REVISTA PERSPECTIVAS

VOL. 8, N°1 / ENERO - JUNIO 2026 / e-ISSN: 266-6688

DOI: https://doi.org/10.47187/perspectivas.8.1.252

DE LA DETECCIÓN A LA CONVERSACIÓN: EVOLUCIÓN DE UN CHATBOT DE APOYO EMOCIONAL CON COMPETENCIA CULTURAL

From Detection to Conversation: Evolution of a Culturally Competent Emotional Support Chatbot

Yan Josue Tepancal Martinez	yan.tepancal@gmail.com
María del Carmen Santiago Díaz	marycarmen.santiago@correo.buap.mx
Ana Claudia Zenteno Vázquez	ana.zenteno@correo.buap.mx
Gustavo Trinidad Rubín Linares	gustavo.rubin@correo.buap.mx

Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla, Puebla, México.

RESUMEN

Este artículo presenta la evolución de PsicoloBot, un chatbot de soporte emocional desarrollado con competencia cultural para el entorno hispanohablante de México. El trabajo se realizó en dos fases. En la primera, se implementó un algoritmo de detección léxica basado en reglas, que alcanzó una precisión del 82% en la clasificación de emociones, aunque manifestó limitaciones importantes en la comprensión contextual y la calidad del diálogo. Para asegurar la adaptación cultural, esta fase requirió un exhaustivo proceso de minería de modismos en redes sociales y transformación de datos (ETL). Posteriormente, se propone una segunda fase que integra un modelo híbrido, incorporando un Modelo de Lenguaje a Gran Escala (LLM), como Gemini de Google, a fin de potenciar el manejo del contexto. Esta estructura fusiona la rapidez y precisión del motor léxico con las capacidades generativas de la IA, permitiendo una interacción que simula un diálogo empático y culturalmente relevante.

Palabras Clave: Inteligencia Artificial, Salud Emocional, Chatbot, LLM, Adaptación Cultural, Modelo Híbrido.

ABSTRACT

This paper presents the evolution of PsicoloBot, an emotional support chatbot developed with cultural competence for the Spanish-speaking population of Mexico. The study was conducted in two phases. In the first phase, a rule-based lexical detection algorithm was implemented, achieving an accuracy of 82% in emotion classification; however, it exhibited notable limitations in contextual understanding and dialogue quality. To ensure cultural adaptation, this phase required an extensive process of idiom mining from social media sources and data extraction, transformation, and loading (ETL). In the second phase, a hybrid architecture is proposed that integrates a Large Language Model (LLM), such as Google's Gemini, to enhance contextual awareness. This architecture combines the efficiency and precision of the lexical engine with the generative capabilities of artificial intelligence, enabling interactions that more closely simulate empathetic and culturally relevant dialogue.

Keywords: Artificial Intelligence, Emotional Health, Chatbot, LLM, Cultural Adaptation, Hybrid Model.

Licencia: © 2026. Este artículo se publica bajo la licencia **Creative Commons Atribución 4.0 Internacional (CC BY 4.0)**. https://creativecommons.org/licenses/by/4.0/

Fecha de Recepción: 26/11/2025. **Fecha de Aceptación:** 02/01/2026. **Fecha de Publicación:** 28/01/2026.

I. Introducción

Los trastornos mentales son una de las causas más importantes de discapacidad en todo el mundo, sobre todo entre los jóvenes [1]. El costo, el estigma social y la escasa disponibilidad de expertos en salud mental son factores que representan obstáculos importantes para obtener un apoyo a tiempo. En este contexto, la inteligencia artificial (IA), especialmente los agentes conversacionales (chatbots), surgen como una alternativa con mucho potencial para brindar un primer contacto empático, confidencial y de fácil acceso.

Este estudio documenta el diseño, la implementación y la evolución de PsicoloBot, un prototipo de chatbot creado para ser un puente tecnológico hacia el bienestar emocional, con una atención particular en la adaptación cultural y lingüística al español mexicano.

II. Metodología

A. Marco Teórico y Estado del Arte

El trabajo se basa en tres fundamentos teóricos: el estado actual de las tecnologías de IA en la salud mental, la estrategia para gestionar el Mindfulness y la esencia de la emoción.

Para crear un sistema que tenga la capacidad de "comprender" las emociones, es necesario primero determinar qué son. La investigación comenzó con los fundamentos filosóficos, examinando "Las pasiones del alma", una obra de Descartes que diferencia entre las funciones físicas y las espirituales (el pensamiento). Para Descartes [2], las pasiones (o emociones) son tipos de percepción que se originan a partir de la experiencia.

Desde una perspectiva contemporánea, definió la emoción como un proceso dinámico que se produce a partir de la interacción entre el cuerpo, la mente y el entorno. Este proceso se distingue por las reacciones fisiológicas que buscan dar sentido a nuestras vivencias.

Mindfulness (atención plena): Después de definir la emoción, el trabajo se centró en su manejo. La estrategia principal de intervención fue la Atención Plena (Mindfulness) [3], pues se ha evidenciado que es muy efectiva para disminuir la reactividad emocional y el estrés, así como para superar emociones negativas.

Los principios de Mindfulness que el bot, en su interacción ideal, intentaría promover son:

- Identificar y nombrar la emoción: El proceso de identificar y etiquetar una emoción (como la ira o la ansiedad) puede disminuir su intensidad y crear una sana distancia.
- Permitir y aceptar: Se fomenta una actitud de aceptación, reconociendo que una emoción es temporal, en vez de juzgarla como "buena" o "mala".
- Desidentificación de ideas: Mirar los pensamientos que vienen con la emoción como sucesos temporales de la mente, no como verdades absolutas.
- Contestar, no reaccionar: El propósito último es generar un espacio de conciencia que posibilite al usuario optar por su comportamiento, en vez de ser guiado por impulsos automáticos.

Competencia Cultural: El desarrollo de chatbots para la salud mental no es un área reciente. La efectividad de la terapia cognitivo-conductual (TCC) se ha evidenciado mediante herramientas como Woebot [4], que emplean agentes automatizados. Iniciativas como SERMO [5] han investigado arquitecturas que funcionan de manera local con el fin de asegurar la privacidad del usuario.

No obstante, la competencia cultural sigue siendo un desafío constante. La lengua, que está repleta de expresiones idiomáticas y modismos, tiene un papel importante en el reconocimiento de emociones. Para abordar esto, nuestra investigación inicial se basó en la utilización de léxicos estandarizados, como NRC-EmoLex [6], una herramienta que vincula palabras con emociones fundamentales, para desarrollar un sistema de detección propio.

B. Arquitectura de la Aplicación

La versión inicial del sistema se centró en un algoritmo de detección. Un pipeline de

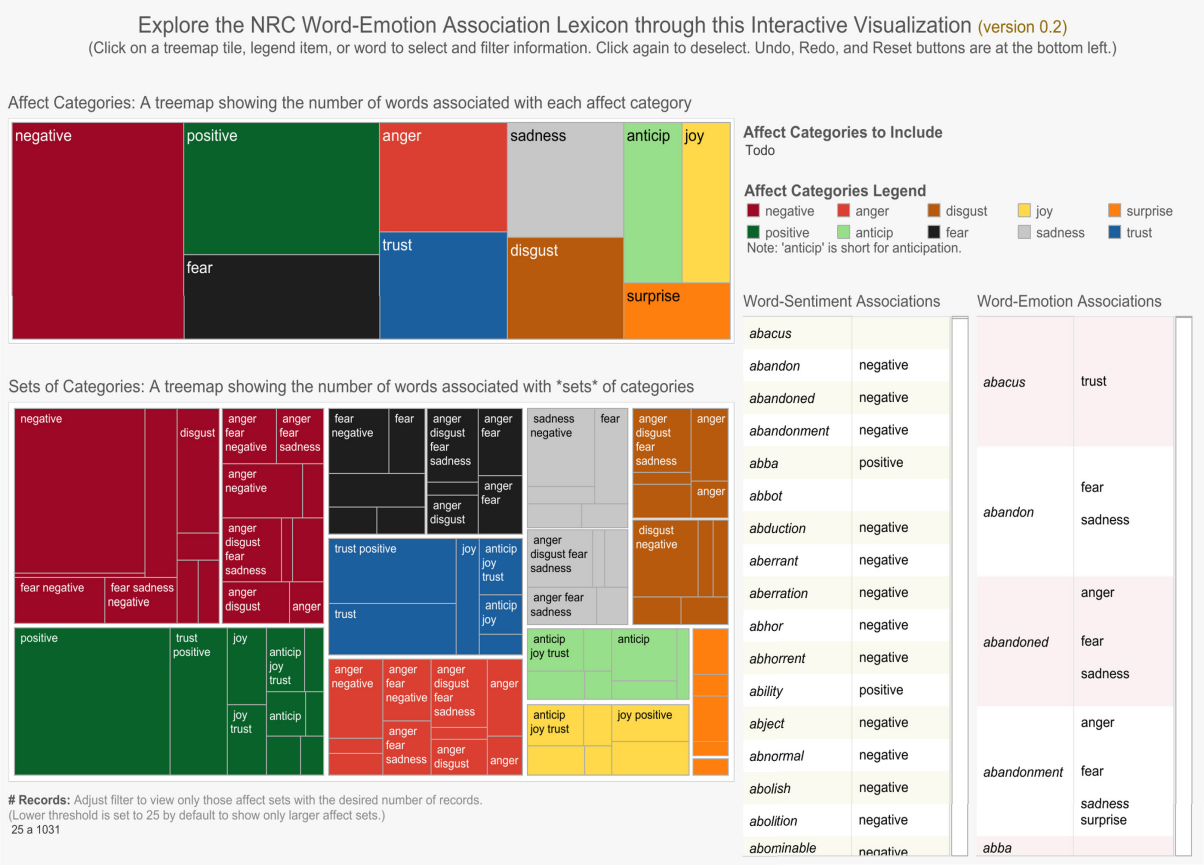


Fig. 1: *Léxico de asociación palabra-emoción del NRC.*

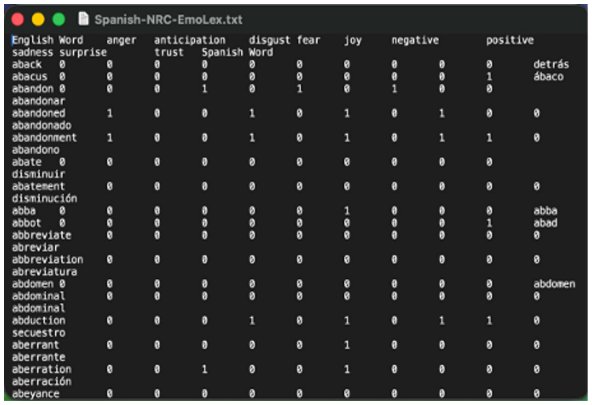


Fig. 2: *Spanish-NRC-EmoLex.*

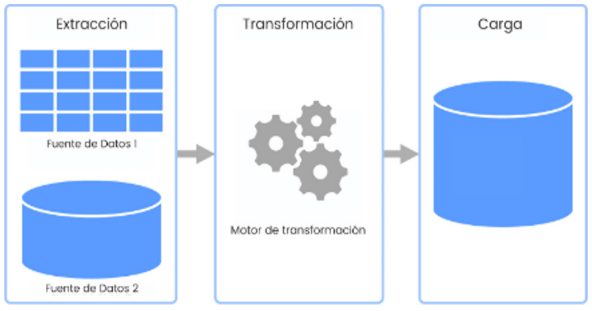


Fig. 3: *Proceso de transformación ETL.*

Procesamiento de Lenguaje Natural (PLN) implementado en C# y basado en reglas. La meta de esta estructura era habilitar la fase inicial del Mindfulness, que consiste en identificar y etiquetar las emociones.

Base de Conocimientos: El dataset de 3,871 términos fue un componente crítico desarrollado manualmente para asegurar la relevancia cultural. Su construcción se basó en tres pasos:

Uso de un Léxico Estandarizado: NRC-EmoLex es un recurso lingüístico esencial para el estudio de sentimientos, que fue creado por Saif M. Mohammad en el Consejo Nacional de Investigación canadiense (NRC). Su creación se fundamentó en el crowdsourcing (microtareas en línea), por medio del cual los participantes vincularon palabras con un grupo de emociones. El léxico resultante asocia ocho emociones fundamentales (asco, alegría, ira, tristeza, temor, sorpresa, confianza y anticipación) y dos

sentimientos (positivo y negativo) con más de 14 000 palabras en su versión en inglés (ver Fig. 1). Este léxico se traduce a más de cien idiomas diferentes, entre ellos el español. La asociación se presenta de manera binaria, 1 si la palabra evoca la emoción, 0 en caso contrario (ver Fig. 2).

Transformación del Léxico (De Binario a Relacional): Un paso metodológico crucial fue la transformación del formato del NRC-EmoLex para hacerlo compatible con el algoritmo de búsqueda de PsicoloBot. El léxico original era una matriz binaria ancha, ineficiente para las búsquedas directas y de similitud del algoritmo. Se desarrolló un proceso ETL (ver Fig. 3) para "des-pivotar" esta matriz. El script fue programado para Extraer (E) y leer el archivo de entrada .txt. Luego, su lógica Transformación (T) consistió en 'des-pivotar' la matriz binaria (iteró sobre cada palabra en español del léxico original); por cada palabra, revisó las columnas de las emociones objetivo (ira, miedo, tristeza, alegría, asco) y, si el valor era '1', generaba una nueva fila en el archivo CSV de destino con el formato “Entidad”, “Termino” (ej. miedo, abandonar). Una palabra asociada a múltiples emociones en el original (ej. 'miedo' y 'tristeza') generó múltiples filas en el nuevo archivo. Finalmente, se cargó (L) todos estos nuevos registros en un archivo .csv de destino, creando así la base de conocimientos compatible con el formato requerido por el algoritmo.

Adaptación cultural (minería de modismos): Después de la automatización del léxico formal, se llevó a cabo el paso más crucial para la competencia cultural: mejorar el conjunto de datos. Primeramente, se llevó a cabo un scraping automatizado (ver Fig. 4) de información pública en las plataformas que tienen una mayor difusión entre los jóvenes: Facebook, TikTok y X (anteriormente Twitter). El propósito de este algoritmo no es crear el conjunto final de datos, sino reunir un conjunto extenso de texto crudo que muestre el lenguaje coloquial, informal y natural empleado por la población meta (los jóvenes en México) para expresar sus sentimientos. Este material será la "materia prima" para el subsiguiente procedimiento de limpieza manual y curación.

Las plataformas que fueron elegidas para la extracción fueron aquellas con la mayor presencia en la demografía objetivo:

- X (antes Twitter): Es ideal por su naturaleza pública y textual, así como por su API (en la etapa de diseño), que hace más sencillo recoger flujos de texto en tiempo real.
- Facebook: En particular, los segmentos de comentarios en publicaciones virales, páginas públicas y grupos, que representan un abundante recurso de conversación coloquial.
- TikTok: El sistema más predominante para la audiencia más joven. La recolección se enfoca en la parte de comentarios de vídeos populares.

Definición de "Palabras Clave”: El scraping no se lleva a cabo de manera aleatoria. Se fundamenta en un listado de "palabras clave" previamente establecido. No se trata de modismos, sino de palabras o frases "gancho" que suelen ir antes o están relacionadas con una expresión emocional.

Ejemplos de semillas: "me siento", "estoy triste", "ando feliz", "qué coraje", "me da asco", "tengo miedo", "qué oso", "me agüita", "me encabrona", "qué chido", "me hierve la sangre".

Herramientas utilizadas:

- Lenguaje de Programación: Python
- Librerías de Extracción:
- Librerías de Limpieza Básica
  - Para X: Tweepy
  - Para Facebook y TikTok: Selenium
- Librerías de Limpieza Básica
  - re (Expresiones Regulares) para la limpieza preliminar.

C. Recolección Automatizada en Redes Sociales

Inicialización: carga la lista de palabras semilla en memoria y define el archivo de salida para el corpus crudo (ej. raw_corpus.txt), después inicializa los controladores de automatización (Selenium/Playwright).

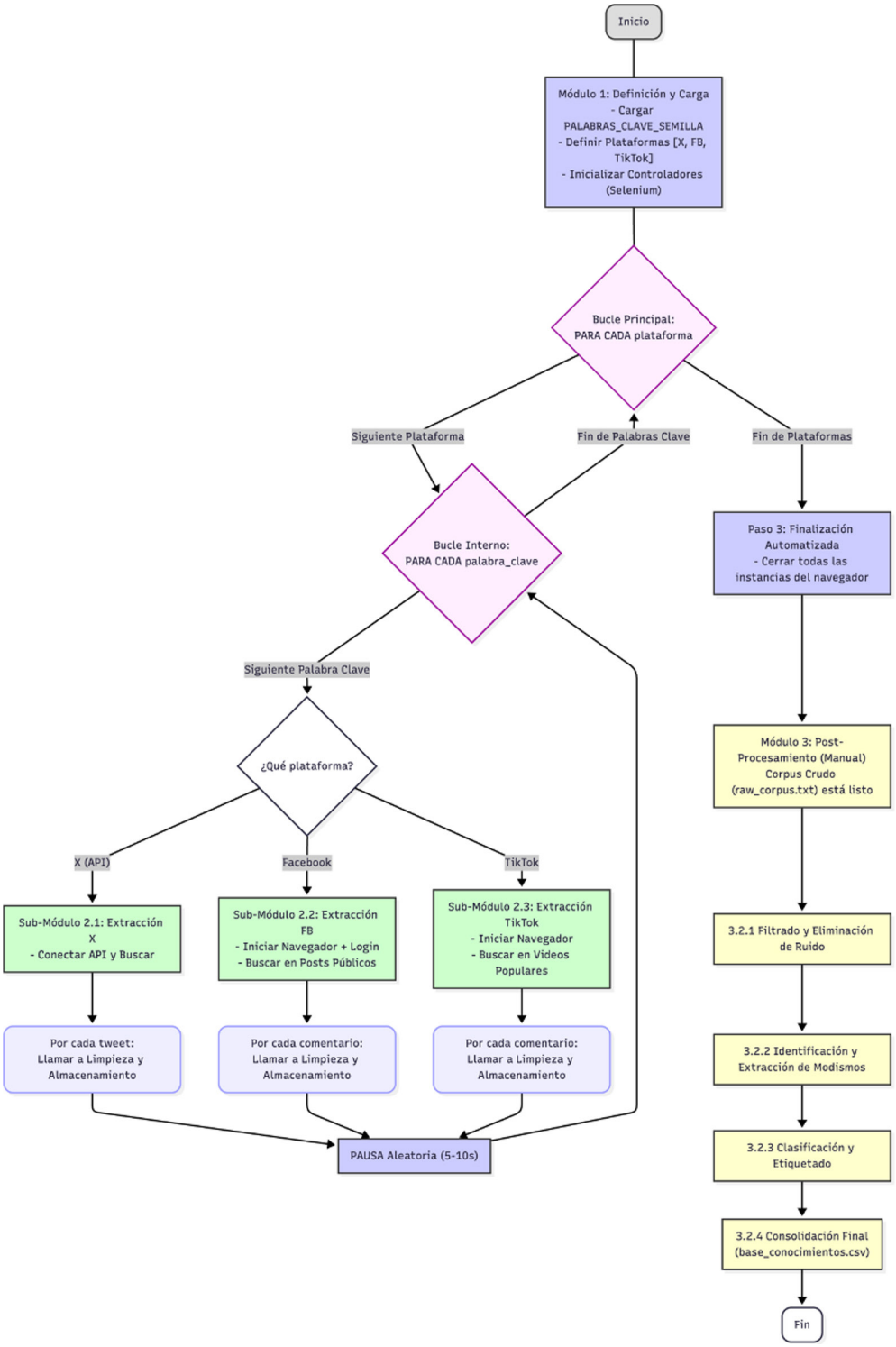


Fig. 4: *Algoritmo de Extracción y Recolección.*

Bucle Principal de Búsqueda: para cada plataforma [X, Facebook, TikTok] y para cada palabra clave se ejecuta el submódulo de búsqueda y extracción correspondiente (Módulo 2.1, 2.2 o 2.3). Pausa la ejecución por un tiempo aleatorio para evitar bloqueos de IP o rate limiting.• Sub-Módulo 2.1: Extracción de X (API), conecta con la API de X y ejecuta una búsqueda con la palabra clave, filtrando por idioma (lang='es') y un número máximo de resultados (ej. 200).

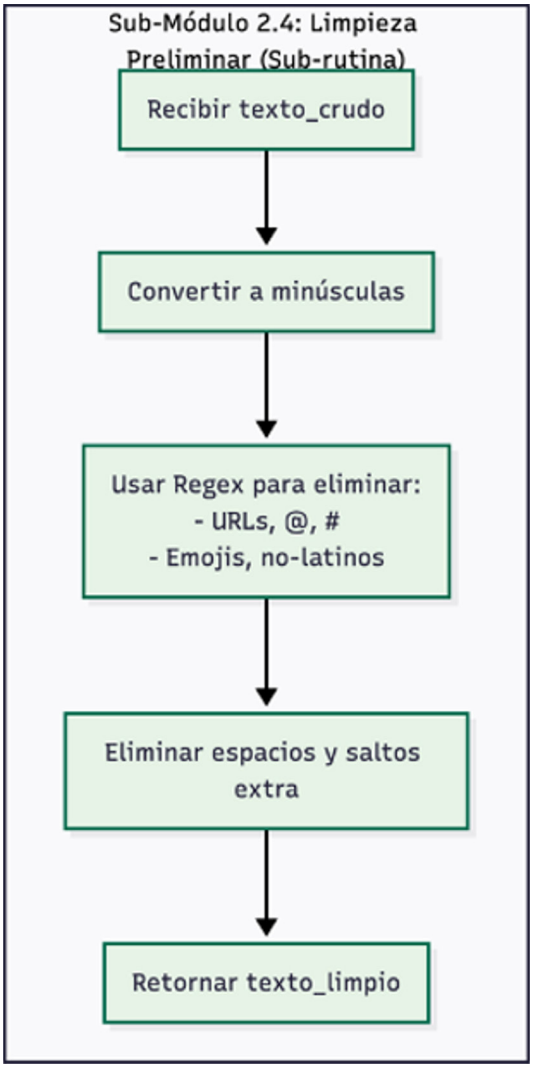


Fig. 5: *Limpieza Preliminar.*

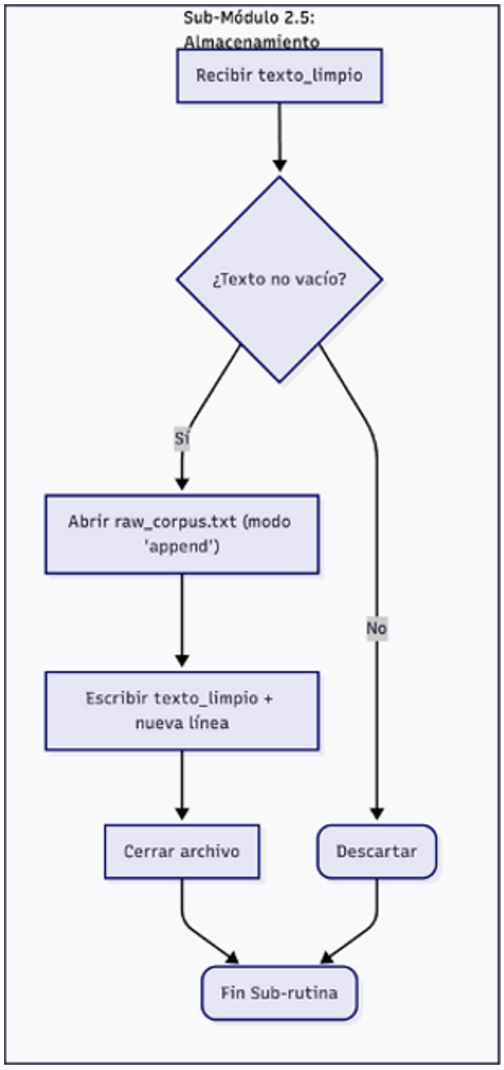


Fig. 6: *Almacenamiento.v*

Para cada tweet en los resultados de la búsqueda obtiene el texto completo del tweet y pasa el texto al Módulo 2.4 (Limpieza Preliminar), para posteriormente pasar el texto limpio al Módulo 2.5 (Almacenamiento).

- Sub-Módulo 2.2: Extracción de Facebook (Automatización de Navegador). Inicia el navegador controlado por Selenium, navega a Facebook y realiza el login (necesario para ver comentarios), con la función de búsqueda, busca la palabra clave, y filtra los resultados para mostrar solo "Publicaciones" de páginas públicas.

Para cada publicación en los resultados hace scroll para cargar los comentarios. Extrae el texto de todos los comentarios visibles y para cada comentario extraído pasa el texto al Módulo 2.4 - Limpieza Preliminar (ver Fig. 5) y al Módulo 2.5 – Almacenamiento (ver Fig. 6).

- Sub-Módulo 2.3: Extracción de TikTok (Automatización de Navegador). Inicia el navegador controlado por Selenium, navega a TikTok y busca la palabra clave (a menudo como hashtag), obteniendo una lista de los videos más populares o relevantes. Para cada video en la lista abre el video y hace scroll en su sección de comentarios para cargar dinámicamente las respuestas. Extrae el texto de todos los comentarios cargados y para cada comentario extraído pasa el texto al Módulo 2.4 y al Módulo 2.5.

Filtrado y Eliminación de Ruido: Se lee el raw_corpus.txt. Se eliminan manualmente entradas que son frases irrelevantes, o texto que usa la palabra semilla sin un contexto emocional (ej. "me siento a esperar"), se identifica y extrae las expresiones coloquiales genuinas que sí expresan una emoción (ej. "me siento agüitado", "qué oso", "me hierve la sangre"). Finalmente, cada expresión extraída se asocia manualmente con su Entidad (emoción) correspondiente.

Estas nuevas filas (Entidad, Término) se añaden al archivo base_conocimientos.csv (que ya contenía los datos transformados del NRC-EmoLex), formando así el dataset final y culturalmente adaptado de 3,871 términos.

D. Algoritmo de Detección Emocional

Herramientas utilizadas:

- Lenguaje: C#
- Lógica: Basada en reglas, procesamiento de cadenas (strings) y algoritmos de similitud de texto.

Requerimientos y Carga de Datos (Inicialización): Se prepara el sistema para el análisis, cargando en memoria la base de conocimientos que actúa como el "cerebro" léxico del algoritmo.

- Componente Clave: Base de Conocimientos
- Fuente: Este archivo CSV de 3,871 términos es el resultado de un proceso de ETL y curación.
- Contenido: Se compone de dos columnas principales: Entidad (la emoción, ej. "tristeza") y Término (la palabra o modismo asociado, ej. "me siento agüitado").

Como se puede observar en la Fig. 7, el primer paso es leer el archivo CSV y cargarlo en una estructura de datos eficiente en memoria (como un Dictionary o List en C#) para permitir búsquedas rápidas. Posteriormente se procesa la entrada del usuario:

- Pre-procesamiento de Texto (Limpieza), transforma el texto crudo (raw text) introducido por el usuario en una lista de palabras limpias ("términos de interés"), eliminando el "ruido" gramatical para preparar el análisis.

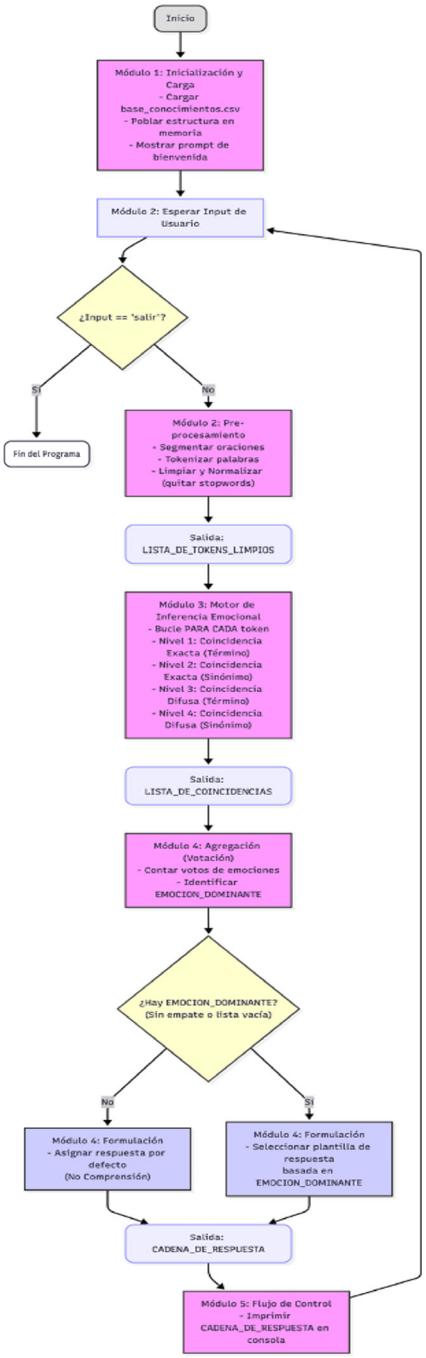


Fig. 7: *Arquitectura del algoritmo de detección emocional.*

- Captura de Entrada: El sistema recibe el texto completo del usuario.
- Segmentación de Oración: El bloque de texto se divide en oraciones individuales. Este paso es crucial si el usuario escribe varios pensamientos a la vez.
- Tokenización: Cada oración se descompone en sus palabras y símbolos individuales (tokens).

Normalización y Filtrado: Se itera sobre la lista de tokens para limpiarla

- Conversión a minúsculas: Todos los tokens se convierten a minúsculas para estandarizar la comparación.
- Eliminación de Puntuación: Se descartan todos los tokens que son signos de puntuación.
- Eliminación de Stopwords: Se compara cada token con una lista predefinida de "palabras vacías" (ej."la","hoy", "me", "muy", "el"). Si el token está en la lista de stopwords, se descarta.

Salida del Módulo: El resultado es una lista final de "términos de interés", que son las únicas palabras que se enviarán al motor de inferencia.

Módulo de Inferencia Emocional (Motor de 4 Niveles):

- Nivel 1 (Coincidencia exacta): El sistema buscaba una correspondencia directa y literal entre una palabra del usuario y un término emocional principal en la base de datos.
- Nivel 2 (Coincidencia difusa): En caso de no hallar una coincidencia directa, el sistema verifica si la palabra del usuario era idéntica a alguno de los sinónimos enumerados para un término emocional (como "pánico" con un sinónimo de "miedo").
- Nivel 3 (Coincidencia difusa - Variaciones): Si los niveles anteriores no funcionaban, el algoritmo emplea una búsqueda "difusa" o aproximada para hallar términos del usuario que tuviesen gran semejanza con una palabra clave, tolerando ligeras variaciones o errores tipográficos (por ejemplo, "mido" podría corresponderse con "miedo").
- Nivel 4 (Coincidencia difusa - Sinónimos): Como último recurso, el sistema realizaba una búsqueda difusa contra

E. Votación y respuesta:

Tras el análisis de tokens, la etapa de "Votación" consolida los resultados. El sistema cuenta el número de votos acumulados para cada emoción y declara una "emoción predominante" para formular una respuesta.

- Votación: Se cuentan las ocurrencias de cada emoción encontrada en la lista de salida del Módulo 3.
- Determinación de la Dominante: Se selecciona la emoción con el conteo más alto.
- Selección de Plantilla de Respuesta: El sistema utiliza la emoción dominante como clave para buscar en una lista de respuestas predefinidas.
- Presentación de Respuesta: Se muestra en la pantalla la emoción detectada al usuario.
- El algoritmo vuelve al Módulo 2 y espera la siguiente respuesta del usuario, reiniciando el ciclo hasta que el usuario escriba "salir

III. Resultados

En un conjunto de pruebas controlado, el algoritmo alcanzó una precisión del 82% en la clasificación correcta de la emoción dominante. Esta métrica, aunque alta, solo refleja la capacidad del sistema para clasificar palabras clave, no para comprender el contexto.

Algunas limitaciones que se encontraron fueron:

1. Ceguera al contexto: el sistema es incapaz de entender matices, sarcasmo, ironía o la relación entre diferentes partes del texto.
2. Incapacidad para Interpretar Negaciones: Una frase como "No estoy triste" sería incorrectamente interpretada, ya que el algoritmo identificaría el token "triste", pero ignoraría el token "no", resultando en un falso positivo para "tristeza".
3. Respuestas Genéricas y Robóticas: Al depender de plantillas fijas, las respuestas resultan repetitivas y superficiales, fallando en generar conexión o confianza. El ejemplo del poema (Ver figuras 8 y 9) ilustra esto: la detección de "alegría" (basada en la frecuencia de palabras como "amor", "mejor", "dulce") llevaba a la respuesta insustancial: "Percibo que podrías estar sintiendo algo como 'alegría'. ¿Quieres hablar de eso?" Esta respuesta ignora la complejidad del poema y falla en validar los sentimientos mixtos expresados.

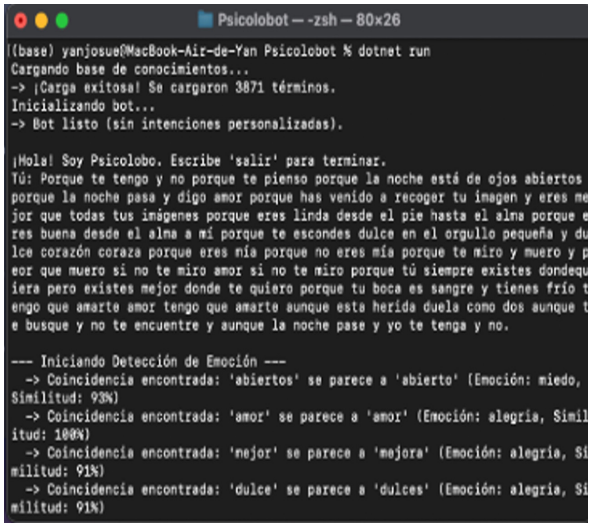


Fig. 8: *Respuesta del algoritmo basado en Léxico.*

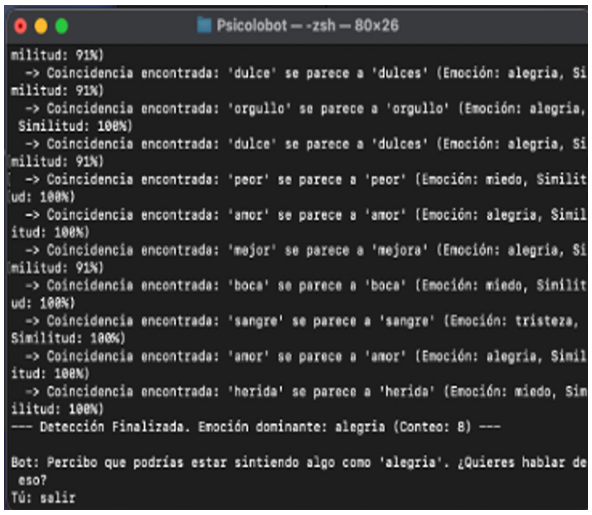


Fig. 9: *Respuesta del algoritmo basado en Léxico.*

IV. Conclusiones

Este trabajo de investigación evidencia que el procedimiento de clasificar emociones en un contexto lingüístico particular, usando un enfoque léxico híbrido que combina un léxico formal (NRC-EmoLex) con un corpus adaptado culturalmente (por medio de la curación manual y scraping), es una alternativa robusta y factible, logrando una precisión elevada (82%). La curación manual del conjunto de datos y la metodología ETL en Python fueron etapas clave para el éxito de esto.

No obstante, la conclusión más importante es que clasificar emociones no equivale a entender una conversación. La creación de una herramienta que simule un diálogo humano y establezca confianza no es posible con el enfoque léxico debido a sus limitaciones intrínsecas (incapacidad para manejar negaciones, ceguera al contexto o respuestas genéricas). Para superar las limitaciones de la Fase 1 sin descartar el valioso trabajo léxico y cultural realizado, el siguiente paso lógico es la evolución hacia una arquitectura híbrida. Esta propuesta combina la velocidad y especificidad cultural del motor de Fase 1 con la fluidez contextual y la capacidad de razonamiento de un Modelo de Lenguaje a Gran Escala (LLM) como Gemini de Google.

En lugar de que el LLM reemplace al motor léxico, este último actuaría como un "enriquecedor de contexto":

- Paso 1 (Análisis Léxico): El texto del usuario (ej. "neta, me siento bien agüitado") pasaría primero por el motor léxico de Fase 1.
- Paso 2 (Generación de Metadatos): Este motor no generaría una respuesta, sino un objeto de metadatos de análisis preliminar (ej. {"emocion_detectada": "tristeza", "termino_clave": "agüitado", "confianza": "alta", "tipo": "coloquial"}).
- Paso 3 (Inyección en Prompt): Estos metadatos se inyectarán en el Prompt de Sistema que se envía a la API de Gemini, junto con el texto original del usuario.
- Paso 4 (Respuesta Generativa): El LLM (Gemini) generaría la respuesta final. Al tener la ayuda del motor léxico, puede confirmar que "agüitado" es, en efecto, tristeza, y usar ese contexto para generar una respuesta mucho más precisa y empática.

Por lo tanto, el trabajo futuro se centrará en la implementación y evaluación de esta arquitectura híbrida, seguida de:

- Creación de una interfaz para el usuario (UX/UI): Desarrollar una aplicación móvil atractiva que incorpore ejercicios de Mindfulness.
- Pruebas formales de usabilidad: Evaluar cualitativamente la nueva arquitectura híbrida con usuarios reales.

IV. Referencias

[1] Organización Mundial de la Salud. (2022). Salud mental del adolescente. Obtenido de https://www.who.int/es/news-room/fact-sheets/detail/adolescent-mental-health

[2] Descartes, R. (1649), Las pasiones del alma, Alianza Editorial, Madrid, España. Partes I y II.

[3] Bresó, E., Gabriela, M. R. Y. and Andriani, J. (2013), 'La inteligencia emocional y la atención plena (mindfulness) como estrategia para la gestión de emociones negativas', Psicogente 16 (30).

[4] Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. JMIR Mental Health, 4(2), e19.

[5] Denecke, K., Vaaheesan, S., & Arulnathan, A. (2021). A Mental Health Chatbot for Regulating Emotions (SERMO): Concept and Usability Test. IEEE Transactions on Emerging Topics in Computing, 9(3), 1170-1182.

[6] Mohammad, S. M. (n.d.). NRC Emotion Lexicon. National Research Council Canada. https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm