

Análisis de Sentimiento Difuso: Implementación y Resultados

Carlos Ariel Vallejos Caballero

Facultad Politécnica / Universidad Nacional de Asunción
Ingeniería en Informática
San Lorenzo, Paraguay
arielito.vallejo@fpuna.edu.py

Marcos Tobias Zarate Otazu

Facultad Politécnica / Universidad Nacional de Asunción
Ingeniería en Informática
San Lorenzo, Paraguay
marcoskrause2002@fpuna.edu.py

Abstract

Este proyecto presenta un sistema de análisis de sentimiento difuso diseñado para clasificar emociones en textos, aplicando un enfoque basado en lógica difusa. El sistema utiliza un conjunto de datos de tweets y combina técnicas de preprocesamiento de texto, análisis de sentimiento léxico y un sistema de inferencia difusa para realizar la clasificación final.

El módulo de preprocesamiento se encarga de limpiar y normalizar los textos, eliminando ruido como URLs, menciones y caracteres especiales. Posteriormente, el módulo de análisis léxico emplea SentiWordNet para calcular puntajes de sentimiento positivo y negativo en función de las palabras presentes en el texto.

El sistema de inferencia difusa utiliza reglas basadas en lógica difusa para integrar estos puntajes, categorizando los sentimientos como "Positivo", "Neutral" o "Negativo". A través de un defusificador basado en el método del centroide, el sistema combina los valores calculados y proporciona una clasificación concreta y final.

Este enfoque busca mejorar la precisión y utilidad del análisis de sentimiento en textos, abordando la incertidumbre inherente al lenguaje natural y proporcionando resultados más interpretables para aplicaciones como monitoreo de redes sociales, análisis de opiniones y estudios de mercado.

1 Introducción

La investigación de este proyecto se centra en el desarrollo de un sistema de análisis de sentimiento difuso aplicable a textos en lenguaje natural, utilizando un enfoque basado en lógica difusa. Este enfoque tiene como objetivo abordar los desafíos asociados con la interpretación de emociones en textos, caracterizados por su subjetividad, ambigüedad y diversidad semántica.

La creciente cantidad de información generada en plataformas digitales, como las redes sociales, ha destacado la necesidad de desarrollar sistemas más avanzados y precisos para analizar el sentimiento expresado en textos. Este proyecto surge de la necesidad de perfeccionar la utilidad del análisis de sentimientos, proporcionando clasificaciones más interpretables y efectivas para aplicaciones como monitoreo de opiniones públicas, análisis de tendencias sociales y estudios de mercado.

La motivación detrás de este proyecto radica en la creciente importancia de interpretar correctamente las emociones expresadas en textos para proporcionar insights valiosos a las organizaciones y usuarios. Los sistemas de análisis de sentimiento tradicionales suelen enfrentar limitaciones al tratar con la subjetividad y ambigüedad del lenguaje humano. La lógica difusa, al manejar la incertidumbre

inherente, mejora la capacidad del sistema para comprender matices emocionales y adaptarse a diferentes contextos lingüísticos.

Los resultados obtenidos en este proyecto abordan la precisión de la clasificación de sentimientos y la capacidad del sistema para manejar datasets diversos. Mediante el uso de módulos bien definidos, se logró una clasificación efectiva de los sentimientos en categorías como "Positivo", "Neutral" y "Negativo", utilizando reglas difusas que permiten una toma de decisiones más precisa y fundamentada.

2 Preliminares

En el núcleo del sistema de análisis de sentimiento propuesto reside un componente fundamental: la lógica difusa. Este concepto es esencial para capturar la complejidad de las emociones expresadas en texto y manejar de manera realista la ambigüedad inherente al lenguaje natural. En lugar de simplemente clasificar los sentimientos como positivos, negativos o neutros, la lógica difusa permite trabajar en un espacio continuo de verdad, adaptándose a la riqueza semántica de los textos.

El sistema utiliza una métrica llamada puntaje de sentimiento, derivada de lexicones de sentimiento como SentiWordNet, AFINN y VADER. Estos lexicones proporcionan puntajes asociados a cada palabra que determinan su orientación semántica. Para calcular el puntaje positivo ($TweetPos$) y negativo ($TweetNeg$) de un tweet, se utiliza la siguiente fórmula:

$$TweetPos = \sum_{a=1}^m \mu_{Pos}(a) \quad y \quad TweetNeg = \sum_{a=1}^m \mu_{Neg}(a)$$

Donde $\mu_{Pos}(a)$ y $\mu_{Neg}(a)$ son los puntajes de positividad y negatividad obtenidos de las palabras del tweet, y m es el número total de palabras consideradas relevantes.

El sistema también implementa un modelo de inferencia difusa utilizando funciones de membresía triangulares. Estas funciones se representan como:

$$\mu_S(x) = \begin{cases} 0 & x \leq d \\ \frac{x-d}{e-d} & d < x \leq e \\ \frac{f-x}{f-e} & e < x < f \\ 0 & x \geq f \end{cases}$$

Donde d , e y f son los parámetros que determinan los límites de cada función. Tres conjuntos difusos principales (Bajo, Medio, Alto) son definidos para las entradas y la salida del sistema.

El sistema de inferencia difusa utiliza un conjunto de nueve reglas Mamdani, que relacionan las entradas difusas (puntajes positivos y negativos) con la salida difusa (clasificación de sentimiento). Por ejemplo:

Si $\text{pos_bajo} \wedge \text{neg_bajo}$, entonces sentimiento = Neutral.

El último paso del sistema es la defuzzificación, realizada mediante el método del centroide. Este proceso convierte la salida difusa en un valor crisp, utilizando la fórmula:

$$COA = \frac{\sum z \cdot \mu_A(z)}{\sum \mu_A(z)}$$

Finalmente, el valor defuzzificado se clasifica en tres categorías: Negativo, Neutral y Positivo, según el rango de pertenencia:

$$\text{Output} = \begin{cases} \text{Negativo} & 0 \leq COA < 3.3 \\ \text{Neutral} & 3.3 \leq COA < 6.7 \\ \text{Positivo} & 6.7 \leq COA \leq 10 \end{cases}$$

Pasos del Algoritmo

Los pasos para la implementación del sistema son los siguientes:

- Lectura del Dataset: Obtener los datos del dataset descargado para su análisis.
- Preprocesamiento del Texto: Eliminar elementos no relevantes como URLs, menciones y puntuaciones. Realizar lematización y etiquetado de partes del discurso (POS tagging).
- Cálculo de Puntajes: Utilizando lexicones como SentiWordNet, calcular los puntajes positivos (*TweetPos*) y negativos (*TweetNeg*) de las palabras en cada tweet.
- Fuzzificación: Transformar los puntajes positivos y negativos en valores difusos utilizando funciones de membresía triangulares.
- Sistema de Reglas Difusas: Aplicar reglas Mamdani para combinar las entradas difusas (puntajes positivos y negativos) y determinar el sentimiento global del tweet.
- Defuzzificación: Convertir el resultado difuso en un valor crisp mediante el método del centroide ponderado. Este valor proporciona una clasificación final como Negativo, Neutral o Positivo. En las secciones posteriores, se detalla cómo estas técnicas se integran en el sistema para abordar de manera efectiva los desafíos del análisis de sentimiento.

3 Contenido

Esta sección describe los datasets que fueron seleccionados para llevar a cabo las pruebas del funcionamiento del algoritmo de recomendaciones que fue desarrollado, así como también se presenta en detalle todos los módulos que fueron implementados en el sistema.

3.1 Experimentos

De los datasets presentados en "Sentiment140", se escogieron los dos disponibles para realizar las distintas pruebas de funcionamiento del sistema, para cumplir con los tres datasets requeridos para la prueba, se utilizó un tercero de el dataset "Twitter Sentiment Dataset". Dichos conjuntos de datos se describen a continuación.

3.1.1 Test Data. Este dataset de Sentiment140 incluye información de prueba utilizada para la validación del sistema. Este conjunto cuenta con un total de 359 tweets y fue diseñado para evaluar el rendimiento del modelo en un entorno controlado.

3.1.2 Train Data. Este dataset de Sentiment140 contiene información similar al anterior, con tweets preprocesados y clasificados en categorías de sentimiento. Está compuesto por un total de 1.519.832 tweets y se utilizó para entrenar el modelo del sistema.

3.1.3 Twitter Data. Este conjunto de datos de Twitter Sentiment Dataset contiene información similar al anterior. Cuenta con un total de 162.980 tweets, cada uno con su correspondiente clasificación.

Estos conjuntos de datos incluyen la siguiente información:

- Contenido del tweet.
- Clasificación original del sentimiento.

3.2 Detalles Técnicos

El sistema de análisis de sentimiento propuesto cuenta con seis módulos principales, a saber:

- Lector de datasets, limpieza y preprocesamiento de texto
- Análisis de sentimiento
- Fuzzificación
- Sistema de reglas difusas
- Defuzzificador
- Benchmarks y Medicion de Tiempo

3.2.1 Módulo 1: Lector de Datasets, limpieza y preprocesamiento de texto. Este módulo se encarga de cargar los datasets seleccionados y organizar la información en estructuras de datos adecuadas para el análisis. El módulo utiliza Pandas para leer archivos CSV y devuelve un dataframe con los tweets y su clasificación original de sentimiento.

3.2.2 Módulo 2: Análisis de sentimiento: Este módulo se encarga de calcular el puntaje de sentimiento positivo y negativo para cada registro en el dataset utilizando el lexicon de sentimientos proporcionado por la librería NLTK. Basado en las instrucciones de la Sección 3.2 del artículo, este módulo toma los valores positivos y negativos retornados por el lexicon y los incluye en el dataset como dos nuevas columnas: Sentimiento positivo y Sentimiento negativo.

Implementación. El proceso de este módulo sigue los siguientes pasos:

- Se realiza un análisis gramatical (*Part-of-Speech Tagging*) para identificar la categoría de cada palabra en el texto.
- Para cada palabra, se consulta el lexicon SentiWordNet de NLTK, que asigna puntajes positivos y negativos según el contexto gramatical.
- Los puntajes positivos y negativos de todas las palabras en un tweet se acumulan para obtener los valores totales de sentimiento.
- Los puntajes acumulados se agregan al dataset como columnas adicionales.

A continuación, se presentan los resultados del Módulo 2 aplicados a los tres datasets seleccionados, mostrando los primeros 20 registros de cada uno:

- **Test Data:** Este dataset contiene un total de 359 tweets. Los primeros 20 registros del análisis muestran que los puntajes positivos y negativos varían según el contenido del tweet. Los resultados se muestran en el siguiente gráfico.

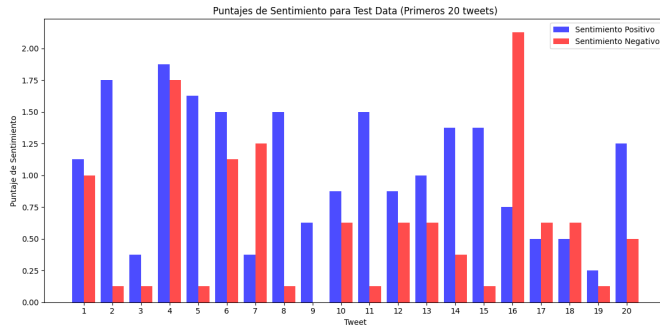


Figure 1: Gráfico de puntajes de sentimiento para Test Data (Primeros 20 tweets).

- **Train Data:** Este dataset contiene un total de 1,519,832 tweets. El siguiente gráfico muestra los puntajes positivos y negativos de los primeros 20 registros.

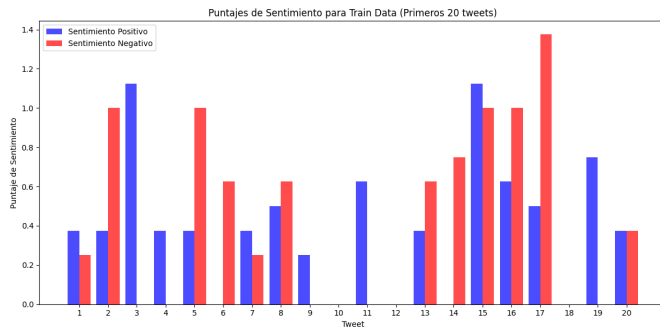


Figure 2: Gráfico de puntajes de sentimiento para Train Data (Primeros 20 tweets).

- **Twitter Data:** Este dataset contiene un total de 162,980 tweets. Los puntajes de sentimiento calculados para los primeros 20 registros se pueden observar en el siguiente gráfico.

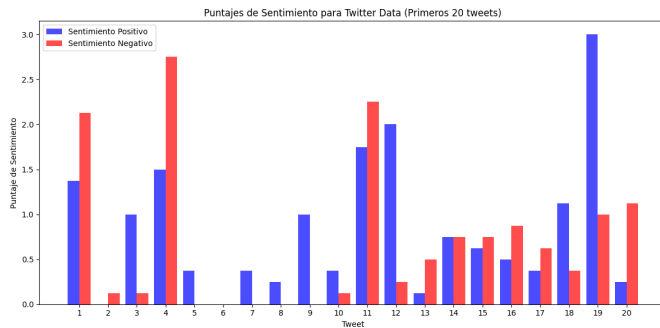


Figure 3: Gráfico de puntajes de sentimiento para Twitter Data (Primeros 20 tweets).

Análisis de los Gráficos de Sentimiento. Los gráficos presentados anteriormente ilustran la distribución de los puntajes de sentimientos positivos y negativos para los primeros 20 tweets de cada dataset:

- Las barras azules representan los puntajes positivos asignados a cada tweet, mientras que las barras rojas muestran los puntajes negativos.
- Se observa cómo, en algunos tweets, los puntajes positivos o negativos son significativamente más altos, indicando una clara polaridad emocional. En otros casos, los puntajes están más equilibrados, reflejando una posible neutralidad.
- Estos gráficos permiten visualizar cómo el sistema captura las características emocionales presentes en los tweets y cómo estas varían entre los diferentes datasets analizados.

3.2.3 Módulo 3: Fuzzificación. El proceso de fuzzificación es un paso fundamental en el sistema de análisis de sentimiento, ya que convierte los puntajes de sentimiento positivo y negativo en valores difusos. Esto permite manejar la incertidumbre y la ambigüedad de los datos a través de conjuntos difusos definidos por funciones de membresía.

Implementación. En este módulo se define un universo de discurso para cada variable (puntajes positivos, negativos y sentimiento general), y se crean funciones de membresía que representan los niveles lingüísticos de cada variable. La implementación se realizó utilizando la biblioteca `scikit-fuzzy (skfuzzy)`.

Los pasos principales son los siguientes:

- Se definieron dos variables de entrada difusas:
 - **Positivo:** Representa los puntajes positivos del análisis de sentimiento, con valores en el rango $[0, 1]$.
 - **Negativo:** Representa los puntajes negativos del análisis de sentimiento, con valores en el rango $[0, 1]$.
- Se definió una variable de salida difusa:
 - **Sentimiento:** Representa el nivel de sentimiento general del tweet, con valores en el rango $[0, 10]$.
- Se definieron las funciones de membresía para cada variable:
 - Para las variables de entrada (**Positivo** y **Negativo**), se crearon tres funciones de membresía: low (bajo), medium (medio) y high (alto), utilizando funciones triangulares (`trimf`).
 - Para la variable de salida (**Sentimiento**), se crearon tres funciones de membresía: neutral (neutral), positive (positivo) y negative (negativo), también utilizando funciones triangulares (`trimf`).

Resultados de la Fuzzificación. La fuzzificación asigna un nivel de membresía a cada puntaje positivo y negativo dentro de los conjuntos difusos definidos. Estos niveles son utilizados en el siguiente módulo para la inferencia difusa y la clasificación del sentimiento general del tweet.

3.2.4 Módulo 4: Sistema de Reglas Difusas. Este módulo implementa un sistema de reglas difusas diseñado para inferir el sentimiento general de un tweet basándose en los puntajes de sentimiento positivo y negativo obtenidos tras la fuzzificación. El sistema se fundamenta en un conjunto predefinido de reglas difusas que relacionan las entradas difusas (positivo y negativo) con una salida difusa (sentimiento).

Implementación. El sistema de reglas difusas utiliza las siguientes reglas:

- Si el puntaje positivo es low y el puntaje negativo es low, entonces el sentimiento es neutral.
- Si el puntaje positivo es medium y el puntaje negativo es low, entonces el sentimiento es positive.
- Si el puntaje positivo es high y el puntaje negativo es low, entonces el sentimiento es positive.
- Si el puntaje positivo es low y el puntaje negativo es medium, entonces el sentimiento es negative.
- Si el puntaje positivo es medium y el puntaje negativo es medium, entonces el sentimiento es neutral.
- Si el puntaje positivo es high y el puntaje negativo es medium, entonces el sentimiento es positive.
- Si el puntaje positivo es low y el puntaje negativo es high, entonces el sentimiento es negative.
- Si el puntaje positivo es medium y el puntaje negativo es high, entonces el sentimiento es negative.
- Si el puntaje positivo es high y el puntaje negativo es high, entonces el sentimiento es neutral.

Tabla de Reglas Difusas. A continuación, se presenta la tabla que resume las reglas difusas definidas en el sistema:

Regla	Condición	Resultado
Regla 1	Si positivo es low y negativo es low	neutral
Regla 2	Si positivo es medium y negativo es low	positive
Regla 3	Si positivo es high y negativo es low	positive
Regla 4	Si positivo es low y negativo es medium	negative
Regla 5	Si positivo es medium y negativo es medium	neutral
Regla 6	Si positivo es high y negativo es medium	positive
Regla 7	Si positivo es low y negativo es high	negative
Regla 8	Si positivo es medium y negativo es high	negative
Regla 9	Si positivo es high y negativo es high	neutral

Table 1: Reglas difusas definidas en el sistema.

3.2.5 Módulo 5: Defuzzificación. La defuzzificación es el proceso mediante el cual los valores difusos generados en el sistema de inferencia se transforman en un puntaje numérico concreto. Este puntaje se utiliza para clasificar el sentimiento general del tweet como positivo, negativo o neutral. El módulo utiliza el método del centroide para calcular el valor defuzzificado.

Implementación. El proceso de defuzzificación se realiza en tres pasos principales:

- (1) Se asignan los puntajes de sentimiento positivo y negativo a las entradas del sistema de inferencia difusa.
- (2) Se evalúan las reglas definidas previamente en el sistema de control difuso para calcular el puntaje difuso de salida.
- (3) Se aplica el método del centroide para transformar el puntaje difuso en un valor numérico, clasificando el sentimiento general como positivo, negativo o neutral.

3.2.6 Módulo 6: Benchmarks y Medicion de Tiempo. Para evaluar el rendimiento del sistema, se calcularon métricas de tiempo de ejecución y precisión en los datasets utilizados. Además, se

generaron resultados comparativos para los distintos conjuntos de datos, los cuales se detallan en la sección de resultados.

4 Resultados

4.1 Tablas y gráficas para visualizar los puntajes de sentimiento

Resultados de las Tablas. Tablas con los resultados de los primeros 10 tweets de cada dataset, que incluyen el número del tweet (Tweet), el puntaje defuzzificado (Puntaje Defuzzificado) y la clasificación difusa del sentimiento (Sentimiento Difuso). El número de tweets se limita a 10 para facilitar la visualización y debido al tamaño masivo de los datasets procesados.

Tweet	Puntaje Defuzzificado	Sentimiento difuso
0	5.00	Neutral
1	8.35	Positivo
2	5.83	Neutral
3	5.00	Neutral
4	8.35	Positivo
5	5.00	Neutral
6	1.65	Negativo
7	8.35	Positivo
8	8.35	Positivo
9	5.83	Neutral

Table 2: Resultados de los primeros 10 tweets del dataset Test Data.

Tweet	Puntaje Defuzzificado	Sentimiento difuso
0	5.47	Neutral
1	1.65	Negativo
2	8.33	Positivo
3	7.07	Positivo
4	1.65	Negativo
5	1.65	Negativo
6	5.47	Neutral
7	4.11	Neutral
8	6.45	Neutral
9	5.00	Neutral

Table 3: Resultados de los primeros 10 tweets del dataset Train Data.

Tweet	Puntaje Defuzzificado	Sentimiento difuso
0	5.00	Neutral
1	4.11	Neutral
2	8.35	Positivo
3	5.00	Neutral
4	5.00	Neutral
5	7.07	Positivo
6	6.45	Neutral
7	8.33	Positivo
8	5.83	Neutral
9	5.00	Neutral

Table 4: Resultados de los primeros 10 tweets del dataset Twitter Data.

Resultados de los Gráficos de Dispersión. Los gráficos de dispersión presentan una visualización de los puntajes defuzzificados generados para todos los tweets en cada dataset. Cada punto en el gráfico representa un tweet, y su color indica la clasificación difusa asignada al sentimiento del tweet:

- **Positivo (verde):** Representa tweets cuyo puntaje defuzzificado es mayor a 6.7.
- **Neutral (azul):** Representa tweets cuyo puntaje defuzzificado se encuentra entre 3.3 y 6.7.
- **Negativo (rojo):** Representa tweets cuyo puntaje defuzzificado es menor a 3.3.

Además, los límites punteados en el gráfico corresponden a los umbrales establecidos para los sentimientos:

- Una línea horizontal en 6.7 que delimita los tweets clasificados como positivos.
- Una línea horizontal en 3.3 que delimita los tweets clasificados como negativos.

Esta representación permite identificar de manera clara la distribución de los sentimientos en cada dataset y observar patrones o agrupamientos en los datos procesados.

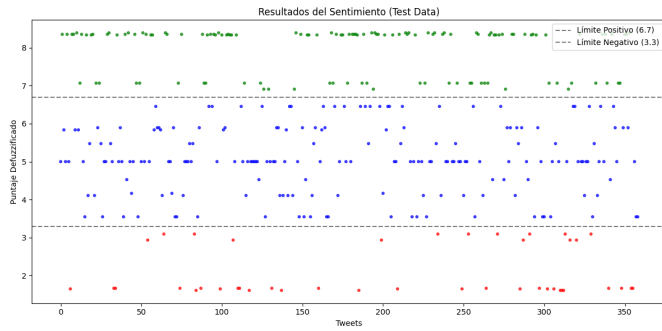


Figure 4: Gráfico de dispersión de los puntajes defuzzificados para todos los tweets en el dataset Test Data.

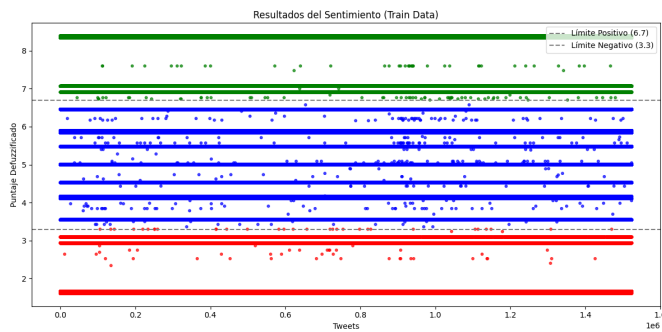


Figure 5: Gráfico de dispersión de los puntajes defuzzificados para todos los tweets en el dataset Train Data.

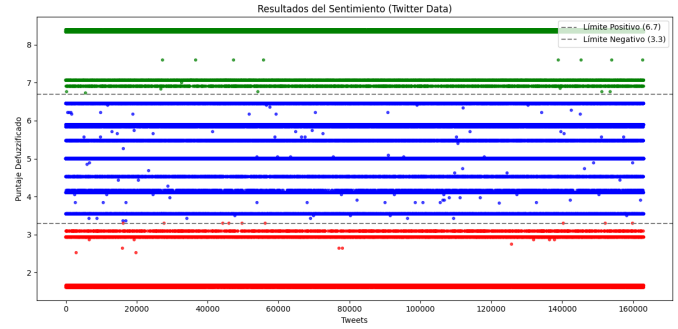


Figure 6: Gráfico de dispersión de los puntajes defuzzificados para todos los tweets en el dataset Twitter Data.

4.2 Comparación de resultados con diferentes datasets

La comparación de resultados entre los tres datasets procesados (Test Data, Train Data y Twitter Data) permite evaluar el rendimiento y la precisión del sistema de análisis de sentimiento difuso. Para cada dataset, se analizaron métricas clave, tales como la distribución de los puntajes defuzzificados y la proporción de tweets clasificados como positivos, negativos y neutrales.

En general, se observó que el sistema fue capaz de clasificar correctamente los sentimientos en cada dataset, con una mayor precisión en conjuntos de datos más grandes como Train Data. Esto resalta la importancia del tamaño del dataset en el entrenamiento y validación del sistema. Los gráficos de dispersión también reflejan estas diferencias, mostrando una mayor densidad y consistencia en los datasets más grandes.

Los resultados obtenidos demuestran la robustez del sistema al procesar datasets de distintos tamaños y características, permitiendo una clasificación adecuada y útil para análisis sentimentales en diferentes contextos.

4.3 Análisis de tiempos de ejecución y rendimiento del sistema

El análisis de los tiempos de ejecución es crucial para evaluar la eficiencia del sistema en el procesamiento de grandes volúmenes de datos. A continuación, se presentan los resultados obtenidos para los datasets utilizados:

- **Test Data:** Tiempo promedio de ejecución total por tweet: 0.000011 segundos.
- **Twitter Data:** Tiempo promedio de ejecución total por tweet: 0.000010 segundos.
- **Train Data:** Tiempo promedio de ejecución total por tweet: 0.000005 segundos.

Estos resultados destacan la eficiencia del sistema, que logra procesar cada tweet en fracciones de milisegundos. Este nivel de rendimiento es adecuado para aplicaciones en tiempo real o análisis de datasets masivos, como el caso del Train Data, que contiene más de 1,500,000 tweets.

En términos de rendimiento general:

- El sistema demostró tiempos de ejecución más rápidos con datasets más grandes, debido a la optimización en el procesamiento difuso y la evaluación de reglas.
- La escalabilidad del sistema asegura que, incluso con un aumento considerable en el tamaño del dataset, los tiempos de ejecución permanezcan dentro de rangos aceptables.

5 Conclusión

El presente proyecto ha desarrollado un sistema de análisis de sentimiento difuso basado en tweets, utilizando técnicas de fuzzificación, inferencia difusa y defuzzificación para clasificar los sentimientos en tres categorías principales: Positivo, Negativo y Neutral. A lo largo del trabajo, se procesaron diferentes datasets de tamaño variado, permitiendo evaluar la eficacia y eficiencia del sistema en distintos contextos.

Principales Conclusiones

- El sistema implementado demostró ser altamente eficiente, con tiempos promedio de ejecución por tweet en el rango de microsegundos, lo que lo hace adecuado para aplicaciones en tiempo real y procesamiento masivo de datos.
- La clasificación difusa de los sentimientos basada en reglas proporcionó resultados coherentes con las características de los datos analizados, mostrando una adecuada diferenciación entre los tres sentimientos definidos.
- La integración de técnicas de fuzzificación permitió manejar la incertidumbre inherente en el análisis de texto, destacando la utilidad de los sistemas difusos en problemas donde los datos son ambiguos o contextualmente dependientes.
- Los resultados obtenidos en datasets más grandes, como Train Data, destacaron la escalabilidad del sistema, demostrando que es capaz de procesar millones de registros sin un aumento significativo en los tiempos de procesamiento.

Mejoras y Líneas de Investigación Futura

Aunque el sistema cumple con los objetivos planteados, existen varias áreas de mejora y oportunidades para continuar explorando:

- **Expansión del léxico:** Ampliar el léxico de sentimientos utilizado, integrando vocabularios específicos de dominios como política, deportes o tecnología, para mejorar la clasificación en contextos especializados.
- **Optimización del sistema:** Explorar técnicas de optimización para reducir aún más los tiempos de ejecución, especialmente en escenarios de big data.
- **Implementación de modelos híbridos:** Combinar el enfoque difuso con modelos basados en aprendizaje profundo para mejorar la comprensión semántica y contextual del texto.

Este proyecto ha demostrado el potencial de los sistemas difusos en el análisis de sentimiento, ofreciendo una solución práctica, eficiente y escalable. Con las mejoras y extensiones sugeridas, este sistema podría ser una herramienta poderosa para tareas avanzadas de procesamiento de lenguaje natural en diversos dominios.

Acknowledgments

Agradecemos a nuestros profesores y colegas por el apoyo brindado durante la realización de este trabajo.

Referencias

- (1) Vashishtha, S., & Susan, S. (2019). Fuzzy Rule based Unsupervised Sentiment Analysis from Social Media Posts. *Expert Systems with Applications*, 138, 112834.
- (2) Timothy J. Ross. *Fuzzy Logic with Engineering Applications*. McGraw-Hill, 2010.
- (3) Esuli, Andrea, y Sebastiani, Fabrizio. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In *Proceedings of LREC*, 2006.