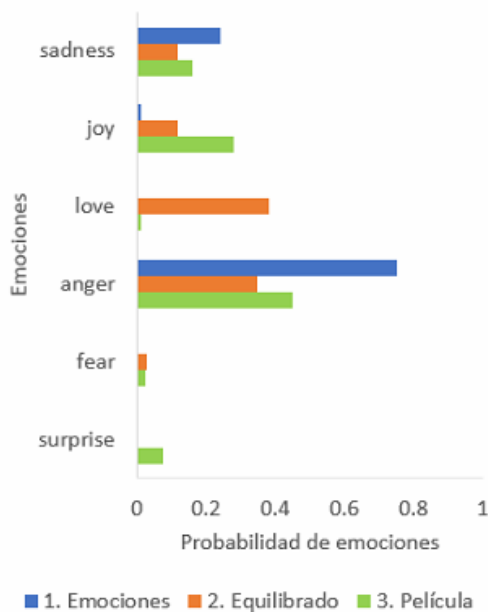
**Figura 8.** Recomendación de película con emoción del usuario altamente positiva.

La **Figura 9** ilustra una situación en la que el usuario experimenta una emoción combinada. El sistema modifica el vector de emociones del usuario, reduciendo la intensidad de las emociones negativas como tristeza, rabia y temor, y potenciando las positivas, específicamente alegría y amor. Luego, el sistema avanza a la fase de identificación de la película utilizando la distancia de coseno, donde se obtuvo una similitud de 0.88. De acuerdo con la retroalimentación del usuario, la sugerencia fue satisfactoria.

Según se puede observar en la **Figura 10**, ante la presencia de una emoción predominantemente negativa, el sistema recalibra el vector emocional del usuario, minimizando las emociones negativas (rabia y tristeza) y amplificando las positivas (alegría y amor). Una vez realizado este ajuste, la elección de la película se realiza a través del método de distancia de coseno, arrojando una similitud de 0.74 en este caso particular. La validación por parte del usuario indica que la elección propuesta fue coherente con sus expectativas.



**Figura 9.** Recomendación de película con emociones mixtas del usuario.



**Figura 10.** Recomendación de película con emoción del usuario altamente negativa.

## VI. DISCUSIÓN

En base a los resultados obtenidos puede evidenciarse algunos aspectos relevantes:

1. El modelo entrenado para la detección de emociones en texto ha proporcionado resultados satisfactorios en la matriz de confusión que han permitido clasificar estados de ánimo de las personas de forma muy eficaz.
2. La detección de las emociones inherentes a las reseñas de consumidores previos de películas junto a la clusterización de ellas ha permitido presentar recomendaciones novedosas que dependen exclusivamente de las opiniones de los usuarios.
3. Ante la presencia de emociones particularmente

negativas, el sistema recomienda películas orientadas a mejorar el estado de ánimo.

## VII. CONCLUSIONES

Las contribuciones de este trabajo se reflejan en la creación de un conjunto de datos de películas específicamente orientado a las emociones detectadas por espectadores previos y a las características de cada una de ellas (año, género, clasificación y sinopsis), así como la implementación de un modelo de análisis emocional y la construcción del sistema de recomendación integrado en el chatbot. Estos componentes se han diseñado para ofrecer recomendaciones de películas que se adapten mejor a las emociones de los usuarios, mejorando así la personalización y la satisfacción de estos.

La técnica de clusterización permitió agrupar películas con distintas variables sujetas a consideración, una de las cuales era el vector promedio de emociones de las reseñas realizadas por espectadores previos de las mismas, este enfoque innovador logró determinar un perfil emocional para cada película, lo que permitió al chatbot ofrecer recomendaciones a los usuarios basadas en la concordancia entre sus emociones y las evocadas por las películas.

En términos de evaluación del modelo de detección de emociones, se entrenó exitosamente un modelo de Transformers con alta precisión capaz de clasificar adecuadamente la mayoría de los estados de ánimo que fueron sujetos a estudio, como lo demostraron las métricas del modelo presentadas en la sección correspondiente.

La evaluación realizada al sistema de recomendación ha demostrado pertinencia en la adaptación de las recomendaciones de películas a las emociones de los usuarios. Los resultados obtenidos en las pruebas funcionales internas y de usuario (terceros) respaldan la eficacia de la solución propuesta.

## REFERENCIAS

- [1] M. Dahiya, "A Tool of Conversation: Chatbot," *International Journal of Computer Sciences and Engineering*, vol. 5, no. 5, pp. 158–161, 2017.
- [2] A. Vaswani *et al.*, "Attention Is All You Need," *Thirty-First International Conference on Neural Information Processing Systems*, pp. 6000–6010, Jun. 2017.
- [3] J. Wei, S. Kim, H. Jung, and Y.-H. Kim, "Leveraging Large Language Models to Power Chatbots for Collecting User Self-Reported Data," *ArXiv*, p. 22, Jan. 2023.
- [4] R. Plutchik, *Emotion: A Psychoevolutionary Synthesis*. Harper & Row, 1980.
- [5] F. Å. Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," *ESWC2011 Workshop on "Making Sense of Microposts"*, vol. 718, pp. 93–98, Mar. 2011.

- [6] R. Socher *et al.*, “Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank,” *2013 Conference on Empirical Methods in Natural Language Processing*, pp. 1631–1642, 2013.
- [7] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *ArXiv*, p. 16, May 2019.
- [8] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving language understanding by generative pre-training,” *OpenAI*, p. 12, 2018.
- [9] M. Hu and B. Liu, “Mining and summarizing customer reviews,” *Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177, 2004.
- [10] C. Strapparava and R. Mihalcea, “SemEval-2007 Task 14: Affective Text,” *Fourth International Workshop on Semantic Evaluations*, pp. 70–74, 2007.
- [11] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? Sentiment Classification using Machine Learning Techniques,” *Conference on Empirical Methods in Natural Language Processing*, pp. 79–86, 2002.
- [12] E. Kouloumpis, T. Wilson, and J. Moore, “Twitter Sentiment Analysis: The Good the Bad and the OMG!,” *Fifth International AAAI Conference on Web and Social Media*, vol. 5, no. 1, pp. 538–541, Aug. 2011.
- [13] S. K. Bharti *et al.*, “Text-Based Emotion Recognition Using Deep Learning Approach,” *Comput Intell Neurosci*, vol. 2022, p. 8, 2022.
- [14] N. Majumder, S. Poria, D. Hazarika, R. Mihalcea, A. Gelbukh, and E. Cambria, “DialogueRNN: An Attentive RNN for Emotion Detection in Conversations,” *Thirty-Third AAAI Conference on Artificial Intelligence*, pp. 6818–6825, Nov. 2018.
- [15] Y. Y. Tang, B. K. Hölzel, and M. I. Posner, “The neuroscience of mindfulness meditation,” *Nat Rev Neurosci*, vol. 16, no. 4, pp. 213–225, Mar. 2015.
- [16] L. Zhang, S. Wang, and B. Liu, “Deep Learning for Sentiment Analysis : A Survey,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 8, no. 4, p. e1253, Jan. 2018.
- [17] F. Ricci, L. Rokach, and B. Shapira, “Introduction to Recommender Systems Handbook,” *Recommender Systems Handbook*, pp. 1–35, 2011.
- [18] K. Falk, *Practical Recommender Systems*. 2019.
- [19] D. Jannach, P. Pu, F. Ricci, and M. Zanker, “Recommender systems: Trends and frontiers,” *AI Magazine*, vol. 43, no. 2, John Wiley & Sons, Ltd, pp. 145–150, 2022.
- [20] M. McTear, Z. Callejas, and D. Griol, *The conversational interface: Talking to smart devices*. Springer Cham, 2016.
- [21] J. Weizenbaum, “ELIZA - a computer program for the study of natural language communication between man and machine,” *Commun ACM*, vol. 9, no. 1, pp. 36–45, Jan. 1966.
- [22] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language Models are Unsupervised Multitask Learners,” *OpenAI blog*, vol. 1, p. 9, 2019.
- [23] A. Pal, A. Barigidad, and A. Mustafi, “IMDb Movie Reviews Dataset,” *IEEE Dataport*, 2020.
- [24] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter,” *ArXiv*, p. 5, Oct. 2019.