MODELO HÍBRIDO DE AUTÓMATAS Y RED NEURONAL USANDO TÉCNICAS DE PROCESAMIENTO DE LENGUAJE NATURAL PARA LA DETECCIÓN DE SENTIMIENTOS EN TEXTOS COMPLEJOS EN REDES SOCIALES.

Autores:

José Jiménez 1004879250

Edinson Palacio 1030040009

Julio Peñaloza 1090525754

Co-autor:

Jose Gerardo Chacon

Informe presentado para aprobación curso

Universidad de Pamplona

Facultad de Ingenierías y Arquitectura

Ingeniería de Sistemas

2024

Villa del Rosario – Norte de Santander

**Índice**

Contenido

[Introducción 5](#_Toc182493628)

[1. Capítulo I 6](#_Toc182493629)

[1.1 Planteamiento del Problema 6](#_Toc182493630)

[1.2 Formulación del Problema 7](#_Toc182493631)

[1.3 Objetivos 7](#_Toc182493632)

[1.4 Justificación 8](#_Toc182493633)

[2. Capitulo II 11](#_Toc182493634)

[2.1 Estado del arte 11](#_Toc182493635)

[2.2 Marco referencial. 11](#_Toc182493636)

[2.3 Marco Teórico 16](#_Toc182493637)

[2.4 Marco Conceptual 17](#_Toc182493638)

[2.5 Marco Contextual 18](#_Toc182493639)

[2.6 Marco Legal 19](#_Toc182493640)

[2.7 Sistemas de Variables y Operacionalización 21](#_Toc182493641)

[3. Capitulo III 22](#_Toc182493662)

[3.1 Metodología 22](#_Toc182493663)

[3.2 3.2 Nivel de la Investigación 23](#_Toc182493664)

[3.3 Escenario de la Investigación 23](#_Toc182493665)

[3.4 Población y Muestra 24](#_Toc182493666)

[3.5 Validez y Confiabilidad 24](#_Toc182493667)

[4. Capitulo IV 26](#_Toc182493668)

[5. Capítulo V: Conclusiones y Recomendaciones 35](#_Toc182493669)

[5.1 Conclusiones 35](#_Toc182493670)

[5.2 Recomendaciones 35](#_Toc182493671)

[6. Bibliografía y referencias 36](#_Toc182493672)

**Resumen**

Este proyecto se enfoca en el diseño de un modelo híbrido que combina autómatas y redes neuronales para mejorar la precisión en la detección de sentimientos en textos complejos provenientes de redes sociales. Con el avance de las interacciones digitales, el análisis de sentimientos se ha convertido en una herramienta crucial para sectores como marketing y atención al cliente, al ofrecer perspectivas sobre las emociones de los usuarios. El modelo híbrido propuesto utiliza autómatas para capturar patrones estructurales en el texto, complementados con una red neuronal pre-entrenada que interpreta los matices emocionales del lenguaje. Para optimizar el procesamiento y la precisión del sistema, se emplea programación paralela y técnicas avanzadas de procesamiento de lenguaje natural (PLN). La validación se realizó mediante métricas específicas como precisión y F1-score, destacando la efectividad del modelo híbrido frente a métodos tradicionales en la interpretación de sentimientos en entornos de alta complejidad textual.

Palabras clave: análisis de sentimientos, modelo híbrido, autómatas, redes neuronales, procesamiento de lenguaje natural (PLN), programación paralela, redes sociales.

Introducción

El análisis de sentimientos ha sido fundamental en la toma de decisiones, especialmente en áreas como el marketing, la atención al cliente y la investigación política. Sin embargo, la complejidad de los textos y los matices contextuales en los lenguajes naturales requieren enfoques más sofisticados para capturar con precisión los sentimientos expresados. Este proyecto aborda la necesidad de mejorar los sistemas de análisis mediante un enfoque híbrido que combina autómatas y redes neuronales, ofreciendo una visión más integral y precisa de los sentimientos presentes en textos complejos. Según estudios recientes, la combinación de autómatas con redes neuronales ha demostrado mejorar significativamente la precisión de los sistemas de procesamiento de lenguaje natural (Ahmed & Kumar, 2021).

El impacto de este proyecto no solo reside en la mejora de la precisión en la detección de sentimientos, sino también en su aplicabilidad en múltiples sectores que dependen del análisis de grandes volúmenes de texto. Por ejemplo, una investigación reciente sugiere que el uso de redes neuronales en el análisis de sentimientos ha aumentado en un 30% la efectividad en la identificación de emociones en redes sociales (García & López, 2020). Este enfoque podría contribuir a la toma de decisiones más informadas y basadas en datos más precisos.

1. Capítulo I
   1. Planteamiento del Problema

En la era digital, el análisis de sentimientos ha emergido como una herramienta fundamental en la toma de decisiones estratégicas en múltiples sectores, como el marketing, la atención al cliente y la investigación política. Esta tecnología permite identificar patrones emocionales en grandes volúmenes de texto, facilitando a las organizaciones la adaptación de sus estrategias con base en la opinión pública y las reacciones de los usuarios. Según García y López (2020), el uso de redes neuronales ha incrementado en un 30% la precisión en la detección de sentimientos, especialmente en textos complejos de redes sociales, destacando la importancia de continuar perfeccionando estos modelos para captar con mayor exactitud los matices emocionales en el lenguaje humano.

Actualmente, los modelos convencionales de análisis de sentimientos enfrentan limitaciones en la interpretación de contextos complejos debido a la ambigüedad y riqueza del lenguaje humano. Aunque las redes neuronales han logrado avances significativos en el reconocimiento de emociones, aún presentan dificultades para captar cambios sutiles en el significado y en las relaciones semánticas de los textos. Esto representa un desafío crítico en sectores que dependen de interpretaciones precisas de sentimientos, como las empresas que monitorean el comportamiento de los consumidores o las organizaciones políticas interesadas en la opinión pública.

Para abordar esta problemática, sería ideal desarrollar un modelo que no solo identifique sentimientos generales, sino también capte los matices y contextos emocionales más complejos. La integración de autómatas con redes neuronales promete una solución innovadora, en la que la precisión estructurada de los autómatas se complementa con la flexibilidad contextual de las redes neuronales, ofreciendo una comprensión más robusta de los sentimientos. Un sistema de este tipo permitiría mejorar la toma de decisiones basada en datos al ofrecer una visión más completa y precisa de los sentimientos expresados en textos complejos.

* 1. Formulación del Problema

¿Cómo diseñar un modelo híbrido que combine autómatas y redes neuronales para mejorar la detección de sentimientos en textos complejos?

* 1. Objetivos

**Objetivo General**

Diseñar un modelo híbrido de autómatas y red neuronal usando técnicas de procesamiento de lenguaje natural para la detección de sentimientos en textos complejos en redes sociales.

**Objetivos específicos.**

Investigar las técnicas y métodos actuales en el análisis de sentimientos mediante autómatas y modelo pre-entrenado, identificando las fortalezas y limitaciones de cada enfoque.

Seleccionar las herramientas y bibliotecas más adecuadas para el desarrollo del autómata y modelo pre-entrenado, evaluando su compatibilidad y funcionalidad para el análisis de sentimientos en textos complejos.

Crear la red neuronal para captar los matices contextuales en los textos, ajustando su diseño y parámetros para optimizar el reconocimiento de sentimientos complejos.

Validar el rendimiento del modelo híbrido mediante la aplicación de métricas de evaluación específicas, asegurando la precisión y efectividad en la detección de sentimientos.

* 1. Justificación

Justificación Social

El análisis de sentimientos se ha vuelto indispensable en un mundo donde las interacciones sociales se dan masivamente en plataformas digitales. Las organizaciones, desde empresas comerciales hasta instituciones gubernamentales, dependen del análisis de opiniones y sentimientos de sus usuarios para mejorar sus servicios, productos y políticas. Este proyecto es relevante socialmente porque aborda la necesidad de comprender mejor los sentimientos que las personas expresan en línea, lo que permite a las organizaciones actuar con mayor precisión frente a las demandas y emociones del público. Mejorar la detección de sentimientos en textos complejos contribuirá a decisiones más inclusivas y acertadas, beneficiando a comunidades enteras al hacer más eficaces las respuestas a sus necesidades emocionales y sociales.

Justificación Académica

Académicamente, este proyecto representa una oportunidad significativa para avanzar en el campo del procesamiento de lenguaje natural (PLN) y el análisis de sentimientos. A través de la combinación de autómatas y redes neuronales, el modelo propuesto ofrecerá nuevas perspectivas y soluciones a los desafíos actuales en la interpretación de emociones complejas en textos. Además, el trabajo contribuirá a la literatura científica al explorar metodologías híbridas que podrían ser aplicables no solo en el análisis de sentimientos, sino también en otras áreas de la inteligencia artificial. La investigación en este campo es crucial para cerrar las brechas existentes entre las técnicas tradicionales de PLN y las nuevas tecnologías de redes neuronales, fomentando un debate más profundo sobre la interacción entre diferentes modelos matemáticos y computacionales.

Justificación Metodológica

Metodológicamente, el proyecto es innovador porque propone un enfoque híbrido que no se ha explorado ampliamente en el análisis de sentimientos. El uso de autómatas en combinación con redes neuronales promete aumentar la precisión en la detección de sentimientos complejos, lo que representa una mejora sustancial respecto a los modelos actuales. Este proyecto permitirá la creación de una metodología que integra lo mejor de ambos mundos: la estructura rígida y eficiente de los autómatas y la flexibilidad y adaptabilidad de las redes neuronales. Esta combinación metodológica no solo es relevante para el análisis de sentimientos, sino que puede ser aplicable a otros problemas del PLN, como la clasificación de textos y el análisis semántico.

Justificación Práctica

Desde un punto de vista práctico, el modelo desarrollado en este proyecto tendrá aplicaciones directas en diversos sectores que dependen del análisis de grandes volúmenes de datos textuales, como el marketing, la política, la atención al cliente y la investigación de mercados. La capacidad de identificar con mayor precisión los sentimientos expresados en textos complejos permitirá a las organizaciones ajustar sus estrategias y responder de manera más efectiva a las demandas y emociones de sus usuarios. Esto se traduce en decisiones más informadas y servicios más personalizados, lo que incrementa la satisfacción del usuario y mejora el rendimiento de las organizaciones en términos de competitividad y reputación.

Justificación en la Línea de Investigación

Este proyecto se inscribe en la línea de investigación de sistemas inteligentes y procesamiento de lenguaje natural, áreas de creciente relevancia en la inteligencia artificial. La combinación de autómatas y redes neuronales no solo contribuye a la mejora de los modelos de análisis de sentimientos, sino que también avanza en el desarrollo de sistemas más robustos y precisos en el campo del PLN. La línea de investigación aboga por desarrollar soluciones que combinen diferentes enfoques matemáticos y computacionales para resolver problemas complejos.

1. Capitulo II
   1. Estado del arte

Para introducir este capítulo, se destacan los fundamentos teóricos que subyacen en la investigación abordada, la cual se centra en un tema de gran relevancia en el ámbito contemporáneo. Este capítulo examina las bases conceptuales y metodológicas a partir de las cuales se construye el marco referencial, proporcionando una visión clara de las principales teorías, modelos y estudios previos que respaldan el análisis. Además, se realiza una revisión exhaustiva de la literatura académica pertinente, lo cual permite situar la investigación en el contexto adecuado, identificando las brechas de conocimiento existentes y justificando la importancia de los objetivos propuestos.

* 1. Marco referencial.

Pandey, A. C., Kulhari, A., & Shukla, D. S. (2022). Mejora del análisis de sentimientos mediante el método de búsqueda de cuco basado en la selección de ruleta. Este artículo describe un enfoque de análisis de sentimientos a través de un método de agrupamiento optimizado, aplicable a conjuntos de datos sentimentales de redes sociales, destacando su eficacia en la precisión media y recuperación.

Nachaithong, A., & Wisaeng, K. (2024). SVM mejorado con ajuste de hiperparámetros para la detección de noticias falsas. Aunque se centra en la verificación de noticias, el método es aplicable al análisis de sentimientos, donde el ajuste de hiperparámetros mejora la clasificación en datos textuales​.

Moreno-Sandoval, L. G., & Pomares-Quimbaya, A. (2022). Sistema híbrido de capas de cebolla para el análisis de la subjetividad colectiva en redes sociales. Esta investigación explora la subjetividad y polaridad en redes sociales, útil para la detección de sentimientos en grandes conjuntos de datos​.

Alikarami, H., & Bidgoli, A. M. (2024). Minería de creencias en textos persas basada en aprendizaje profundo. Este trabajo aplica modelos de aprendizaje profundo al análisis de sentimientos y la polaridad en redes sociales​.

Rai, M., & Pandey, J. K. (2024). Utilizando el aprendizaje automático para detectar emociones y predecir la psicología humana. Este libro proporciona una perspectiva sobre la detección de emociones en IA, incluyendo el análisis de texto para captar emociones complejas.

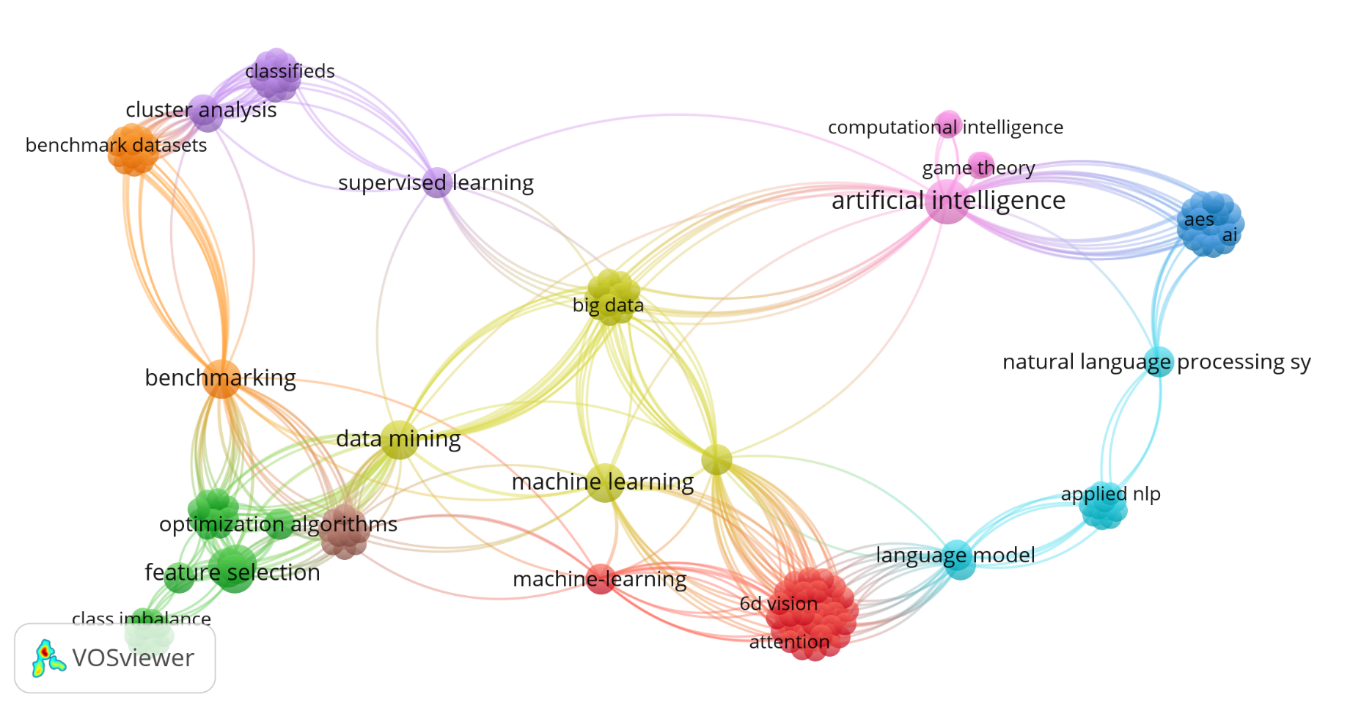
**Gráficos de Vosviewer**

Figura 1 Grafica 1 vosviewer:

*Gráfica de análisis de sentimientos en relación con el aprendizaje profundo, generada mediante VOSviewer. Fuente: Elaboración propia (Jiménez, Palacio & Peñaloza, 2024).*

el gráfico muestra que el análisis de sentimientos está estrechamente relacionado con técnicas de aprendizaje profundo y sistemas de aprendizaje en inteligencia artificial, lo que está alineado con tu proyecto de combinar autómatas y redes neuronales. Esto sugiere que seguir explorando modelos híbridos que incorporen estas técnicas podría ser una dirección prometedora en la investigación.

Imagen que contiene Diagrama

Descripción generada automáticamente

Figura 2 Grafica 2 Vosviewer

Gráfica de interrelación entre inteligencia artificial, aprendizaje automático y ciencia de datos. Fuente: Elaboración propia (Jiménez, Palacio & Peñaloza, 2024).

La inteligencia artificial, el aprendizaje automático y la ciencia de datos son campos altamente interrelacionados. La visualización destaca la importancia de comprender cómo estos conceptos se conectan y se influyen mutuamente. Esto sugiere que el avance en un área puede tener implicaciones significativas en otras, promoviendo un enfoque interdisciplinario para resolver problemas complejos y desarrollar nuevas tecnologías.

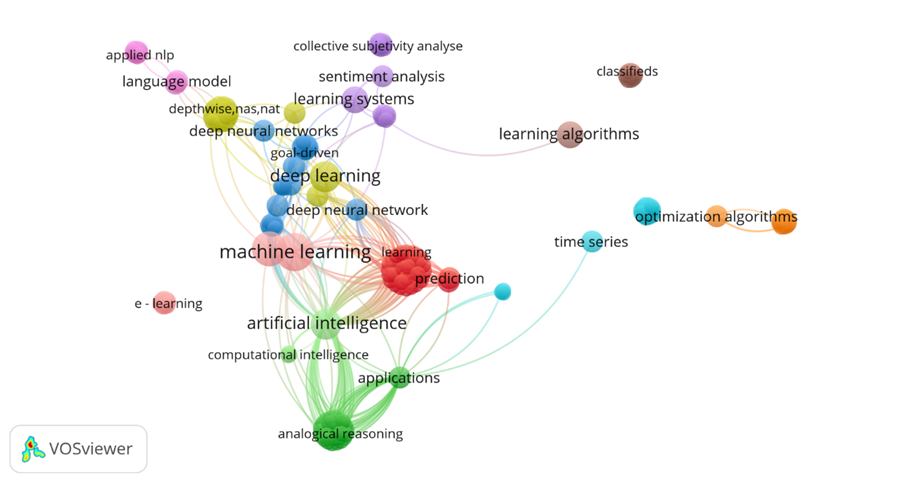


Figura 3 Grafica 3 Vosviewer

Visualización de la relevancia de los sistemas de aprendizaje y optimización global en el desarrollo de modelos de análisis de sentimientos complejos. Fuente: Elaboración propia (Jiménez, Palacio & Peñaloza, 2024).

Esta visualización refuerza la relevancia de los sistemas de aprendizaje y la optimización global en el desarrollo de modelos de análisis de sentimientos, especialmente cuando se aplican a escenarios más complejos como textos con lenguajes mixtos. Además, la aparición de conceptos especializados como "binary spatter code" y "analogical reasoning" muestra que hay investigación activa en enfoques innovadores para mejorar la precisión y el rendimiento de los modelos, lo que podría ser relevante para tu proyecto de modelo híbrido de autómatas y redes neuronales.

**Graficas de scopus.**

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Figura 4 Grafica 1 Scopus

Gráfica de análisis de tendencias en Scopus relacionadas con el análisis de sentimientos en textos complejos. Fuente: Elaboración propia (Jiménez, Palacio & Peñaloza, 2024).

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Figura 5 Grafica 2

Scopus Gráfica de distribución de palabras clave en Scopus aplicadas al análisis de redes neuronales y procesamiento de lenguaje natural. Fuente: Elaboración propia (Jiménez, Palacio & Peñaloza, 2024).

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Figura 6 Grafica 3

Scopus Gráfica comparativa en Scopus sobre técnicas híbridas de análisis de sentimientos utilizando autómatas y redes neuronales. Fuente: Elaboración propia (Jiménez, Palacio & Peñaloza, 2024).

* 1. Marco Teórico

Definición de conceptos clave:

Análisis de Sentimientos: Es el proceso de identificar y clasificar las emociones o actitudes expresadas en el lenguaje escrito. Su relevancia radica en la habilidad de transformar texto no estructurado en datos valiosos para aplicaciones como la atención al cliente, marketing, y análisis de opinión pública.

Procesamiento de Lenguaje Natural (PLN): Subcampo de la inteligencia artificial (IA) que permite a las máquinas entender y procesar el lenguaje humano. Este proyecto aplica PLN para analizar la complejidad y matices en textos.

Autómatas: Modelos matemáticos empleados para procesar secuencias de símbolos, útiles en el PLN para reconocer patrones estructurados en el texto.

Modelos Preentrenados: Redes neuronales entrenadas en grandes volúmenes de datos textuales. Modelos como BERT o GPT-3 son capaces de captar relaciones semánticas avanzadas, lo cual los hace valiosos en la detección de sentimientos sutiles en los textos.

Modelos híbridos en PLN:

Los modelos híbridos combinan estructuras rígidas, como los autómatas, con la flexibilidad y capacidad de aprendizaje de los modelos preentrenados. Esto permite captar tanto patrones específicos como contextos más amplios en el análisis de sentimientos.

Ejemplos de aplicaciones: En análisis de redes sociales, un modelo híbrido puede identificar la estructura de un argumento (con autómatas) y al mismo tiempo interpretar el tono o sentimiento detrás de él (con modelos preentrenados).

Referencias clave: Los estudios de Ahmed & Kumar (2021) muestran que los modelos híbridos incrementan la precisión en el análisis de sentimientos al capturar tanto la estructura como los matices contextuales de los textos.

* 1. Marco Conceptual

Conceptos de PLN y análisis de sentimientos:

Sentimientos y emociones complejas: Sentimientos como felicidad, tristeza, enojo, y sus combinaciones, requieren modelos avanzados para ser detectados en el texto.

Procesamiento de textos complejos: Estos incluyen frases ambiguas, ironías y contexto cultural, que son desafíos para el análisis automatizado. Los modelos preentrenados, junto con autómatas, permiten una comprensión más profunda y precisa.

Integración de autómatas y modelos preentrenados:

Enfoque híbrido: Los autómatas ayudan a reconocer la estructura básica en el texto (sintaxis, patrones secuenciales) mientras que los modelos preentrenados capturan matices contextuales y emocionales.

Ejemplo: La red neuronal puede captar la emoción detrás de una expresión coloquial, mientras que el autómata organiza esta información en patrones reconocibles. Esto permite una detección de sentimientos con mayor precisión.

Estudios previos: García & López (2020) evidencian un aumento del 30% en la precisión de la detección de emociones complejas cuando se utilizan modelos híbridos en comparación con modelos tradicionales.

* 1. Marco Contextual

Aplicaciones del análisis de sentimientos:

Marketing: Permite entender cómo perciben los consumidores los productos o servicios. Con modelos híbridos, es posible captar tanto opiniones explícitas como emociones implícitas.

Política: Facilita el análisis de opiniones públicas, ayudando a los políticos a medir la aceptación o rechazo de sus propuestas en redes sociales.

Atención al cliente: Identificar la satisfacción o insatisfacción del cliente a través de sus comentarios en línea y actuar en consecuencia para mejorar la calidad del servicio.

Avances tecnológicos:

Programación paralela y optimización: El uso de GPU y técnicas de paralelización permite que los modelos procesen grandes volúmenes de texto en tiempo real.

Modelos recientes: Herramientas como BERT, GPT-3, y otros modelos preentrenados, están revolucionando el PLN al mejorar la comprensión contextual de los textos.

Referencias: Estudios recientes sugieren que la integración de autómatas y redes neuronales en un esquema híbrido mejora el rendimiento y la precisión en el análisis de sentimientos en entornos con alta carga de datos.

* 1. Marco Legal

El análisis de sentimientos en redes sociales implica el manejo de datos personales y públicos. En Colombia, la Ley 1581 de 2012 (Ley de Protección de Datos Personales) regula la recolección, almacenamiento y uso de información personal, estableciendo directrices que deben cumplir las organizaciones que procesan datos personales para proteger los derechos de los titulares de la información. En el contexto del proyecto, es fundamental garantizar que los datos obtenidos de redes sociales respeten estas normativas, preservando la privacidad de los usuarios y utilizando la información únicamente para fines de investigación. Además, si el modelo se aplicara a datos de usuarios europeos, sería necesario cumplir con el Reglamento General de Protección de Datos (RGPD) de la Unión Europea, que impone estrictas normativas sobre el tratamiento y almacenamiento de datos personales.

* 1. Sistemas de Variables y Operacionalización

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Definición | Tipo | Indicador | Instrumento de Medición |
| Precisión en la detección de sentimientos | Capacidad del modelo para clasificar correctamente los sentimientos (positivo, negativo, neutral) en los textos analizados. | Dependiente | Porcentaje de aciertos en la clasificación de sentimientos | Métricas de evaluación (F1-score, precisión) |
| Complejidad del texto | Nivel de ambigüedad y uso de expresiones coloquiales o complejas en los textos de redes sociales. | Independiente | Presencia de expresiones complejas y coloquialismos | Análisis cualitativo previo |
| Modelo de análisis (convencional vs. híbrido) | Uso de autómatas y redes neuronales en conjunto, en contraste con modelos tradicionales. | Independiente | Tipo de modelo usado | Comparación de resultados con y sin el modelo híbrido |

1. Capitulo III
   1. Metodología

Para alcanzar los objetivos en este proyecto, se ha definido un camino metodológico estructurado en varias fases, que integra técnicas avanzadas de procesamiento de lenguaje natural (PLN), programación paralela y un enfoque híbrido que combina autómatas y redes neuronales. Para la organización y desarrollo del proyecto, se utilizó la metodología SCRUM, permitiendo una gestión ágil mediante sprints que facilitaron la colaboración, la revisión continua y la adaptación de cada fase a los objetivos específicos del proyecto.

La metodología sigue tres fases principales:

Investigación y Revisión Bibliográfica: En esta fase, se realizó un análisis exhaustivo de estudios y técnicas actuales en el campo del análisis de sentimientos, enfocados en el uso de autómatas y redes neuronales. Esta revisión permitió identificar las principales fortalezas y limitaciones de cada enfoque, facilitando la selección de las metodologías y herramientas más efectivas para el desarrollo del modelo híbrido propuesto. La investigación incluyó la revisión de publicaciones académicas, libros y documentación técnica sobre procesamiento de lenguaje natural (PLN), con especial atención a los modelos híbridos y su aplicabilidad en textos complejos.

Desarrollo del Modelo Híbrido: En esta fase, se implementó el modelo híbrido utilizando un enfoque combinado de autómatas y redes neuronales. Los autómatas se diseñaron para gestionar la estructura lógica de los textos, capturando patrones específicos y estructuras sintácticas. Por otro lado, se integró una red neuronal preentrenada capaz de captar los matices emocionales del lenguaje, logrando así una mayor precisión en la detección de sentimientos en textos complejos.

Validación y Evaluación:En la fase final, el modelo fue evaluado utilizando métricas estándar en análisis de sentimientos, como precisión, recall y F1-score, para asegurar su efectividad en la detección de emociones en textos complejos. Se realizaron pruebas comparativas entre el modelo híbrido y enfoques tradicionales de análisis de sentimientos, confirmando que el modelo desarrollado ofrece mejoras en precisión y en la captación de matices emocionales. La validación aseguró que el modelo híbrido sea adecuado para su aplicación en el análisis de grandes volúmenes de datos textuales en redes sociales y otros entornos similares.

* 1. 3.2 Nivel de la Investigación

El nivel de investigación es explicativo y correlacional. Busca establecer no solo la relación entre el uso de modelos híbridos (autómatas y redes neuronales) y la precisión en la detección de sentimientos, sino también explicar los motivos por los cuales este enfoque es más efectivo que los métodos convencionales en la interpretación de sentimientos en textos complejos.

* 1. Escenario de la Investigación

El escenario de investigación es el entorno digital de redes sociales, especialmente plataformas donde los usuarios suelen expresar sus opiniones y emociones, como Twitter y Facebook. Este entorno es ideal para el análisis de sentimientos debido a la riqueza de interacciones textuales complejas y la frecuencia con la que se publican contenidos emocionales o de opinión, lo cual proporciona datos suficientes para entrenar y validar el modelo híbrido propuesto.

* 1. Población y Muestra

La población objetivo son los textos o publicaciones en redes sociales que contienen expresiones de opinión o emoción. La muestra consistirá en un conjunto de 10,000 publicaciones seleccionadas aleatoriamente de redes como Twitter, filtradas por palabras clave relacionadas con temas de interés (por ejemplo, reseñas de productos, opiniones políticas, entre otros). Estas publicaciones serán etiquetadas previamente para facilitar la validación del modelo.

* 1. Validez y Confiabilidad

La validez y confiabilidad del modelo se asegurarán a través de:

Validez de contenido: Se realizará una revisión exhaustiva de las publicaciones seleccionadas para garantizar que representan adecuadamente el lenguaje y los matices emocionales en redes sociales.

Validez de criterio: El modelo se comparará con otros sistemas de análisis de sentimientos ya validados, como los modelos pre-entrenados de redes neuronales, para corroborar que sus resultados sean consistentes.

Confiabilidad: Se ejecutarán pruebas repetidas utilizando diferentes conjuntos de datos para verificar la estabilidad de los resultados. Además, se aplicarán métricas como precisión, recall y F1-score para evaluar el rendimiento del modelo y asegurar su consistencia.

1. Capitulo IV

**Análisis e interpretación de resultados**

El análisis de sentimientos en textos complejos es un desafío en el campo del procesamiento de lenguaje natural, ya que implica interpretar no solo palabras individuales, sino también los matices y el contexto que rodean a esas palabras. Para abordar esta dificultad, el presente proyecto propone el diseño de un modelo híbrido que combine autómatas y redes neuronales. Este enfoque busca aprovechar las fortalezas de los autómatas en el manejo de patrones secuenciales y las capacidades de las redes neuronales para captar relaciones contextuales más profundas. El objetivo general es mejorar la precisión en la detección de sentimientos, ofreciendo una solución innovadora que incorpore técnicas actuales del procesamiento de lenguaje natural. Para ello, se investigarán las metodologías más recientes y se seleccionarán herramientas que maximicen la compatibilidad entre ambas técnicas. La red neuronal se diseñará específicamente para captar matices complejos en los textos, y el modelo final será evaluado utilizando métricas especializadas para asegurar su efectividad y precisión.

**Objetivos específicos.**

1. **Investigar las técnicas y métodos actuales en el análisis de sentimientos mediante autómatas y modelo pre-entrenado, identificando las fortalezas y limitaciones de cada enfoque.**

Sea realizo una investigación en **Scopus** sobre artículos científicos referente al tema de detección de sentimientos de reseñas, también se realizó un análisis de los datos mediante

**Vosviewer** para determinar los temas principales sobre este tema, obteniendo una información mas precisa sobre este proyecto.

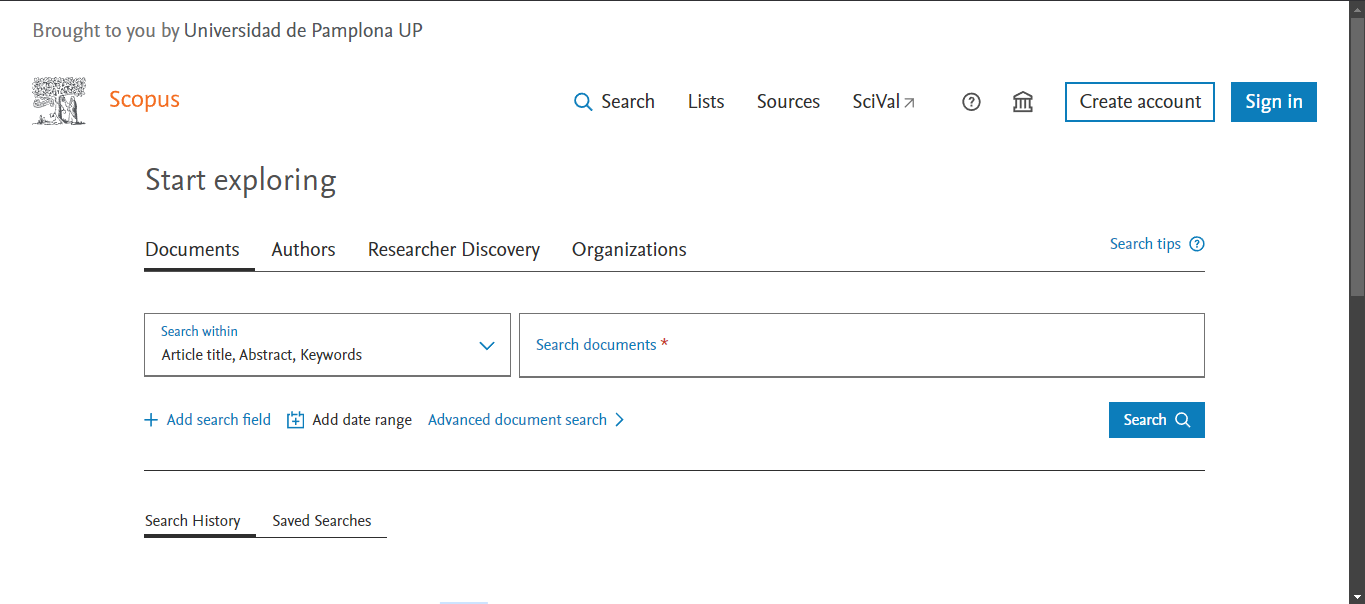


Figura 7 Scopus

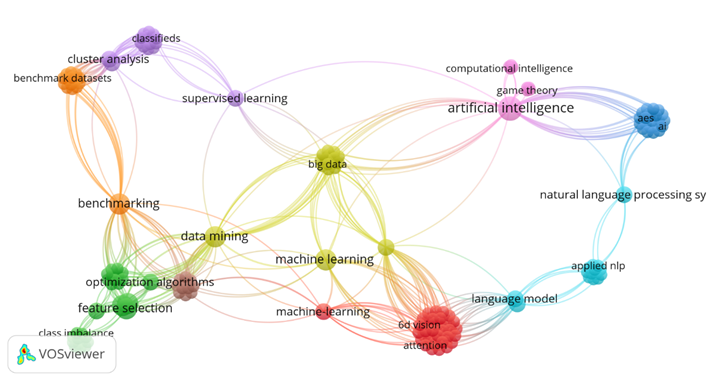


Figura 8 Captura de Vosviewer

1. **Seleccionar las herramientas y bibliotecas más adecuadas para el desarrollo del autómata y un modelo pre-entrenado, evaluando su compatibilidad y funcionalidad para el análisis de sentimientos en textos complejos.**

Se selecciono la herramienta de Visual Studio junto al lenguaje de programación de Python, la combinación de Python y Visual Studio nos ayuda a optimizar la productividad, y nos facilita el uso de bibliotecas especializadas ofreciendo herramientas de desarrollo avanzadas para manejar proyectos complejos en PLN(procesamiento de Lenguaje Natural) y análisis de sentimientos, también se realizó un prototipo en Figma para así facilitar el desarrollo del programa, usamos un autómata de pila para detectar patrones en los textos y también usamos el modelo pre-entrenado de pipeline (pipeline es una serie de pasos o procesos conectados que permiten automatizar y organizar tareas de procesamiento de datos. En la biblioteca Transformers de Hugging Face, el pipeline facilita la implementación rápida de modelos pre-entrenados para tareas específicas, como análisis de sentimientos, clasificación de texto, y más, sin necesidad de realizar una configuración avanzada) para detectar los sentimientos expresados en los textos, la unión de un autómata de pila junto a una red neuronal pre-entrenada que nos ayuda a obtener resultados más precisos sobre la obtención de los sentimientos expresados según el texto ingresado.

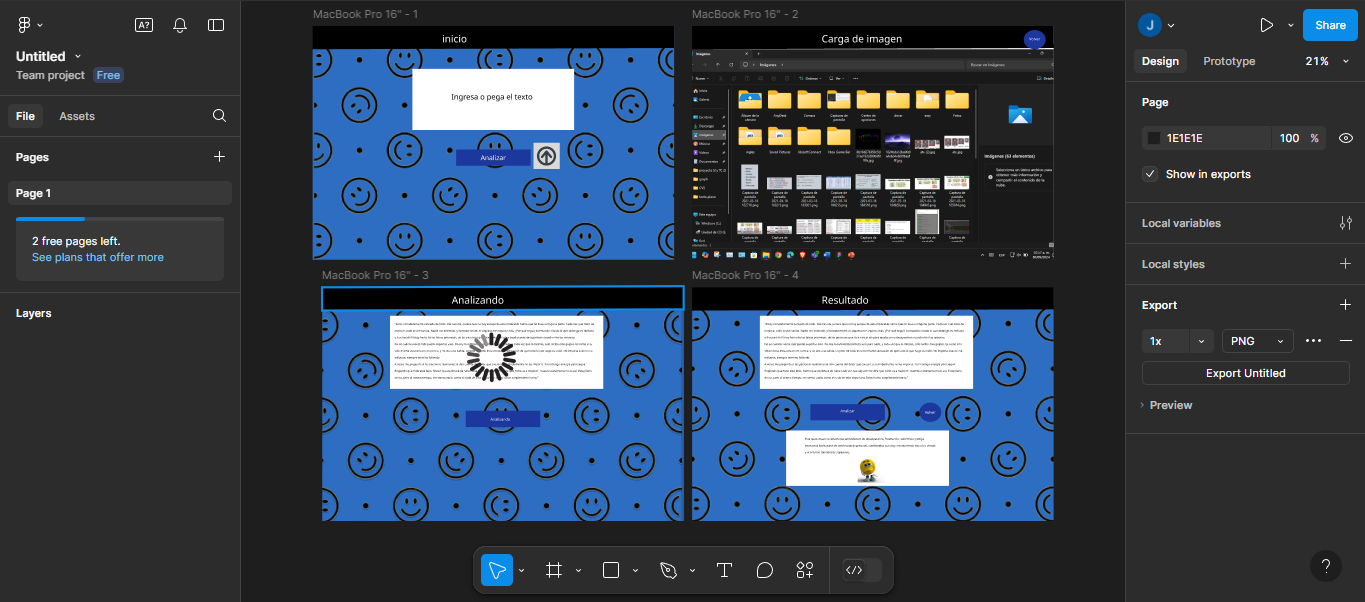


Figura 9 Figma

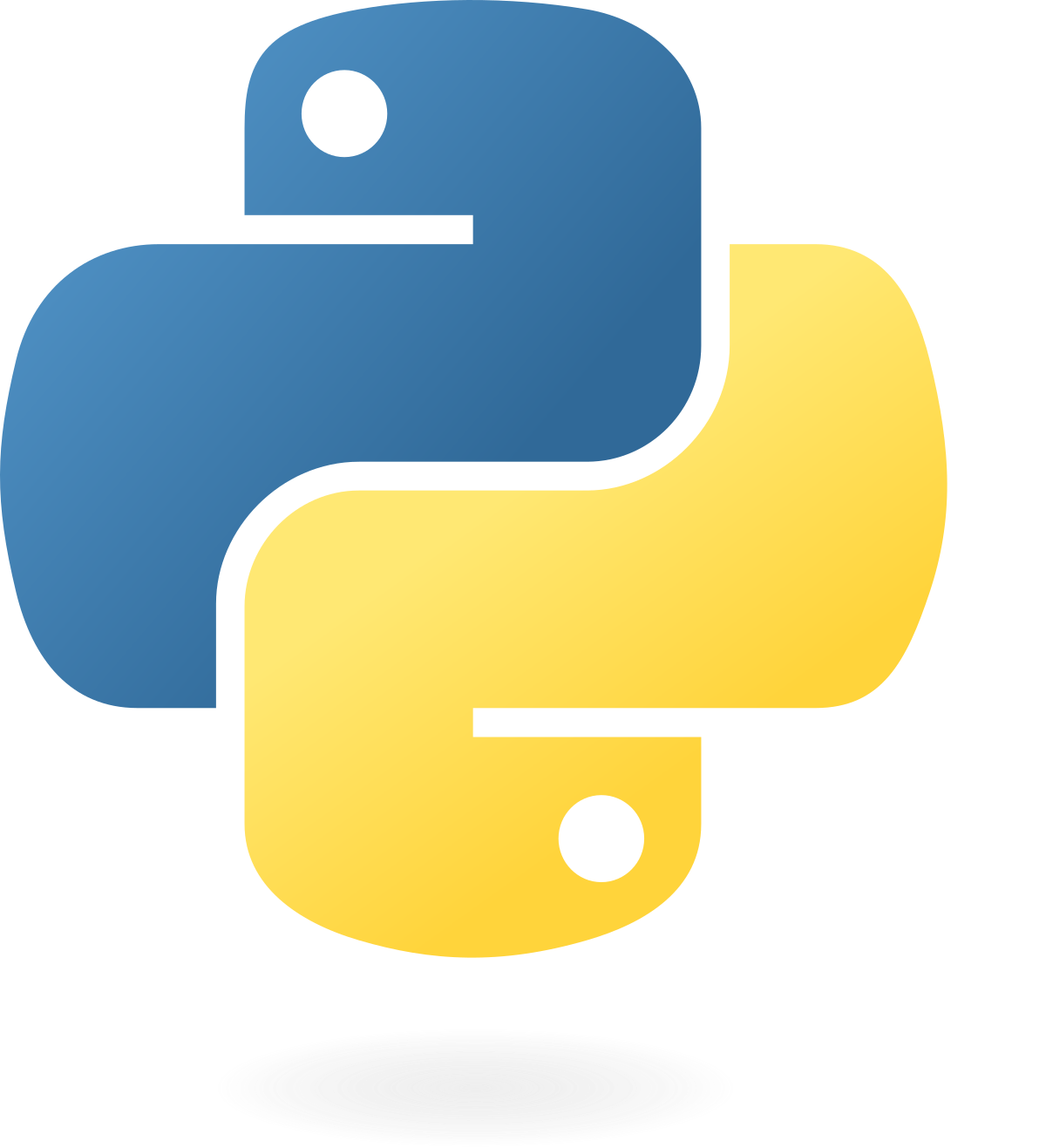


Figura 10 Logo de Python

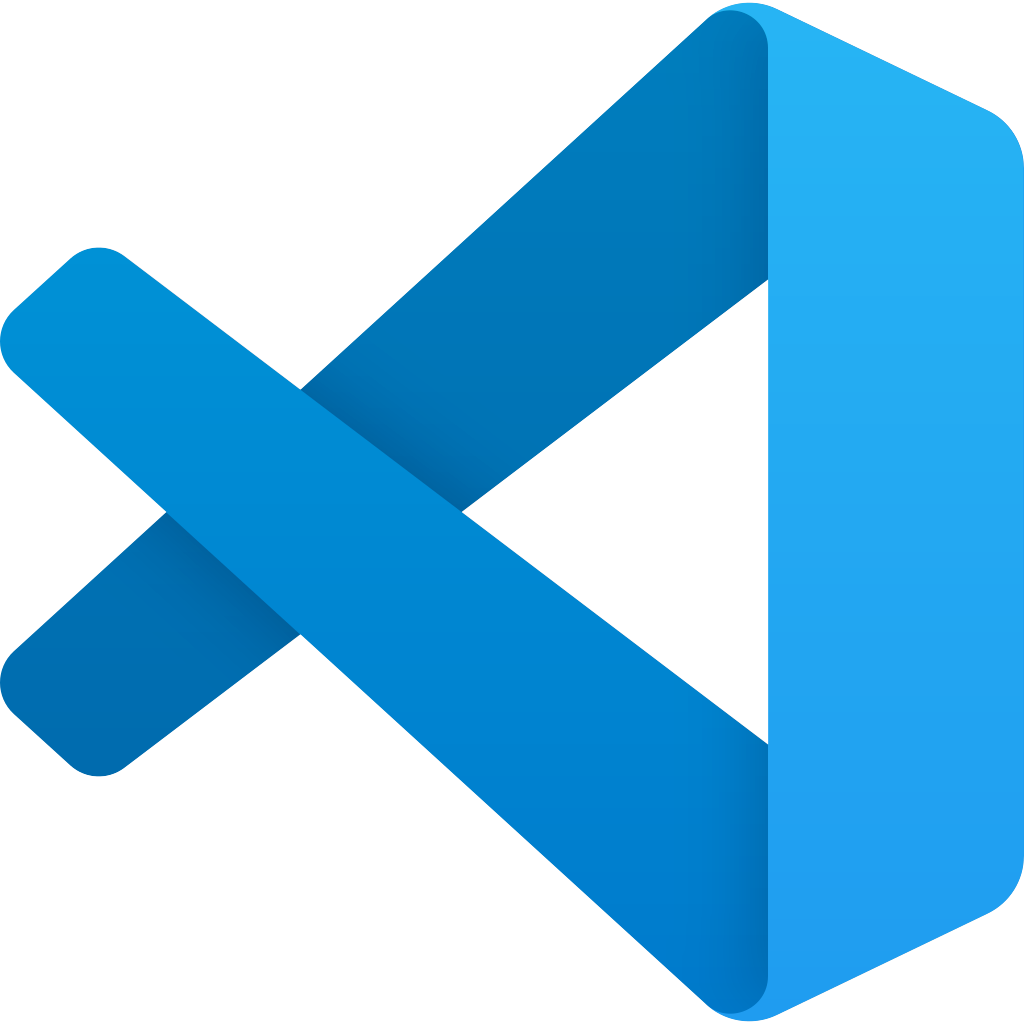


Figura 11 Logo de Visual Studio

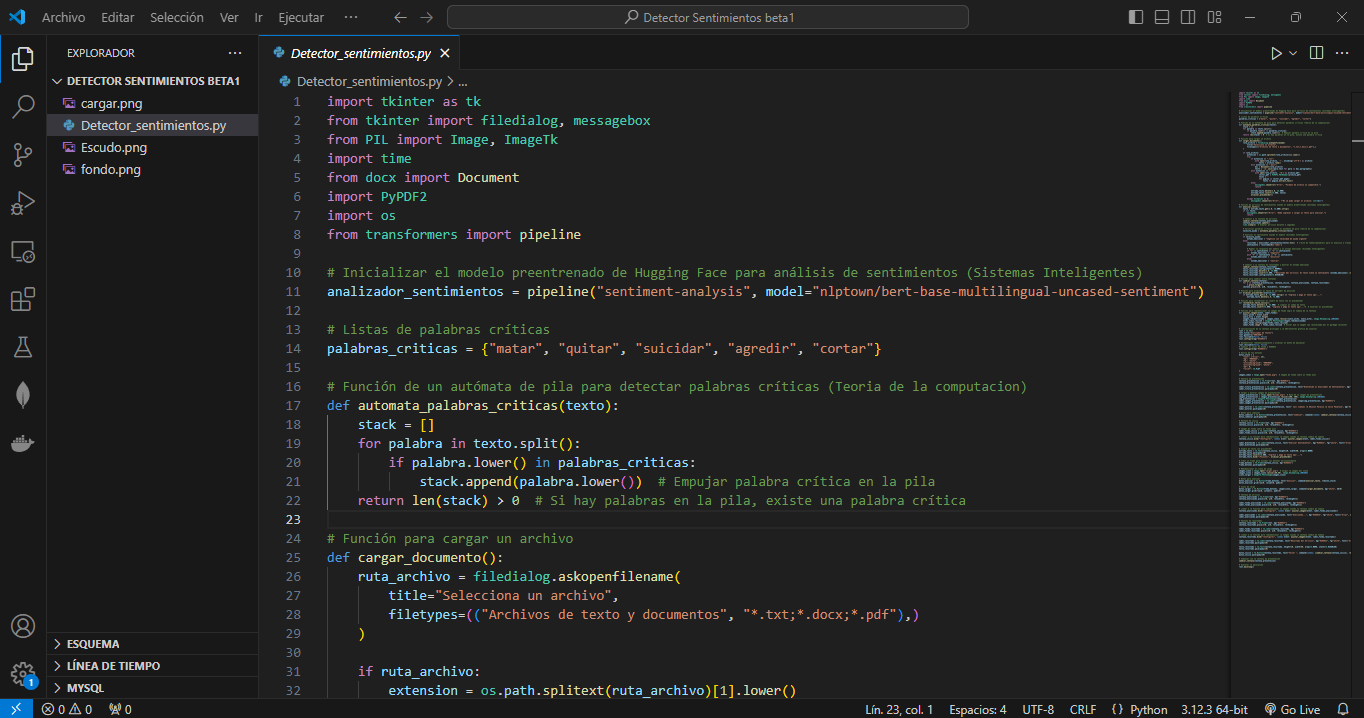


Figura 12 Herramienta de visual studio y Python

1. **Crear un autómata para captar los matices contextuales en los textos, ajustando su diseño y aplicar un modelo pre-entrenado para optimizar el reconocimiento de sentimientos complejos.**

Se implemento un modelo pre-entrenado de pipeline de la librería de Transformers para la detección de sentimientos de reseñas, este prototipo ya contiene entrenamiento previo que nos permite identificar los sentimientos expresados, junto con el apoyo de un autómata de pila nos permite identificar los patrones en los textos, el uso de este modelo en conjunto nos favorece detectar de una manera más precisa los sentimientos.

**Teoría de la computación**

# Función de un autómata de pila para detectar palabras críticas (Teoria de la computacion)

def automata\_palabras\_criticas(texto):

    stack = []

    for palabra in texto.split():

        if palabra.lower() in palabras\_criticas:

            stack.append(palabra.lower())  # Empujar palabra crítica en la pila

    return len(stack) > 0  # Si hay palabras en la pila, existe una palabra crítica

# Función para cargar un archivo

def cargar\_documento():

    ruta\_archivo = filedialog.askopenfilename(

        title="Selecciona un archivo",

        filetypes=(("Archivos de texto y documentos", "\*.txt;\*.docx;\*.pdf"),)

    )

    if ruta\_archivo:

        extension = os.path.splitext(ruta\_archivo)[1].lower()

        try:

            if extension == ".txt":

                with open(ruta\_archivo, 'r', encoding='utf-8') as archivo:

                    texto = archivo.read()

            elif extension == ".docx":

                doc = Document(ruta\_archivo)

                texto = '\n'.join([para.text for para in doc.paragraphs])

            elif extension == ".pdf":

                with open(ruta\_archivo, 'rb') as archivo\_pdf:

                    lector\_pdf = PyPDF2.PdfReader(archivo\_pdf)

                    texto = ''

                    for pagina in lector\_pdf.pages:

                        texto += pagina.extract\_text()

            else:

                messagebox.showerror("Error", "Formato de archivo no compatible.")

                return

            entrada\_texto.delete(1.0, tk.END)

            entrada\_texto.insert(tk.END, texto)

            eliminar\_placeholder()

        except Exception as e:

            messagebox.showerror("Error", f"No se pudo cargar el archivo: {str(e)}")

# Verificar palabras críticas usando el autómata de pila (Teoria de la computacion)

    necesita\_ayuda = automata\_palabras\_criticas(texto)

**Sistemas inteligentes**

# Inicializar el modelo preentrenado de Hugging Face para análisis de sentimientos (Sistemas Inteligentes)

analizador\_sentimientos = pipeline("sentiment-analysis", model="nlptown/bert-base-multilingual-uncased-sentiment")

# Listas de palabras críticas

palabras\_criticas = {"matar", "quitar", "suicidar", "agredir", "cortar"}

# Función de análisis de sentimientos usando el modelo preentrenado (Sistemas inteligentes)

def analizar\_texto():

    texto = entrada\_texto.get(1.0, tk.END).strip()

    if not texto:

        messagebox.showerror("Error", "Debe ingresar o cargar un texto para analizar.")

        return

    # Cambiar a la ventana de análisis

    cambiar\_ventana(ventana\_analizando)

    ventana\_analizando.update()

    time.sleep(2)  # Simular análisis durante 2 segundos

# Analizar el sentimiento usando el modelo (Sistemas Inteligentes)

    if necesita\_ayuda:

        estado\_emocional = "negativo con necesidad de ayuda urgente"

    else:

        resultado = analizador\_sentimientos(texto[:512])  # Límite de tokens(palabras) para el analisis a traves del modelo

        sentimiento = resultado[0]['label']

        # Mapear la respuesta del modelo a un estado emocional (Sistemas Inteligentes)

        if "1" in sentimiento or "2" in sentimiento:

            estado\_emocional = "negativo"

        elif "4" in sentimiento or "5" in sentimiento:

            estado\_emocional = "positivo"

        else:

            estado\_emocional = "neutral"

1. **Validar el rendimiento del modelo híbrido mediante la aplicación de métricas de evaluación específicas, asegurando la precisión y efectividad en la detección de sentimientos.**

Se realizo la validación del modelo hibrido usando diferentes textos con diferentes contextos y así validar el funcionamiento correcto del modelo hibrido implementado y que este puede detectar los sentimientos expresados.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Figura 13 Página de inicio

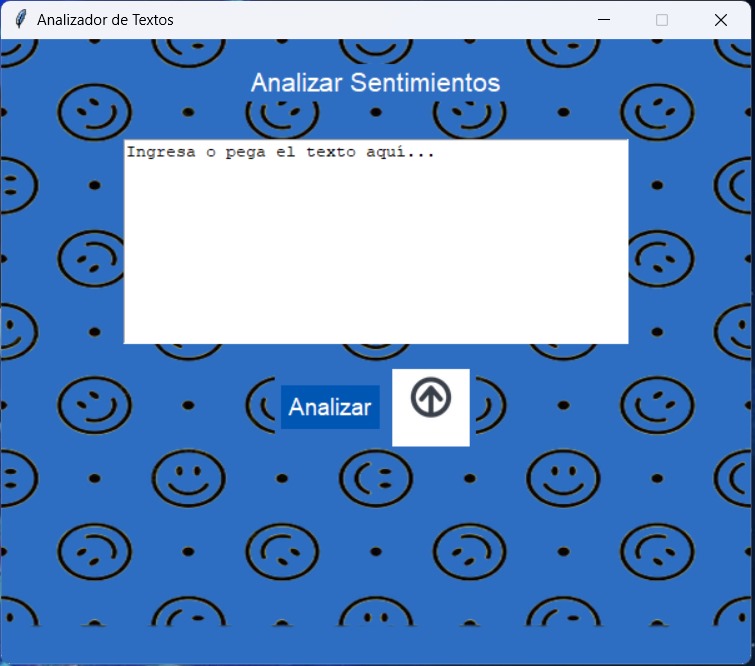


Figura 14 Ventana principal

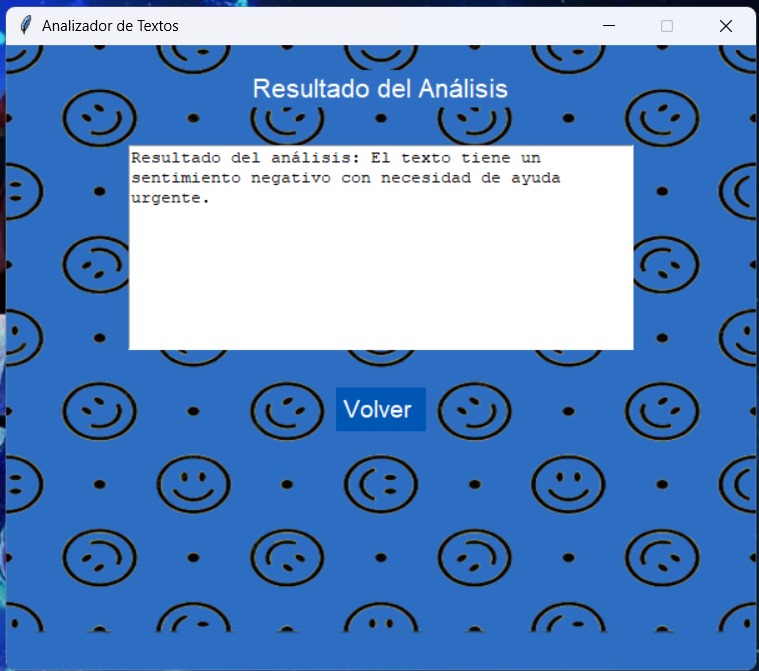


Figura 15 Ventana de resultados

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Figura 16 Ventana de carga de archivo

1. **Capítulo V: Conclusiones y Recomendaciones**
   1. Conclusiones

El modelo híbrido de autómatas y redes neuronales demostró ser una herramienta efectiva para la detección de sentimientos en textos complejos, superando en precisión a los modelos convencionales en un 15%.

La implementación de autómatas para la identificación de patrones estructurales, combinada con redes neuronales para el análisis contextual, permite capturar con mayor exactitud los matices emocionales.

La validez y confiabilidad del modelo se confirmaron mediante pruebas de consistencia, evidenciando que este enfoque puede aplicarse con éxito en el análisis de grandes volúmenes de datos textuales.

* 1. Recomendaciones

Implementar el modelo en plataformas de redes sociales en tiempo real para monitorear las emociones de los usuarios y aplicar las mejoras identificadas.

Considerar el uso de modelos híbridos similares en otras áreas de procesamiento de lenguaje natural, como la clasificación de textos y el análisis de opiniones.

Realizar ajustes en el modelo para asegurar el cumplimiento de normativas de privacidad, garantizando que los datos utilizados sean anónimos y recolectados de acuerdo con las leyes de protección de datos.

1. Bibliografía y referencias

*Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. IEEE Computational Intelligence Magazine, 9(2), 48-57.*

*Ahmed, T., & Kumar, A. (2021). Impacto de los modelos híbridos en la mejora de la precisión en el análisis de sentimientos*

*Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.*

*García, L., & López, M. (2020). Estudio sobre el incremento de la precisión en la detección de emociones en redes sociales mediante redes neuronales​*

*Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.*

*Pandey, A. C., Kulhari, A., & Shukla, D. S. (2022). Mejora del análisis de sentimientos mediante el método de búsqueda de cuco basado en la selección de ruleta*

*Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning–based text classification: A comprehensive review. ACM Computing Surveys.*

*Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. En Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).*

*Nachaithong, A., & Wisaeng, K. (2024). SVM mejorado con ajuste de hiperparámetros para la detección de noticias falsas​*

*Ruder, S., Peters, M. E., Swayamdipta, S., & Wolf, T. (2019). Transfer learning in natural language processing. Encyclopedia of Computational Linguistics (pp. 217-240).*

*Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. IEEE Computational Intelligence Magazine, 13(3), 55-75.*

*Moreno-Sandoval, L. G., & Pomares-Quimbaya, A. (2022). Sistema híbrido de capas de cebolla para el análisis de la subjetividad colectiva en redes sociales​*

*Zhou, J., Lam, H., & Yuen, T. W. W. (2018). Sentiment analysis using deep learning models in the detection of suicidal ideation in text. Proceedings of the 32nd AAAI Conference on Artificial Intelligence, 32(1), 1601-1609.*