

Aplicación de Redes Neuronales y n-Gramas para el Análisis de Sentimientos en Críticas de Cine

Application of Neural Networks and n-Grams for Sentiment Analysis in Movie Reviews

Cesar Alexis Estrada Palacios^{1,*}  y José Luis Tapia Fabela¹ 

¹Universidad Autónoma del Estado de México
Paraje el Tejocote, 52640, San Pedro Tlaltizapan, Estado de México, México
[*kirdrazler@gmail.com](mailto:kirdrazler@gmail.com)

PALABRAS CLAVE: RESUMEN

Optimización, aprendizaje automático, procesamiento de lenguaje natural, construcción de características

El análisis de sentimientos en críticas de cine desempeña un papel crucial tanto en la industria del cine como en la investigación académica. Este artículo, se enfoca en la combinación de n-gramas y el uso de redes neuronales para mejorar la clasificación para el análisis de sentimientos en críticas de cine en español. El método propuesto utiliza la combinación de n-gramas como modelo de representación y se evalúa en el corpus Muchocine, comparándolo con los métodos del estado del arte. Los resultados destacan un rendimiento superior en términos de precisión, recall y f-measure, demostrando así la eficacia de la combinación de n-gramas en el rendimiento del clasificador.

KEYWORDS:

Optimization, machine learning, natural language processing, feature engineering

ABSTRACT

Sentiment analysis in movie reviews plays a crucial role in both the film industry and academic research. This article focuses on the combination of n-grams and the use of neural networks to enhance sentiment classification in Spanish-language movie reviews. The proposed method utilizes n-gram combination as a representation model and is evaluated on the Muchocine corpus, comparing it with state-of-the-art methods. The results highlight superior performance in terms of precision, recall, and f-measure, demonstrating the effectiveness of n-gram combination in the classifier's performance.

• Recibido: 13 de agosto de 2023 • Aceptado: 21 de agosto de 2023 • Publicado en línea: 1 de octubre de 2024

1. INTRODUCCIÓN

Este artículo científico aborda el análisis de sentimientos en críticas de cine en español utilizando el corpus Muchocine, el cual es una colección de críticas de cine en español. Se enfoca en la combinación de n-gramas secuenciales contiguos de n elementos (en este caso, palabras) y el uso de redes neuronales para mejorar la clasificación de emociones y actitudes en el lenguaje natural. El estudio demuestra la importancia de adaptar los enfoques existentes a diferentes idiomas y contextos lingüísticos, con resultados superiores a métodos previos en el estado del arte.

El análisis de sentimientos ha sido un campo de investigación en constante evolución, centrándose en la comprensión y clasificación de las emociones y actitudes expresadas en el lenguaje natural [1]. Esta área de estudio es fundamental en diversos campos, como el procesamiento del lenguaje natural, la inteligencia artificial y el aprendizaje automático. Permite extraer información valiosa de grandes cantidades de datos no estructurados, como publicaciones en redes sociales, reseñas de productos, opiniones de usuarios y comentarios en blogs, entre otros.



En los últimos años, ha habido un creciente interés en el análisis de sentimientos debido a su amplia gama de aplicaciones en campos como la toma de decisiones empresariales, el monitoreo de la reputación de una marca, la detección de tendencias de mercado y la evaluación de la satisfacción del cliente. Comprender las opiniones y actitudes de los usuarios es esencial para mejorar la calidad de los productos y servicios, y para desarrollar estrategias de marketing efectivas [2].

A pesar de las numerosas investigaciones sobre el análisis de sentimientos en inglés, existe una clara necesidad de investigar y desarrollar enfoques específicos para otros idiomas, como el español. El español es el segundo idioma más hablado a nivel mundial, con 493 millones de personas [3], y tiene una presencia significativa en las redes sociales, además de un crecimiento constante en el uso de Internet. Sin embargo, la mayoría de los métodos y recursos disponibles se han diseñado principalmente para el idioma inglés, lo que limita su aplicabilidad y precisión en el contexto del español.

El objetivo de este artículo científico fue abordar esta brecha y contribuir al campo del análisis de sentimientos en español. Nuestro enfoque se centró en desarrollar un sistema de análisis de sentimientos preciso y efectivo para textos en español, utilizando técnicas avanzadas de procesamiento del lenguaje natural y aprendizaje automático. En particular, se puso énfasis en el uso de redes neuronales, que han demostrado su eficacia en diversos problemas relacionados con el procesamiento del lenguaje natural [4].

Para lograr estos objetivos, se llevaron a cabo varios pasos metodológicos. En primer lugar, se recopiló un corpus de críticas de cine. Este corpus se utilizó para entrenar y evaluar nuestro método de análisis de sentimientos propuesto.

A continuación, se aplicaron técnicas de preprocesamiento específicas para el español,

teniendo en cuenta las particularidades lingüísticas y gramaticales del idioma. Esto permitió optimizar el rendimiento del sistema al considerar aspectos como la lematización, y la detección de palabras clave relevantes para el análisis de sentimientos.

Posteriormente, se desarrollaron y evaluaron modelos de aprendizaje automático basados en redes neuronales para realizar el análisis de sentimientos en español. Estos modelos se entrenaron utilizando el corpus de datos recopilado y se ajustaron mediante técnicas de validación cruzada, el cual es un método que divide el corpus en subconjuntos para entrenar y evaluar el modelo de manera iterativa y la optimización de hiperparámetros de la red neuronal, aprovechando el potencial de las redes neuronales en la captura de patrones complejos en los datos [5].

En resumen, este artículo científico abordó la necesidad de un análisis de sentimientos específico para el español y propuso un sistema basado en redes neuronales para lograr una clasificación precisa y efectiva de las emociones y actitudes en el lenguaje natural. Los resultados obtenidos demuestran el potencial de las técnicas de aprendizaje automático y resaltan la importancia de adaptar los enfoques existentes a los diferentes idiomas y contextos lingüísticos.

2. ESTADO DEL ARTE

En este apartado se presenta el estado del arte en el campo del análisis de sentimientos y la clasificación de documentos basados en la opinión. El enfoque principal de esta investigación es analizar los sentimientos expresados en las críticas de cine en español, aunque dado el limitado número de trabajos específicos en esta área, también se exploran estudios relevantes en la clasificación de opiniones en otros dominios, los cuales pueden aportar conocimiento valioso para nuestro estudio.

La escasez de información en el estado del arte se atribuye a la marcada predominancia de investigaciones centradas en el idioma inglés en el ámbito del análisis de sentimientos especialmente en el ámbito del análisis de sentimientos en críticas de cine. Esta limitación puede explicarse por la hegemonía del inglés en la investigación científica y tecnológica, así como por la disponibilidad restringida de recursos y conjuntos de datos ya que el corpus Muchocine es el único corpus en español de críticas de cine.

Los autores [6] realizaron un trabajo fundamental en el análisis de sentimientos en críticas de cine en español. Propusieron un corpus llamado "Muchocine" que consiste en 3,878 críticas de cine extraídas del sitio web Muchocine. Este corpus se creó debido a la falta de datos disponibles para el análisis de sentimientos en críticas de cine en español. En este estudio, se utilizó un subconjunto de 400 críticas, divididas en 200 críticas positivas y 200 críticas negativas, excluyendo las críticas neutras. Para la optimización de parámetros, se dividió este conjunto en un 80% para entrenamiento y un 20% para pruebas. Los resultados obtenidos mostraron que el mejor desempeño alcanzado fue del 77%. Cabe destacar que en este estudio se emplearon bigramas de palabras, aunque no se realizaron combinaciones más complejas como se propone en nuestra investigación. Además, no se utilizaron clasificadores como SVM o redes neuronales debido a limitaciones de recursos.

Los autores [7] realizaron investigaciones relevantes en la clasificación de opiniones en español utilizando el corpus Muchocine. Obtuvieron resultados destacados al emplear SVM como clasificador, logrando una *precision* del 87.73%. Los mejores resultados se obtuvieron utilizando SVM en combinación con TF-IDF y sin aplicar técnicas de *stemming* ni eliminación de *stopwords*.

Otro trabajo relevante es el presentado por [8], que se centró en la atribución de autoría,

pero exploró características basadas en frecuencia de palabras y n-gramas, así como características lingüísticas y estilísticas. Aunque su enfoque difiere ligeramente de nuestra investigación, sus hallazgos pueden ser útiles para nuestro contexto.

3. MÉTODO PROPUESTO

3.1. Análisis y etiquetado de los datos

Para llevar a cabo la tarea de clasificación basada en redes neuronales, es necesario contar con la crítica y su respectiva clase en el corpus Muchocine. Estos elementos vienen en el mismo archivo, por lo que es necesario separarlos. Para obtener la crítica, se extrajo el contenido de las etiquetas <body> y <summary>, ya que combinar la crítica y su resumen mejora la clasificación [7]. Para obtener la clase, se extrajo el valor numérico de la etiqueta *rank*. Si el valor de *rank* es 1 o 2, se etiqueta como clase 0 (crítica negativa), y si el valor es 3 o 4, se etiqueta como clase 1 (crítica positiva). Al finalizar este proceso, se obtiene una tabla con los datos etiquetados.

3.2. Construcción y pesado de características

Antes de construir las características deseadas, se realiza un preprocesamiento del texto. En este caso, se eliminan saltos de línea, números, acentos, signos de puntuación y símbolos, dejando solo los caracteres que son letras. Además, se aplica la técnica de *stemming* para obtener la raíz de las palabras, unificando términos que aportan la misma información al clasificador. Sin embargo, cabe mencionar que aplicar *stemming* puede dar peores resultados en algunos casos [9].

Una vez preprocesado el texto, se procede a construir las características que conformarán la bolsa de n-gramas. Para el método propuesto, se generan n-gramas de diferentes tamaños y se combinan en una sola bolsa de n-gramas. Se exploran cuatro opciones de combinación: unigramas +

bigramas, unigramas + trigramas, bigramas + trigramas y unigramas + bigramas + trigramas. El objetivo es determinar si la combinación de n-gramas mejora los resultados y cuál de estas combinaciones obtiene los mejores resultados [8].

3.3 Pesado de características

En esta etapa, se asigna un valor a cada característica (n-grama) mediante el pesado de términos. Se utiliza la técnica TF-IDF, que asigna un peso a cada término en función de su frecuencia en el documento y su frecuencia inversa en el corpus. Esta técnica ha demostrado obtener buenos resultados en análisis de sentimientos [7]. A partir de los resultados del pesado de términos, se obtiene una tabla con los valores asignados a cada característica.

3.4 Selección de características

Dado el gran número de características generadas, es necesario reducir su cantidad para agilizar el proceso de aprendizaje de la inteligencia artificial. En este caso, se seleccionan las 25000 características más frecuentes, siguiendo el enfoque utilizado en otros trabajos [10].

3.5 Entrenamiento del modelo de clasificación

En esta etapa, se procede al entrenamiento del modelo de clasificación utilizando principalmente Redes Neuronales, una técnica poderosa y ampliamente reconocida en tareas de análisis de sentimientos y clasificación de texto. Las Redes Neuronales han demostrado su eficacia en diversos contextos y ofrecen la capacidad de capturar patrones y características complejas en los datos.

La combinación de n-gramas también se utiliza en conjunto con las Redes Neuronales para mejorar la *precision* de la clasificación. Los n-gramas permiten capturar la estructura y secuencia de las palabras en los textos, lo

que puede ser útil para la tarea de análisis de sentimientos.

3.6 Evaluación del método

Para evaluar la calidad de la solución proporcionada por los clasificadores, se utilizan métricas como *precision* y *accuracy*. La métrica *precision* se refiere a la proporción de casos correctamente clasificados en relación con todos los casos clasificados. Es una métrica importante que indica la capacidad del clasificador para identificar de manera precisa las clases objetivo.

El *accuracy* es la proporción de casos correctamente clasificados sobre el total de casos. Esta métrica nos permite evaluar el rendimiento general del modelo en términos de aciertos y errores en la clasificación.

La evaluación del método se realiza utilizando una matriz de confusión, que proporciona una visión detallada de los resultados del clasificador, mostrando los verdaderos positivos, falsos positivos, falsos negativos y verdaderos negativos. A partir de esta matriz, se calculan las métricas de *precision* y *accuracy*.

4. EXPERIMENTOS Y RESULTADOS

En esta sección, presentaremos una serie de experimentos y resultados para evaluar el rendimiento de las Redes Neuronales en la clasificación de textos. Compararemos estos resultados con los obtenidos por el método propuesto y otros en el estado del arte.

4.1. Comparación con los métodos del estado del arte

Realizamos una comparación entre los mejores resultados obtenidos por el método propuesto basado en Redes Neuronales y los métodos del estado del arte aplicados previamente al corpus Muchocine. Los resultados se presentan en la Tabla 1.

Tabla 1. Comparación de los resultados del método propuesto (Estrada) y los métodos de [6], [7].

| Autor | Clasificador | Precision | Recall | F-measure |
|----------|------------------|-----------|--------|-----------|
| Estrada | Redes neuronales | 90.62 | 90.28 | 90.45 |
| Martinez | SVM Lineal | 87.73 | 87.69 | 87.71 |
| Martinez | Naive Bayes | 84.08 | 84.01 | 84.04 |
| Cruz | SOv1-S | 77.50 | N/A | N/A |
| Cruz | SOv2-NS | 69.05 | N/A | N/A |

Los resultados muestran que el método propuesto basado en Redes Neuronales logra una *precision*, *recall* y puntuación *F-measure* superiores a los métodos del estado del arte [6], [7]. Específicamente, el método propuesto alcanza una *precision* del 90.52%, un *recall* del 88.75% y una puntuación *F-measure* del 89.62%, lo que representa una mejora significativa en comparación con los métodos de SVM Lineal y Naive Bayes utilizados por otros autores. Además, el método propuesto supera los resultados obtenidos por [6] utilizando los clasificadores SOv1 - S y SOv2 - NS.

Estos hallazgos respaldan la efectividad de las Redes Neuronales en la clasificación de textos y su potencial para superar a otros métodos en términos de rendimiento.

4.2. Experimento 1

Comparación entre diferentes arquitecturas de Redes Neuronales En este experimento, evaluamos diferentes arquitecturas de Redes Neuronales como clasificadores en el método propuesto, los hyper parámetros utilizados en cada red neuronal fueron los siguientes.

Tabla 2. Hyperparámetros de la Red neuronal.

| Hyperparámetro | Red neuronal A | Red neuronal B | Red neuronal C |
|---------------------|----------------------|----------------------|----------------------|
| Activation function | Sigmoid | Tanh | Relu |
| Neurons | 25000 | 25000 | 25000 |
| Loss | Binary cross-entropy | Binary cross-entropy | Binary cross-entropy |
| Optimizer | RMSprop | Adam | RMSprop |
| Learning rate | 0.001 | 0.002 | 0.001 |
| Rho | 0.9 | 0.9 | 0.9 |
| Epsilon | 1.00E-08 | 1.00E-08 | 1.00E-08 |
| Epochs | 1000 | 1000 | 1000 |
| Batch size | 12 | 12 | 12 |

Utilizando la combinación de unigramas + bigramas + trigramas como base de comparación. Los resultados obtenidos se presentan en la Tabla 3.

Tabla 3. Comparación entre diferentes modelos de redes neuronales.

| Modelo de Red Neuronal | Accuracy | Precision | Recall | F-measure |
|------------------------|----------|-----------|--------|-----------|
| Red Neuronal A | 90.52 | 90.62 | 90.28 | 90.45 |
| Red Neuronal B | 88.93 | 88.48 | 88.78 | 88.63 |
| Red Neuronal C | 86.79 | 86.78 | 86.56 | 86.56 |

Los resultados indican que la Red Neuronal A tuvo el mejor rendimiento, con una *precision* del 90.52% y una puntuación *F-measure* del 90.45% La Red Neuronal B también obtuvo buenos resultados, con una *precision* del 88.93% y una puntuación *F-measure* del 88.63%. Por otro lado, la Red Neuronal C tuvo un rendimiento ligeramente inferior, con una *precision* del 86.79% y una puntuación *F-measure* del 86.56%. Estos resultados sugieren que la Red Neuronal A es la mejor opción para la clasificación de textos en este corpus específico.

4.3. Experimento 2

Comparación de rendimiento entre diferentes tamaños de n-gramas de palabra En este experimento, evaluamos el rendimiento de tres modelos diferentes de n-gramas: unigramas, bigramas y trigramas, para identificar cuál proporciona mejores resultados. Los resultados se muestran en la Tabla 4.

Tabla 4. Comparación entre diferentes tamaños de n-gramas de palabra.

| Característica | Accuracy | Precision | Recall | F-measure |
|----------------|----------|-----------|--------|-----------|
| Unigramas | 87.35 | 87.18 | 86.79 | 87.04 |
| Bigramas | 84.92 | 84.74 | 84.35 | 84.54 |
| Trigramas | 82.67 | 82.48 | 82.19 | 82.33 |

De acuerdo con los resultados, los unigramas presentan una mayor *precision* con un 87.35%, seguidos de los bigramas con un 84.74% y los trigramas con un 82.48%. En cuanto al *recall*, los unigramas también tienen el mayor valor con un 86.79%, seguidos de los bigramas con un 84.35% y los trigramas con

un 82.19%. En cuanto a la puntuación *F-measure*, nuevamente los unigramas presentan un valor superior con un 87.04%, seguidos de los bigramas con un 84.54% y los trigramas con un 82.33%. En conclusión, los resultados indican que, en general, el tamaño de los unigramas proporciona un mejor rendimiento en las diferentes métricas evaluadas.

4.4. Experimento 3

Combinación de n-gramas como modelo de representación En este experimento, combinamos diferentes tamaños de n-gramas para crear una sola bolsa de n-gramas. Los resultados se muestran en la Tabla 5.

Tabla 5. Comparación de la combinación de las bolsas de n-gramas.

| Característica | Accuracy | Precision | Recall | F-measure |
|----------------------------------|----------|-----------|--------|-----------|
| Unigramas + Bigramas | 89.23 | 89.47 | 88.95 | 89.21 |
| Unigramas + Trigramas | 88.91 | 88.18 | 87.75 | 87.96 |
| Bigramas + Trigramas | 87.45 | 87.58 | 87.23 | 87.4 |
| Unigramas + Bigramas + Trigramas | 90.52 | 90.62 | 90.28 | 90.45 |

Basándonos en los resultados obtenidos, podemos afirmar que la combinación de unigramas + bigramas + trigramas es la que proporciona el mejor rendimiento, con una *precision* del 90.62%, un *recall* del 90.28% y una puntuación *F-measure* del 90.45%. Esto demuestra que la utilización de tres tamaños de n-gramas es más efectiva que la combinación de solo dos tamaños.

En segundo lugar, se encuentra la combinación de unigramas + bigramas, con una *precision* del 89.47%, un *recall* del 88.95% y una puntuación *F-measure* del 89.21%. Aunque estos resultados son buenos, no superan a la combinación de tres tamaños de n-gramas.

Las combinaciones de unigramas y trigramas, y bigramas y trigramas, presentaron resultados similares, pero aún

están por debajo de la mejor combinación de n-gramas. Esto indica que la combinación de tres tamaños de n-gramas es la que proporciona el mejor rendimiento en este corpus.

En conclusión, la combinación de tres tamaños de n-gramas puede ser de gran utilidad para mejorar la *precision* de los modelos de redes neuronales en el futuro. Los resultados obtenidos en este estudio proporcionan una guía útil para la elección de la combinación de n-gramas en la construcción de modelos de clasificación de texto.

A pesar de los resultados prometedores obtenidos en este estudio, todavía hay margen de mejora para alcanzar una *precision* del 99.9%. Por lo tanto, se espera que este trabajo sirva como punto de partida para futuros estudios que busquen optimizar la *precision* y aplicarlo en la vida real. Es importante mencionar que se probaron diferentes variantes del método, como distintas arquitecturas de redes neuronales, tamaños de n-gramas, combinación de n-gramas y el uso de técnicas de reducción de dimensionalidad.

CONCLUSIONES

En este artículo científico, se abordó la necesidad de desarrollar un sistema de análisis de sentimientos específico para el español, utilizando técnicas avanzadas de procesamiento del lenguaje natural y aprendizaje automático.

Los resultados obtenidos demostraron que el enfoque propuesto basado en redes neuronales supera a los métodos del estado del arte, logrando una *precision*, *recall* y puntuación *F-measure* superiores. Esto demuestra la efectividad de las redes neuronales en la clasificación de textos y su potencial para capturar patrones complejos en los datos.

Además, se exploraron diferentes arquitecturas de redes neuronales para obtener los mejores resultados en términos de *precision*, *recall* y *f-measure*. Esto resalta la importancia de seleccionar la arquitectura adecuada para cada tarea de análisis de sentimientos.

En cuanto a los trabajos futuros, se podrían considerar las siguientes direcciones de investigación:

- 1 Mejora del preprocesamiento y la construcción de características: Se podría investigar en técnicas más avanzadas de preprocesamiento y construcción de características específicas para el español, teniendo en cuenta las particularidades lingüísticas y gramaticales del idioma. Esto podría incluir el uso de técnicas de eliminación de ruido y normalización de texto.
- 2 Exploración de enfoques de transferencia de aprendizaje: Investigar en enfoques de transferencia de aprendizaje para aprovechar los conocimientos y modelos pre entrenados en otros idiomas o dominios relacionados. Esto podría ayudar a mejorar el rendimiento del sistema en situaciones de datos limitados o dominios específicos.
- 3 Evaluación en otros dominios y géneros textuales: Si bien este estudio se centró en el análisis de sentimientos en críticas de cine en español, sería interesante evaluar el sistema en otros dominios y géneros textuales, como noticias, comentarios en redes sociales o reseñas de productos. Esto permitiría generalizar los resultados y evaluar la aplicabilidad del sistema en diferentes contextos.
- 4 Interpretabilidad de los resultados: Sería importante investigar en técnicas de interpretabilidad de los resultados del análisis de sentimientos, especialmente en el contexto de las redes neuronales. Esto permitiría comprender mejor qué características y patrones están influyendo en las predicciones del sistema y brindaría

una mayor confianza en los resultados obtenidos.

En conclusión, este artículo científico contribuyó al campo del análisis de sentimientos en español mediante el desarrollo de un sistema preciso y efectivo basado en redes neuronales. Los resultados obtenidos destacaron la importancia de adaptar los enfoques existentes a diferentes idiomas y contextos lingüísticos. Los trabajos futuros podrían centrarse en mejorar el preprocesamiento y la construcción de características, explorar enfoques de transferencia de aprendizaje, evaluar el sistema en otros dominios y géneros textuales, y desarrollar técnicas de interpretabilidad de los resultados.

REFERENCIAS

- [1] Liu B. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. 2012. Morgan & Claypool Publishers: San Rafael, California. EU. doi: [10.1007/978-3-031-02145-9](https://doi.org/10.1007/978-3-031-02145-9).
- [2] Sprout Social. El índice de Sprout Social Q3 2016. 2016. Disponible en <https://sproutsocial.com/insights/data/es-3t-de-2016>.
- [3] Cervantes I. EL ESPAÑOL EN EL MUNDO 2021. Anuario del Instituto Cervantes. 2021, Instituto Cervantes: Madrid, España.
- [4] Kim S, Koh Y, Cha J, Lee S. Effects of social media on firm value for U.S. restaurant companies. International Journal of Hospitality Management. 2015, 49, 40-46. doi: [10.1016/j.ijhm.2015.05.006](https://doi.org/10.1016/j.ijhm.2015.05.006).
- [5] Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P. Natural Language Processing (Almost) from Scratch. Journal of Machine Learning Research. 2011, 12, 2493-2537.
- [6] Cruz FL, Troyano JA, Enriquez F, Ortega J. Clasificación de documentos basada en la opinión: Experimentos con un corpus de críticas de cine en español. Procesamiento del Lenguaje Natural, 2008, 41, <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/2551>.
- [7] Martínez-Cámara E, Martín-Valdivia MT, Ureña-López LA. Opinion Classification Techniques Applied to a Spanish Corpus. En Muñoz R, Montoyo A, Métails E. (eds). 16th Int. Conf. on Applications of Natural Language to Information Systems, NLDB 2011. 2011, Springer: Alicante, Spain, LNCS, 6716. doi: [10.1007/978-3-642-22327-3_17](https://doi.org/10.1007/978-3-642-22327-3_17).
- [8] Sari Y, Stevenson M, Vlachos A. Topic or Style? Exploring the Most Useful Features for Authorship Attribution. Proc. of the 27th Int. Conf. on

- Computational Linguistics, 2018. 343-353. <http://www.aclweb.org/anthology/C18-1029>.
- [9] Dubiau L, Ale, JM. Análisis de Sentimientos sobre un Corpus en Español: Experimentación con un Caso de Estudio. XIV Argentine Symposium on Artificial Intelligence (ASAI)-JAIIO. SADIO: Cordoba, Argentina, 2013. 36-47.
- [10] Estrada CA, Fabela JLT. Nuevo método para atribución de autoría en muestras de entrenamiento balanceadas y desbalanceadas del corpus C10. Research in Computing Science. 2022, 151(7), 51-62.

lo posicionan como un referente en la vanguardia académica, consolidando su contribución significativa al campo de la ingeniería electrónica y la investigación científica.

ACERCA DE LOS AUTORES



Cesar Alexis Estrada Palacios. Oriundo de Toluca, Estado de México, nació en 1994 y en 2023 obtuvo su Doctorado en la Universidad Autónoma del Estado de México. Su experiencia y especialización se centran en el campo de la inteligencia artificial, con un

enfoque destacado en el procesamiento de lenguaje natural. Haciendo uso de avanzadas técnicas como las redes neuronales y SVM, aborda y resuelve complejos problemas en esta disciplina. En la actualidad, dedica sus esfuerzos a la creación de cursos online, compartiendo su conocimiento sobre análisis de datos e inteligencia artificial con una audiencia global. Su contribución trasciende lo académico, extendiéndose a la promoción del aprendizaje colaborativo en la era digital.



José Luis Tapia Fabela. Nacido en la Ciudad de México en 1977, obtuvo su Doctorado en Ciencias en Ingeniería Electrónica por el Instituto Tecnológico de Toluca (ITT) en 2008. Actualmente es profesor e

investigador en la Universidad Autónoma del Estado de México, donde ha construido una carrera destacada. Su trabajo se centra en el modelado de descargas de alta intensidad, el cómputo aplicado a la resolución de problemas científicos, la electrónica de potencia y las aplicaciones del aprendizaje automático. Su dedicación y vasta experiencia