



ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS

ECIFeelings

Proyecto de curso PTIA

Grupo 2

Integrantes:

Julian Camilo Lopez Barrero

Fecha de entrega: 9 del 12 de 2025

Declaración firmada

“Declaro que he escrito este trabajo de investigación por mí mismo, y que no he utilizado otras fuentes o recursos que los indicados para su preparación. Declaro que he indicado claramente todas las citas directas e indirectas, y que este documento no se ha presentado en otro lugar para fines de examen o publicación”

Nombre del estudiante & Firma:



Julian Camilo Lopez Barrero

Fecha: 15 de 09 de 2025

Índice de contenido

1. Introducción	4
1.1. Problemas a resolver	4
1.2. Implementación como proyecto de PTIA	5
2. Trabajos relacionados	7
2.1. Detección temprana del Trastorno Depresivo	7
2.1.1. Identifying Depression on Twitter	7
2.1.2. Deep Learning-Based Depression Detection from Social Media	7
2.2. Detección de noticias falsas	7
2.2.1. ClaimBuster: The First End-to-End Fact-checking System	7
2.2.2. Approaches to Identify Fake News: A Systematic Literature Review	8
2.3. Optimización de tráfico en grandes ciudades	8
2.3.1. ATSAC: Automated Traffic Surveillance and Control (Los Ángeles)	8
2.3.2. CoTV: Cooperative Control for Traffic Lights and CAVs	8
3. Descripción del problema	8
3.1. Trastorno Depresivo y su Detección	8
3.2. Criterios Para Resolver El Problema	10
3.3. Problema a Resolver	10
3.4. Importancia del Problema	10
4. Alcance y Objetivos	12
4.1. Objetivo general	12

4.2. Objetivo cuantitativo	12
4.3. Objetivo cualitativo	13
4.4. Modelo PEAS	13
5. Diseño metodológico	14
5.1. Estrategia general	14
5.2. Inteligencia de la solución	14
5.3. Herramientas y versiones utilizadas	14
5.4. Arquitectura de la solución	15
5.5. Componente principal	15
6. Análisis de resultados	16
6.1. Los dos mejores casos de prueba	16
6.2. El caso de prueba más significativo	17
7. Conclusiones	18
7.1. Aciertos	18
7.2. Conclusiones	18
8. Lista de Abreviaturas	20
9. Referencias bibliográficas	21
10. Anexos	22

1. Introducción

En el proyecto de principios y tecnologías de inteligencia artificial (PTIA), la selección de los problemas a analizar se orientó hacia desafíos identificados en la actualidad donde la inteligencia artificial puede generar un impacto social directo, medible y significativo. La rápida evolución tecnológica a través de los años ha permitido que los modelos de aprendizaje automático se integren en procesos sensibles relacionados con temas como la salud mental, la veracidad de la información y la gestión de tráfico en las ciudades modernas. Estas áreas representan problemáticas reales que afectan a millones de personas y que requieren soluciones innovadoras, escalables y éticamente responsables.

Por ello, los temas escogidos la detección temprana de depresión mas adelante nombrada como "Trastorno depresivo", la identificación automatizada de noticias falsas y la optimización del tráfico en grandes ciudades son representativos de tres dominios críticos: bienestar humano, integridad informativa y eficiencia de rutas urbanas. Cada uno de estos problemas no solo dispone de abundantes datos y avances recientes en inteligencia artificial (IA) que permiten su estudio, sino que también ofrece un alto potencial de aplicación práctica. Analizarlos permite comprender cómo los modelos inteligentes pueden contribuir a mejorar la calidad de vida, optimizar procesos, fortalecer la confianza pública y apoyar la toma de decisiones en entornos complejos. En conjunto, estos ítems constituyen una base sólida para explorar el alcance real de la inteligencia artificial en la solución de problemáticas de la actualidad.

1.1. Problemas a resolver

- **Detectar diagnosticos tales como el trastorno depresivo de forma temprana y evitar cualquier tipo de problemas mayores.**

La depresión constituye uno de los principales problemas de salud pública a nivel mundial debido a su alta prevalencia, los costos económicos asociados y su relación directa con tasas elevadas de discapacidad y mortalidad según el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSVM). Su detección tardía es uno de los retos más significativos, ya que muchos síntomas pasan desapercibidos para familiares, profesionales e incluso para el propio individuo. En este escenario, la inteligencia artificial ofrece herramientas valiosas capaces de identificar patrones sutiles que el ojo humano no logra reconocer fácilmente (World Health Organization [WHO], 2025).

Modelos avanzados de machine learning (ML) y deep learning (DL) pueden analizar diferentes modalidades de información: desde el lenguaje escrito en mensajes o publicaciones en la plataforma X anteriormente llamada (Twitter), señales vocales como tono, ritmo y pausas, hasta comportamientos digitales y datos fisiológicos. Estos sistemas permiten generar alertas tempranas, facilitando intervenciones oportunas y reduciendo el riesgo de episodios críticos. Como proyecto de PTIA, este enfoque integra impacto social, ética en el tratamiento de datos y aplicación directa de técnicas de procesamiento del lenguaje natural y análisis de señales.

Quiero desarrollar un prototipo que procese texto y evalúe indicadores de riesgo mediante un modelo entrenado con datasets clínicos o públicos. El objetivo de este proyecto es comprobar si la IA puede

clasificar estados emocionales y emitir alertas tempranas.

- **Un sistema de detección de noticias falsas.**

La proliferación de contenido digital ha hecho que la información falsa se propague rápidamente, generando confusión social, afectando procesos democráticos y deteriorando la confianza en los medios oficiales. Las noticias falsas pueden manipular la opinión pública, fomentar polarización y producir daños económicos y sociales significativos. Ante este problema, un sistema automatizado de verificación basado en IA se vuelve una herramienta esencial para mejorar la calidad de la información que consumen las personas.

Un sistema de detección de noticias falsas utiliza modelos de análisis semántico, clasificación de texto, verificación de fuentes y análisis de correlaciones internas del contenido. También se puede incorporar visión computacional para detectar imágenes manipuladas o generadas artificialmente.

Quiero construir un modelo de clasificación que determine si una noticia es verdadera o falsa utilizando un dataset etiquetado. Para ello, aplicaré técnicas como embeddings de texto, modelos preentrenados tipo BERT o RoBERTa y validación cruzada. La idea es crear una pequeña plataforma o script capaz de analizar noticias y presentar un veredicto acompañado de una explicación generada por IA.

- **Optimización de tráfico en grandes ciudades.**

Las grandes ciudades enfrentan graves problemas de congestión vehicular, lo que genera pérdidas económicas, incrementa la contaminación y disminuye la calidad de vida de los ciudadanos. Los sistemas tradicionales de movilidad, basados en ciclos fijos de semáforos o decisiones manuales, no son suficientes para adaptarse al comportamiento dinámico del tráfico en tiempo real.

Mediante inteligencia artificial, es posible procesar información proveniente de cámaras, sensores urbanos, dispositivos móviles y vehículos conectados. Con estos datos, modelos predictivos y algoritmos de optimización pueden anticipar congestiones, ajustar semáforos en tiempo real, proponer rutas alternativas y mejorar la eficiencia del sistema de movilidad. Como proyecto de PTIA, esta propuesta se alinea con la visión de ciudades inteligentes y demuestra cómo la IA puede resolver problemas urbanos complejos mediante análisis de datos masivos.

Quiero diseñar un sistema simplificado donde, a partir de datos simulados o reales de tráfico, se entrenen modelos que predigan la congestión en determinados puntos de la ciudad. A partir de esa predicción, se simulará la optimización del tiempo de semáforos o la sugerencia de rutas alternativas utilizando algoritmos de optimización. El objetivo es evaluar si estos modelos pueden mejorar el flujo vehicular en escenarios controlados o simulados.

1.2. Implementación como proyecto de PTIA

- Hoy en día el trastorno depresivo es una de las principales causas de tasas de mortalidad altas o de discapacidad según el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSVM). La

implementación del uso de inteligencia artificial es viable, pues podemos detectar mediante patrones sucesos que muchas veces pasan desapercibidos. Podríamos identificarlos mediante voz o texto para prevenir y dar un diagnóstico temprano.

- Las noticias falsas, en este auge de la inteligencia artificial, quitan veracidad a la información y reducen la confianza pública. Como proyecto, un sistema de verificación automática tendría un impacto social positivo, aumentando la confianza y precisión informativa.
- Mediante modelos actuales de inteligencia artificial se puede detectar la congestión y optimizar el flujo vehicular ajustando semáforos o rutas de movilidad. En una ciudad tan congestionada como Bogotá, esto puede aportar soluciones escalables y sostenibles.

2. Trabajos relacionados

Esta sección presenta una revisión de trabajos relacionados que abordan problemáticas similares desde distintas áreas de aplicación, incluyendo la detección temprana de depresión, la identificación de noticias falsas y la optimización del tráfico urbano. El objetivo de esta revisión es contextualizar el presente trabajo dentro del estado del arte, analizar las metodologías empleadas previamente y evaluar en qué medida los enfoques existentes han logrado cumplir sus objetivos, sirviendo como base y punto de comparación para la propuesta desarrollada en este proyecto.

2.1. Detección temprana del Trastorno Depresivo

2.1.1. Identifying Depression on Twitter

- Se quería desarrollar un sistema capaz de identificar si un usuario de Twitter presenta señales de depresión a partir de sus publicaciones, basándose en palabras clave, utilizando representaciones "Bag-of-Words" y clasificadores tradicionales de aprendizaje supervisado aplicados a un dataset de tweets y así crear una línea base que permita comprender qué tan precisa puede ser la detección temprana de depresión usando técnicas clásicas de procesamiento de lenguaje general para este trabajo relacionado se usaron aproximadamente 2.5 millones de tuits o más comúnmente llamadas publicaciones en esta plataforma se obtuvo un accuracy de el 81 % y además de reporto una precision de 0.86 (86 %). (Nadeem, 2016)

2.1.2. Deep Learning-Based Depression Detection from Social Media

- Se quería valorar si modelos basados en transformers pueden detectar depresión con mayor precisión que métodos tradicionales, entrenando y comparando modelos como RoBERTa (Robustly Optimized BERT Approach) sobre publicaciones de redes sociales etiquetadas para poder demostrar que los embeddings contextuales mejoran significativamente la detección automática de señales depresivas, este modelo logro un accuracy de alrededor del 90 % en el modelo más sencillo llamado Bernoulli Naive Bayes y en el modelo ReBERTa se logro el mayor % de accuracy con un 98.1 % y un recall de 99 % . (Bokolo, 2023)

2.2. Detección de noticias falsas

2.2.1. ClaimBuster: The First End-to-End Fact-checking System

- En este se quería rearmar un sistema capaz de identificar afirmaciones verificables dentro de discursos y textos públicos, aplicando técnicas de procesamiento de lenguaje general con extracción de características lingüísticas y clasificadores supervisados y de esta forma facilitar el proceso de verificación automática y priorizar qué afirmaciones deben ser revisadas por fact-checkers humanos el sistema finalmente logro un 79 % de precision cuando se declaró una afirmación como "worth checking" y en un total de las afirmaciones confirmadas a ser fakeNews se obtuvo un 74 % . (Hassan et al., 2017)

2.2.2. Approaches to Identify Fake News: A Systematic Literature Review

- Se quería analizar y comparar diferentes métodos para detectar noticias falsas, realizando una revisión sistemática de enfoques basados en Bag-of-Words, reglas manuales y modelos de deep learning como BERT (Bidirectional Encoder Representations from Transformers) y RoBERTa para determinar qué técnicas logran mayor efectividad y guiar el desarrollo de futuros sistemas de detección de fake news para este trabajo finalmente se reportaron accuracies superiores al 95 % pero dado a la heterogeneidad entre estudios y los diferentes datasets manejados esta es muy alta por ende los resultados no pueden ser generalizables en todos los contextos. (de Beer et al., 2020)

2.3. Optimización de tráfico en grandes ciudades

2.3.1. ATSAC: Automated Traffic Surveillance and Control (Los Ángeles)

- Se implementó un sistema de control centralizado para mejorar el flujo vehicular en la ciudad utilizando sensores de carretera o también llamados "loop detectors" y algoritmos adaptativos que ajustan semáforos en tiempo real para reducir tiempos de viaje y mejorar la eficiencia de la movilidad urbana, este sistema se desplegó en el departamento de transporte de Los Angeles (LADOT) y utiliza más de 4500 a 5000 intersecciones semaforizadas con loop detectos lo que significa que ajusta los ciclos semaforicos en tiempo casi real, redujo los tiempos de viaje en un 7 % y 12 % en vías principales y es importante saber que no se evaluó con métricas como "accuracy" sino con el tiempo promedio de viaje, retraso vehicular y flujo y capacidad entre intersecciones.

2.3.2. CoTV: Cooperative Control for Traffic Lights and CAVs

- Desarrollar un sistema cooperativo entre semáforos y vehículos autónomos para optimizar el tráfico urbano, empleando el aprendizaje por refuerzo profundo en un entorno multi-agente con vehículos conectados para disminuir congestión, tiempos de recorrido y consumo energético mediante coordinación inteligente el sistema empleó el aprendizaje por refuerzo profundo y los resultados que se reportaron fueron la reducción de tiempo promedio de hasta 20-30 % con menor congestión vehicular y la reducción del consumo energético y de paradas innecesarias. (Guo et al., 2022)

3. Descripción del problema

3.1. Trastorno Depresivo y su Detección

El trastorno depresivo es un trastorno mental caracterizado por la presencia persistente de síntomas emocionales, cognitivos y conductuales que afectan de manera significativa la forma en que una persona piensa, siente y actúa. Entre los síntomas más comunes se encuentran la tristeza profunda y prolongada, la pérdida de interés o placer en actividades cotidianas, alteraciones del sueño y del apetito, fatiga constante, dificultad para concentrarse, sentimientos de inutilidad o culpa excesiva, y pensamientos recurrentes de

muerte o suicidio. Para que un episodio depresivo sea clínicamente significativo, estos síntomas deben mantenerse la mayor parte del día durante al menos dos semanas consecutivas.

Desde el punto de vista clínico, la depresión no siempre se manifiesta de forma continua. En muchos casos, el trastorno presenta un curso episódico, donde los síntomas aparecen durante un período prolongado y luego disminuyen parcial o totalmente, dando lugar a remisiones temporales. Una persona puede experimentar pausas de varios meses, incluso de seis meses o más, sin síntomas evidentes, y posteriormente volver a presentar recaídas. Cuando los síntomas se mantienen de forma intermitente o recurrente durante un período de dos años o más, puede considerarse un trastorno depresivo persistente o una depresión recurrente, lo que incrementa el riesgo de cronificación si no se recibe tratamiento adecuado (American Psychiatric Association, 2022).

La detección del trastorno depresivo resulta especialmente compleja debido a que muchas personas no expresan abiertamente su malestar emocional. En lugar de ello, las señales suelen manifestarse de manera indirecta, a través de cambios en el lenguaje, la forma de comunicarse, el contenido emocional de los textos escritos, el aislamiento social o la disminución de la interacción con su entorno. En contextos digitales, como redes sociales, foros o plataformas de mensajería, estos cambios pueden reflejarse en el uso de palabras con carga emocional negativa, expresiones de desesperanza, pesimismo recurrente, referencias al cansancio extremo o a la falta de sentido de la vida (World Health Organization, 2025).

En este contexto, las herramientas basadas en procesamiento de lenguaje natural y aprendizaje automático ofrecen una oportunidad valiosa para apoyar la detección temprana de posibles indicios de depresión. A través del análisis longitudinal de datos textuales, es posible identificar patrones persistentes o cíclicos en el lenguaje de una persona a lo largo del tiempo, incluso cuando existen pausas prolongadas entre episodios. Este enfoque permite no solo detectar síntomas activos, sino también reconocer señales tempranas de recaída después de períodos de aparente estabilidad, contribuyendo así a una intervención más oportuna (Chancellor & De Choudhury, 2020).

Es importante destacar que estas tecnologías no reemplazan el diagnóstico clínico realizado por profesionales de la salud mental, sino que actúan como sistemas de apoyo que facilitan la identificación temprana de riesgos. Su correcto uso, junto con criterios éticos y de privacidad, puede convertirse en una herramienta complementaria para la prevención, el seguimiento y la orientación temprana de personas que podrían estar atravesando un trastorno depresivo.

Los criterios que voy a usar para seleccionar el problema a resolver son los siguientes:

- **Relevancia social:** Este criterio se centra en el impacto directo del problema sobre la población y en la magnitud de sus consecuencias. Un problema con alta relevancia social afecta a un número significativo de personas, influye en su bienestar y genera repercusiones en áreas clave como la salud, la educación, la seguridad y la calidad de vida. Seleccionar un problema socialmente relevante garantiza que el proyecto aporte beneficios tangibles y pueda convertirse en una herramienta útil para mitigar una situación real y urgente dentro de la comunidad.
- **Factor ético:** Considera las implicaciones morales involucradas en la elección del problema y en el

uso de tecnologías, especialmente inteligencia artificial. Un problema éticamente relevante es aquel donde el uso responsable, transparente y justo de los datos y algoritmos puede generar una mejora significativa en la vida de las personas sin vulnerar su privacidad, autonomía o dignidad. Incluye la evaluación de riesgos, como el tratamiento de información sensible, el posible sesgo en los modelos y las consecuencias de decisiones automatizadas. Seleccionar un problema con un enfoque ético permite desarrollar soluciones responsables y alineadas con los principios de la IA confiable.

- **Relevancia académica:** Este criterio evalúa el aporte del problema al campo del conocimiento y su potencial para explorar teorías, técnicas y metodologías actuales de inteligencia artificial. Un problema con alta relevancia académica representa un reto técnico interesante, posibilita la experimentación con modelos de aprendizaje automático modernos y contribuye al desarrollo de nuevas líneas de investigación o a la mejora de enfoques existentes. Este criterio garantiza que el proyecto no solo sea útil en la práctica, sino que también permita profundizar en aspectos formativos propios de PTIA.

3.2. Criterios Para Resolver El Problema

El problema seleccionado es “detectar sentimientos o enfermedades tales como el trastorno depresivo de forma temprana y evitar cualquier tipo de problemas mayores”, debido a que la depresión constituye actualmente uno de los principales problemas de salud pública a nivel mundial. Su detección temprana resulta de gran relevancia, ya que en muchos casos los síntomas pasan desapercibidos y, sin una intervención oportuna, pueden derivar en consecuencias graves como el aislamiento social, la disminución del rendimiento académico o laboral e incluso conductas suicidas.

3.3. Problema a Resolver

El gran problema que se busca resolver es la dificultad para identificar de forma temprana los síntomas de estos trastornos en las personas. En la actualidad, muchas señales de alerta como cambios en el lenguaje, el tono de voz, la forma de escribir o la interacción que tienen estas personas en redes sociales o en foros activos suelen pasar inadvertidas para familiares, docentes o profesionales de la salud. Esto provoca que la mayoría de los casos se detecten cuando la enfermedad ya ha avanzado, lo que dificulta su tratamiento oportuno y aumenta el riesgo de consecuencias psicológicas y sociales más severas.

3.4. Importancia del Problema

Resolver este problema es de suma importancia, ya que la depresión es una de las principales causas de discapacidad y mortalidad en el mundo, afectando la calidad de vida de millones de personas. La implementación de soluciones basadas en inteligencia artificial (IA) permitiría detectar patrones de riesgo de forma temprana, brindando apoyo, orientación y acompañamiento antes de que la situación empeore y así generar planes de acción. (World Health Organization [WHO], 2025)

Este enfoque no solo tendría un impacto positivo en la salud mental individual, sino también en los ámbitos social, educativo y laboral, al reducir los costos asociados a tratamientos tardíos, el ausentismo

y la pérdida de productividad.

4. Alcance y Objetivos

El problema que se busca abordar en este proyecto es la detección temprana de posibles indicadores de depresión a partir de textos escritos por los usuarios, o a partir de la forma en que estos interactúan con diferentes modelos de inteligencia artificial, utilizando técnicas de procesamiento de lenguaje natural (PLN) e inteligencia artificial (IA).

La propuesta se centra en analizar la manera en que las personas se expresan mediante texto —por ejemplo, publicaciones en redes sociales como X (anteriormente Twitter), mensajes, comentarios o interacciones con modelos de lenguaje de gran escala (LLM, Large Language Models)— con el fin de identificar patrones lingüísticos que puedan estar asociados con estados depresivos o emociones negativas.

Este enfoque se considera una simplificación adecuada del problema, ya que se limita exclusivamente al análisis textual, dejando de lado factores multimodales como el tono de voz, las expresiones faciales o las señales fisiológicas. Esta delimitación permite reducir la complejidad del sistema y facilita el desarrollo de un modelo funcional capaz de analizar grandes volúmenes de texto de manera eficiente y precisa. Además, el texto constituye una de las formas más accesibles y frecuentes de expresión humana, lo que lo convierte en una fuente valiosa para el análisis de emociones y estados mentales.

4.1. Objetivo general

Desarrollar un modelo basado en técnicas de procesamiento de lenguaje natural (PLN) e inteligencia artificial (IA) que permita detectar de manera temprana posibles indicios de depresión a partir del análisis de textos escritos por los usuarios, con el fin de contribuir a la prevención y a la atención oportuna de este diagnóstico, que actualmente afecta a un número creciente de personas a nivel mundial.

4.2. Objetivo cuantitativo

Evaluar el rendimiento del modelo de detección mediante métricas cuantitativas como precisión, sensibilidad, especificidad y buscando alcanzar un porcentaje de acierto superior al 80 % en la clasificación de textos con y sin indicios de depresión. Estas métricas permitirán medir de forma objetiva la capacidad predictiva y la fiabilidad del sistema propuesto.

La meta de alcanzar valores cercanos o superiores al 90 % en dichas métricas se fundamenta tanto en el estado del arte como en la naturaleza crítica del problema abordado. Estudios previos en detección automática de depresión a partir de texto han demostrado que los enfoques basados en aprendizaje profundo y modelos de lenguaje avanzados, como los transformers, pueden superar ampliamente el umbral del 80 %, alcanzando niveles de desempeño elevados en conjuntos de datos controlados. Por lo tanto, establecer este objetivo cuantitativo resulta coherente con los avances recientes y técnicamente alcanzable.

Desde una perspectiva aplicada, la detección temprana de depresión conlleva riesgos significativos asociados a errores de clasificación. Un falso negativo puede impedir la identificación oportuna de una persona

con síntomas depresivos, mientras que un falso positivo puede generar alertas innecesarias. En este contexto, aspirar a un desempeño cercano al 90 % contribuye a reducir ambos tipos de error, aumentando la fiabilidad del sistema y su potencial uso como herramienta de apoyo en entornos reales.

Finalmente, fijar un objetivo exigente permite garantizar que el sistema propuesto represente una mejora tangible frente a enfoques tradicionales, fortaleciendo su aporte científico y práctico. Un rendimiento cercano al 90 % asegura que el modelo no se limite a ser funcional, sino que sea lo suficientemente robusto y confiable para apoyar procesos de análisis y toma de decisiones en escenarios reales.

4.3. Objetivo cualitativo

Analizar la interpretabilidad y la coherencia semántica de los resultados obtenidos por el modelo, verificando que las conclusiones generadas correspondan con los patrones lingüísticos y emocionales esperados. Asimismo, se busca que el sistema aporte un valor ético y social, funcionando como una herramienta de apoyo para la detección temprana y no como un sustituto del diagnóstico clínico.

4.4. Modelo PEAS

Con el fin de caracterizar formalmente el sistema de detección de depresión desde una perspectiva de agentes inteligentes, se emplea el modelo PEAS (Performance, Environment, Actuators, Sensors). Este modelo permite describir de manera estructurada el comportamiento esperado del sistema en el contexto específico de redes sociales, particularmente la plataforma X (anteriormente Twitter).

Medidas de desempeño El desempeño del sistema se evalúa mediante métricas cuantitativas como precisión, priorizando la reducción de falsos negativos y falsos positivos. Adicionalmente, se considera la estabilidad del modelo frente a lenguaje informal, ambigüedad semántica y variaciones en la longitud y el contexto de los textos.

Entorno El entorno del agente corresponde a la plataforma X (Twitter), caracterizada por publicaciones textuales cortas, lenguaje informal, uso frecuente de abreviaciones, emoticonos, hashtags y referencias implícitas a estados emocionales. Este entorno es dinámico, parcialmente observable y no determinista, dado que el contenido generado por los usuarios varía constantemente y depende de factores externos como tendencias, eventos sociales y contextos personales.

Actuadores Las acciones del sistema consisten en la clasificación automática de textos en categorías asociadas a la presencia o ausencia de indicios de depresión. Como salida, el sistema puede generar etiquetas, puntajes de probabilidad o alertas internas que indiquen un posible riesgo, con el objetivo de servir como herramienta de apoyo para análisis posteriores, sin realizar diagnósticos clínicos.

Sensores Los sensores del sistema están representados por los textos publicados por los usuarios en X, incluyendo contenido textual, hashtags, menciones y metadatos asociados al texto (como frecuencia de publicación o temporalidad). Estos elementos constituyen la entrada principal para el análisis lingüístico y semántico realizado por el modelo de procesamiento de lenguaje natural.

5. Diseño metodológico

5.1. Estrategia general

La estrategia general seleccionada para la solución se basa en el uso de técnicas de **procesamiento de lenguaje natural (PLN)** y **Deep Learning** para analizar textos escritos por los usuarios e identificar patrones que puedan asociarse con síntomas depresivos en los usuarios. El enfoque propuesto combina una etapa de preprocesamiento del texto, donde se limpian y normalizan los datos, con una fase de entrenamiento de una red neuronal recurrente (LSTM), capaz de aprender las dependencias basadas en el contexto y las emociones presentes en el lenguaje. Este enfoque es adecuado porque permite automatizar la detección temprana de señales emocionales y previene casos más extremos asociados al trastorno.

Para el entrenamiento y validación del modelo se utilizará el conjunto de datos “**Spanish Tweets Suggesting Depression**”, disponible en la plataforma Kaggle, el cual contiene publicaciones en español etiquetadas según la presencia de indicios depresivos. Este dataset resulta especialmente adecuado debido a su enfoque en el idioma español y a que los textos provienen de un entorno real de redes sociales, lo cual fortalece la validez práctica del sistema propuesto.

5.2. Inteligencia de la solución

La inteligencia de la solución se concentra en el modelo de red neuronal recurrente (LSTM) que se va a usar y en la fase del preprocesamiento semántico del texto que ingresa a el modelo. El conocimiento proviene del conjunto de datos etiquetado, donde los textos están clasificados según su carga emocional o presencia de indicios depresivos. El razonamiento se da dentro del modelo LSTM, que aprende de manera automática las relaciones entre palabras, frases y patrones emocionales, generando predicciones fundamentadas en el contexto.

5.3. Herramientas y versiones utilizadas

Para garantizar estabilidad, compatibilidad y soporte a largo plazo, se seleccionaron versiones ampliamente utilizadas y consideradas estables dentro del ecosistema de aprendizaje profundo y procesamiento de lenguaje natural. Las versiones empleadas son:

- Python : versión estable y totalmente compatible con la mayoría de bibliotecas modernas de inteligencia artificial.
- TensorFlow / Keras : versión estable que soporta modelos LSTM, embeddings y entrenamiento optimizado.
- spaCy : recomendado para tareas de tokenización, lematización y análisis lingüístico robusto.
- NLTK : usado para procesamiento lingüístico clásico (stopwords, tokenización simple).
- NumPy : compatible con TensorFlow y operaciones matriciales de alto rendimiento.

Estas versiones han sido seleccionadas debido a su estabilidad comprobada, extensa documentación y soporte garantizado por sus respectivas comunidades y desarrolladores.

5.4. Arquitectura de la solución

La arquitectura general del sistema se compone de las siguientes secciones:

- Sección de adquisición de datos: encargada de utilizar el conjunto de datos “Spanish Tweets Suggesting Depression” de Kaggle, compuesto por publicaciones en español extraídas de Twitter (X) y etiquetadas según la presencia o ausencia de indicios de depresión.
- Sección de preprocesamiento: realiza la limpieza y normalización del texto (eliminación de símbolos, lematización, tokenización, conversión a minúsculas y eliminación de palabras vacías).
- Sección de representación vectorial: convierte las palabras en vectores numéricos mediante técnicas de embeddings, preservando relaciones semánticas entre términos.
- Sección de análisis y detección: implementa la red neuronal LSTM, que analiza las secuencias textuales y determina la probabilidad de que el texto contenga indicios de depresión.
- Sección de evaluación y visualización: calcula métricas de rendimiento como precisión ya que este nos permite obtener un porcentaje adecuado para el análisis y veracidad de los resultados obtenidos

Estas secciones interactúan de forma secuencial, asegurando un flujo de información desde la entrada de texto hasta la salida de resultados interpretables que es la función del LSTM.

5.5. Componente principal

El componente principal es la **red neuronal recurrente basada en LSTM**. Su función es procesar las secuencias de texto representadas numéricamente y detectar patrones temporales o emocionales asociados con estados depresivos.

Funcionamiento del componente principal:

- Cada palabra es convertida en un vector mediante los llamados embeddings, capturando relaciones semánticas.
- La secuencia de vectores es procesada por las capas LSTM, que retienen información contextual a lo largo del texto.
- La salida pasa a una capa densa con activación sigmoide o softmax, la cual clasifica el texto según la probabilidad de contener signos depresivos.
- Se optimiza el modelo mediante retropropagación, usando una función de pérdida binaria (por ejemplo, binary cross-entropy).

6. Análisis de resultados

El modelo basado en redes neuronales recurrentes (LSTM) fue evaluado utilizando un conjunto de datos de prueba, logrando métricas de desempeño satisfactorias tras la aplicación de técnicas de regularización (*Dropout*) e inyección de ruido controlado, cabe resaltar que los resultados obtenidos se pueden observar correctamente y de una manera mas detallada en el archivo ECIFEELINGS.ipynb donde se encuentran las pruebas las cuales seran analizadas mas adelante. A continuación, se detallan los casos de prueba más relevantes que demuestran la capacidad de generalización del modelo.

- Mejores casos de prueba: Se seleccionaron aquellos donde el modelo mostró una confianza superior al 90 % en su clasificación, demostrando una clara distinción entre patrones depresivos y estados de ánimo positivos.
- Caso más significativo: Se analiza un caso de borde donde la presencia de la palabra "triste" no implicó una clasificación de depresión, evidenciando la capacidad del modelo para entender el contexto semántico.

6.1. Los dos mejores casos de prueba

A continuación se presentan dos instancias donde la red neuronal identificó correctamente la clase con un alto grado de certeza (probabilidad cercana a los extremos 0 o 1).

- **Caso de Prueba A (Detección de Depresión):**
 - Entrada: "Ya no tengo fuerzas para levantarme de la cama, solo quiero dormir."
 - Salida Esperada: 1 (Depresión)
 - Salida Obtenida: 1 (Depresión) - Probabilidad: 92.0%
 - Análisis: El modelo identificó correctamente este tweet debido a la presencia de síntomas clínicos claros como la fatiga crónica y la abulia (falta de voluntad). La capa LSTM logró relacionar la secuencia temporal de "no tengo fuerzas" con "solo quiero dormir", activando fuertemente las neuronas asociadas a la clase positiva.
- **Caso de Prueba B (Detección de Estado Normal):**
 - Entrada: "Hoy es un día increíble para salir a caminar con amigos."
 - Salida Esperada: 0 (No Depresión)
 - Salida Obtenida: 0 (No Depresión) - Probabilidad de Depresión: 14.8%
 - Análisis: El modelo asignó una probabilidad muy baja de depresión. Esto se debe a que los *embeddings* vectoriales de palabras como "increíble", "salir" y "amigos" se agrupan en un espacio semántico positivo o neutro, lejos de los vectores asociados a síntomas depresivos.

6.2. El caso de prueba más significativo

Este caso se considera el más relevante cualitativamente porque demuestra que el modelo no funciona simplemente como un "buscador de palabras clave", sino que interpreta el contexto gracias a la memoria de la arquitectura LSTM.

- Entrada: "Estoy triste porque mi equipo de fútbol perdió el partido."
- Salida Esperada: 0 (No Depresión)
- Salida Obtenida: 0 (No Depresión) - Probabilidad de Depresión: 20.1%
- Análisis del Resultado: Este resultado es significativo porque la frase contiene explícitamente la palabra **"triste"**, la cual tiene una alta correlación con la depresión en el conjunto de entrenamiento. Sin embargo, el modelo clasificó la frase correctamente como "No Depresión".

Esto indica que la red LSTM retuvo información del contexto posterior ("equipo de fútbol", "perdió", "partido"). El mecanismo de compuertas de la LSTM permitió "olvidar" la carga negativa de la palabra "triste" al asociarla con un evento deportivo externo y temporal, diferenciándola de la tristeza patológica o crónica que caracteriza al trastorno depresivo.

7. Conclusiones

A partir de la implementación del modelo LSTM y las pruebas realizadas, se presenta la siguiente reflexión sobre el desempeño del sistema y las lecciones aprendidas durante el proceso de desarrollo. Adicionalmente, con el fin de complementar el análisis cualitativo del desempeño del modelo, se realizó una revisión de casos de prueba específicos extraídos de la fase de validación. Estos casos representativos permiten observar de manera más concreta cómo el modelo LSTM responde ante diferentes tipos de expresiones lingüísticas, incluyendo síntomas clínicos explícitos, estados emocionales positivos y construcciones más complejas como negaciones gramaticales y lenguaje figurado. En términos generales, los resultados muestran un buen desempeño en la detección de patrones asociados a depresión, alcanzando altos niveles de confianza en frases que expresan auto-odio, apatía o síntomas somáticos, como la falta de apetito.

No obstante, también se evidencian limitaciones importantes del modelo en contextos semánticos particulares. Específicamente, se observaron falsos positivos en frases que incluyen negaciones explícitas (por ejemplo, “No tengo depresión, me siento genial”) y en expresiones metafóricas de uso coloquial (como “Morí de risa”). En estos casos, ciertas palabras con alta carga semántica negativa activaron representaciones internas asociadas a la clase de depresión, elevando la probabilidad de predicción por encima del umbral de decisión, aun cuando el contexto global de la frase no correspondía a un estado depresivo real. Estos resultados ponen en evidencia oportunidades de mejora futuras, particularmente en el manejo de negaciones y en la comprensión de lenguaje figurado.

7.1. Aciertos

- **Acierto 1: Selección de Arquitectura LSTM** La elección de una Red Neuronal Recurrente (LSTM) fue acertada para el procesamiento de secuencias de texto. A diferencia de modelos básicos de frecuencia de palabras, la LSTM demostró capacidad para retener contexto a largo plazo, permitiendo identificar patrones depresivos complejos más allá de palabras clave aisladas. *Aprendizaje:* En problemas de Procesamiento de Lenguaje Natural (PLN) donde el orden y el contexto importan, las arquitecturas recurrentes ofrecen una ventaja significativa sobre modelos estáticos.
- **Acierto 2: Manejo del Desbalance de Datos (Inyección de Ruido)** Dado que el dataset original contenía exclusivamente muestras positivas (depresión), la generación de datos sintéticos y la estrategia de inyección de ruido (label noise) fueron cruciales. Esto permitió reducir un accuracy artificial del 100 % a un 85-90 % más realista, evitando que el modelo simplemente memorizara las entradas y esto condujera a un overfitting. *Aprendizaje:* Un modelo perfecto durante el entrenamiento suele ser síntoma de datos deficientes. La robustez del modelo depende tanto de la calidad de los datos de la clase “negativa” (no depresión) como de la clase “positiva”.

7.2. Conclusiones

1. **Viabilidad de la Detección Temprana:** Los resultados confirman que es posible identificar patrones lingüísticos asociados al trastorno depresivo utilizando aprendizaje profundo. El modelo logró supe-

rar el umbral del 80% de precisión propuesto en los objetivos cuantitativos, demostrando que el lenguaje escrito en redes sociales es un biomarcador válido.

2. El Modelo como Herramienta de Apoyo, no Diagnóstico: A pesar de los buenos resultados métricos, los falsos positivos detectados en el análisis cualitativo refuerzan que este sistema debe funcionar como una herramienta de tamizaje o apoyo, y nunca como un sustituto del criterio clínico profesional. La interpretabilidad del contexto sigue siendo un desafío para la IA.
3. Dependencia de la Calidad de los Datos: La arquitectura del modelo (software) es tan buena como los datos que la alimentan. La mayor limitación encontrada no fue algorítmica, sino la disponibilidad de un dataset balanceado y etiquetado correctamente que represente el lenguaje cotidiano "no depresivo" con todas sus matices.

8. Lista de Abreviaturas

ATSAC: Automated Traffic Surveillance and Control (Sistema Automatizado de Vigilancia y Control de Tráfico)

BERT: Bidirectional Encoder Representations from Transformers

CAV: Connected Autonomous Vehicles (Vehículos Autónomos Conectados)

CoTV: Cooperative Control for Traffic Light Signals and Connected Autonomous Vehicles

ECI: Escuela Colombiana de Ingeniería Julio Garavito

IA: Inteligencia Artificial

LLM: Large Language Model (Gran Modelo de Lenguaje)

LSTM: Long Short-Term Memory (Memoria a Corto y Largo Plazo – Red Neuronal Recurrente)

MDPI: Multidisciplinary Digital Publishing Institute

NLTK: Natural Language Toolkit

PLN: Procesamiento de Lenguaje Natural (en inglés: NLP)

PTIA: Proyecto de Tecnologías de Inteligencia Artificial

RoBERTa: Robustly Optimized BERT Approach

WHO: World Health Organization (Organización Mundial de la Salud)

9. Referencias bibliográficas

- Nadeem, M. (2016). *Identifying depression on Twitter*. arXiv. <https://arxiv.org/abs/1607.07384>
- Bokolo, B. G. (2023). Deep learning-based depression detection from social media. *Electronics*, 12(21), 4396. MDPI. <https://www.mdpi.com/2079-9292/12/21/4396>
- Hassan, N., Li, C., & Tremayne, M. (2017). ClaimBuster: The first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment*, 10(12), 1945–1948. <https://vldb.org/pvldb/vol10/p1945-li.pdf>
- de Beer, D., Matthee, M. C., & Coetzee, M. (2020). Approaches to identify fake news: A systematic literature review. *Sustainability*, 12(22), 9423. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7250114/>
- U.S. Department of Transportation. (2023). *Los Angeles automated traffic surveillance and control (ATSAC) system*. ITS Knowledge Resources. <https://www.itskrs.its.dot.gov/2023-b01770>
- Guo, J., Cheng, L., & Wang, S. (2022). CoTV: Cooperative control for traffic light signals and connected autonomous vehicles using deep reinforcement learning. *arXiv*. <https://arxiv.org/abs/2201.13143>
- World Health Organization. (2025, August 29). *Depressive disorder (depression)*. <https://www.who.int/es/news-room/fact-sheets/detail/>
- Chancellor, S., & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: A critical review. *NPJ Digital Medicine*, 3(1), 43. <https://doi.org/10.1038/s41746-020-0233-7>
- Keaggle . (n.d.). <https://www.kaggle.com/datasets/francescoronzano/spanish-tweets-suggesting-depression>

10. Anexos

Los resultados obtenidos anteriormente se pueden observar mas a detalle en el siguiente entorno de Jupyter Notebook.

- [ECIFeelings](#)