

Artículo:

## Clasificación automatizada de noticias para análisis estratégico FODA en tiempo real usando NLP

### Automated news classification for real-time strategic SWOT analysis using NLP

Gilberto Bojorquez-Delgado<sup>1</sup>, Jesús Bojorquez-Delgado<sup>1</sup>, Marco-Antonio-de-Jesús Rojas-Orozco<sup>1</sup>, Fernanda-Janet Limon-Lisjuan<sup>1</sup>

Revista Interdisciplinaria de  
Ingeniería Sustentable y Desarrollo  
Social (RIISDS)

<sup>1</sup> Tecnológico Nacional de México – ITS de Guasave, Sinaloa, México.

\* Autor correspondiente: [gilberto.bd@guasave.tecnm.mx](mailto:gilberto.bd@guasave.tecnm.mx)

Recibido: 20 de octubre de 2024  
Aceptado: 26 de noviembre de 2024  
Publicado: 20 de diciembre de 2024

Publicación anual editada por el  
**Instituto Tecnológico Superior de  
Tantoyuca**

Desv. Lindero Tametate, S/N  
Col. La Morita  
C.P. 92100  
Tantoyuca, Veracruz, México.  
Teléfono: 789 8931680, Ext.196.

Correo electrónico:  
[revistadigital@itsta.edu.mx](mailto:revistadigital@itsta.edu.mx)

Sitio WEB  
<https://itsta.edu.mx/revistadigital>

ISSN 2448-8003

Editor responsable:  
**Dr. Horacio Bautista Santos**

**Copyright:** Este artículo es de acceso  
abierto distribuido bajo los términos y  
condiciones de la licencia Creative  
Commons  
[https://creativecommons.org/licenses/b  
y/4.0/](https://creativecommons.org/licenses/by/4.0/)

**Resumen:** Este estudio presenta un modelo automatizado para realizar análisis FODA utilizando noticias dinámicas como fuente de datos. Empleando técnicas avanzadas de procesamiento de lenguaje natural (NLP), el modelo analiza y categoriza noticias en tiempo real, permitiendo la actualización continua del análisis estratégico. Esta capacidad de adaptación ofrece a las organizaciones una ventaja competitiva, permitiendo respuestas proactivas a cambios del mercado. El modelo sigue un enfoque sistemático que incluye tokenización, eliminación de ruido y vectorización mediante Term Frequency-Inverse Document Frequency. Se evaluaron varios algoritmos de clasificación, incluidos Naive Bayes, Support Vector Machines, Random Forest y Gradient Boosting, comparando su rendimiento con métricas como precisión y área bajo la curva ROC. Los resultados muestran que el modelo basado en SVM alcanzó una precisión cercana al 100% en la mayoría de las categorías FODA. Además, su escalabilidad asegura su aplicabilidad en entornos empresariales complejos que requieren monitoreo continuo. Este sistema automatizado mejora la eficiencia del análisis FODA tradicional y minimiza el sesgo humano. El modelo propuesto representa un avance en la automatización del análisis estratégico, proporcionando una herramienta adaptable y precisa para evaluar factores clave en diversos sectores.

**Palabras clave:** Análisis FODA, Procesamiento de Lenguaje Natural, Modelos de Clasificación.

## Resumen

Este estudio presenta un modelo automatizado para realizar análisis FODA utilizando noticias dinámicas como fuente de datos. Empleando técnicas avanzadas de procesamiento de lenguaje natural (NLP), el modelo analiza y categoriza noticias en tiempo real, permitiendo la actualización continua del análisis estratégico. Esta capacidad de adaptación ofrece a las organizaciones una ventaja competitiva, permitiendo respuestas proactivas a cambios del mercado. El modelo sigue un enfoque sistemático que incluye tokenización, eliminación de ruido y vectorización mediante Term Frequency-Inverse Document Frequency. Se evaluaron varios algoritmos de clasificación, incluidos Naive Bayes, Support Vector Machines, Random Forest y Gradient Boosting, comparando su rendimiento con métricas como precisión y área bajo la curva ROC. Los resultados muestran que el modelo basado en SVM alcanzó una precisión cercana al 100% en la mayoría de las categorías FODA. Además, su escalabilidad asegura su aplicabilidad en entornos empresariales complejos que requieren monitoreo continuo. Este sistema automatizado mejora la eficiencia del análisis FODA tradicional y minimiza el sesgo humano. El modelo propuesto representa un avance en la automatización del análisis estratégico, proporcionando una herramienta adaptable y precisa para evaluar factores clave en diversos sectores.

**Palabras clave:** Análisis FODA, Procesamiento de Lenguaje Natural, Modelos de Clasificación.

## Abstract

This study presents an automated model for conducting SWOT analysis using dynamic news as a data source. Employing advanced natural language processing (NLP) techniques, the model analyzes and categorizes news in real time, allowing for continuous updates to the strategic analysis. This adaptability provides organizations with a competitive advantage, enabling proactive responses to market changes. The model follows a systematic approach that includes tokenization, noise removal, and vectorization using Term Frequency-Inverse Document Frequency. Several classification algorithms were evaluated, including Naive Bayes, Support Vector Machines, Random Forest, and Gradient Boosting, comparing their performance with metrics such as accuracy and area under the ROC curve. The results show that the SVM-based model achieved nearly 100% accuracy in most SWOT categories. Additionally, its scalability ensures its applicability in complex business environments that require continuous monitoring. This automated system improves the efficiency of traditional SWOT analysis and minimizes human bias. The proposed model represents an advancement in the automation of strategic analysis, providing an adaptable and accurate tool for evaluating key factors across various sectors.

**Keywords:** SWOT Analysis, Natural Language Processing, Classification Models.

## Introducción

El análisis FODA, es una herramienta analítica fundamental para la planificación estratégica en organizaciones de todo tipo, facilitando la identificación y evaluación de factores internos y externos que afectan el rendimiento y la competitividad empresarial (Lu, 2010; Puyt et al., 2023). Esta metodología ha sido ampliamente adoptada en sectores como la gestión empresarial, la planificación de políticas públicas, y el desarrollo de estrategias de marketing, entre otros (Jain et al., 2022). Sin embargo, la implementación tradicional del análisis FODA suele depender de la recopilación manual de datos, lo cual no solo es un proceso intensivo en tiempo y recursos, sino también susceptible a sesgos humanos (Hayati et al., 2023). Con el auge de la digitalización y la creciente disponibilidad de datos en tiempo real, surge la necesidad de desarrollar sistemas automatizados que puedan actualizar el análisis FODA dinámicamente para reflejar cambios rápidos y complejos en el entorno empresarial (González-Cancelas et al., 2020; Greif et al., 2024; Inkinen et al., 2021).

En la era actual de la información, las noticias representan una fuente continua y rica de datos que capturan los desarrollos más recientes en múltiples áreas, incluidas las tendencias del mercado, cambios regulatorios, innovaciones tecnológicas y acontecimientos políticos. La integración de tales datos en tiempo real en un sistema automatizado para el análisis FODA ofrece una ventaja significativa, permitiendo a las organizaciones adaptarse rápidamente a nuevas oportunidades y amenazas, y optimizar sus fortalezas internas mientras mitigan debilidades potenciales (Ali et al., 2024; Houben et al., 1999). Esta capacidad de adaptación en tiempo real es especialmente crítica en sectores donde la agilidad y la toma de decisiones informadas son claves para mantener la competitividad en mercados volátiles y cambiantes (Palazzo, 2024).

El Procesamiento de Lenguaje Natural (NLP, por sus siglas en inglés) ha emergido como una disciplina clave en el desarrollo de sistemas capaces de analizar grandes volúmenes de datos textuales no estructurados (Jose Gonzalez-Gomez et al., 2024). Las técnicas de NLP permiten no solo la extracción de palabras clave y entidades relevantes, sino también la comprensión del contexto semántico y la relación entre conceptos, lo que es esencial para la clasificación precisa de la información en categorías FODA (Wu et al., 2024). Estudios recientes han demostrado que los algoritmos de aprendizaje automático, apoyados en técnicas avanzadas de NLP, pueden superar las limitaciones de los métodos manuales tradicionales al identificar patrones ocultos y relaciones complejas en los datos textuales, ofreciendo un análisis más rápido y objetivo (Killman, 2024). Además, enfoques como los word embeddings y modelos basados en transformers han mostrado una capacidad sin precedentes para

capturar matices semánticos y contextuales, lo que mejora significativamente la precisión de las tareas de clasificación textual (Rizvi et al., 2024).

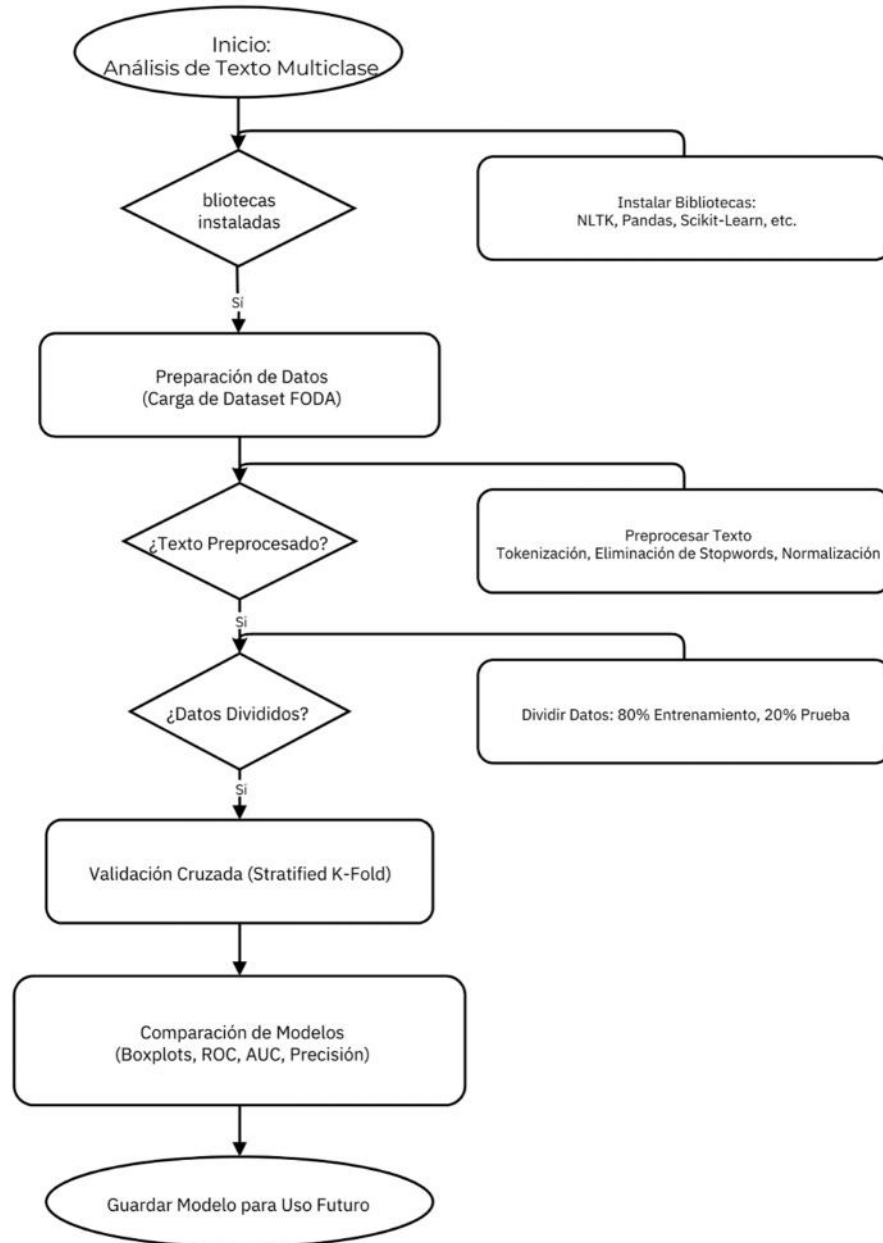
En este artículo se presenta un modelo automatizado que aplica técnicas avanzadas de NLP para la clasificación de noticias en las cuatro categorías FODA, permitiendo así la actualización continua y en tiempo real de los análisis estratégicos. El modelo propuesto sigue un enfoque sistemático que incluye etapas de preprocesamiento exhaustivo de datos, tokenización, y vectorización mediante Term Frequency-Inverse Document Frequency (TF-IDF), para transformar el texto en una representación numérica que pueda ser procesada por algoritmos de aprendizaje automático. Se evaluaron múltiples algoritmos de clasificación, incluyendo Naive Bayes, Support Vector Machines (SVM), Random Forest y Gradient Boosting, comparando su rendimiento a través de métricas robustas como precisión, sensibilidad, especificidad y área bajo la curva ROC (AUC).

Los resultados de este estudio muestran que el modelo puede adaptarse de manera continua y efectiva a la entrada dinámica de noticias, proporcionando un análisis FODA en tiempo real que refleja fielmente las condiciones cambiantes del entorno. Esto representa una mejora significativa con respecto a las metodologías tradicionales, al ofrecer una herramienta estratégica automatizada que permite a las organizaciones responder de forma proactiva y rápida a nuevos desafíos y oportunidades. Además, el sistema es escalable y puede integrarse en arquitecturas más complejas de monitoreo de información, lo que amplía su aplicabilidad práctica en diversos sectores industriales.

Este trabajo contribuye a la literatura sobre la automatización del análisis estratégico al demostrar la viabilidad y efectividad de las técnicas de NLP para la clasificación FODA basada en noticias. Además, se sugieren futuras direcciones de investigación para mejorar y optimizar el modelo, incluyendo la posible implementación de modelos transformer más avanzados que puedan ofrecer una comprensión aún más profunda del contexto y la semántica del texto. Las siguientes secciones describen en detalle la metodología empleada, los resultados obtenidos, y las implicaciones prácticas de la implementación del modelo propuesto.

## **Materiales y Métodos**

**Descripción General del Algoritmo:** El algoritmo desarrollado se basa en técnicas avanzadas de procesamiento de lenguaje natural (NLP) para procesar, analizar y clasificar noticias en categorías específicas del análisis FODA (Fortalezas, Oportunidades, Debilidades, Amenazas). Este enfoque permite la categorización automatizada y en tiempo real de noticias que son relevantes para el análisis estratégico. A continuación, se describen las etapas clave del algoritmo como se observa en la Figura 1.



**Figura 1. Diagrama de flujo del algoritmo para análisis FODA automatizado.**

**Fuente:** Elaboración propia

1. **Instalación y Configuración de Bibliotecas:** Para la implementación del algoritmo, se utilizan bibliotecas especializadas de Python, tales como NLTK, Pandas y Scikit-Learn. La biblioteca NLTK se emplea para el procesamiento y manipulación del texto, Pandas para la gestión y estructura de los datos, y Scikit-Learn para la construcción y evaluación de los modelos de

aprendizaje automático. La instalación y configuración de estas bibliotecas se realiza automáticamente al inicio del proceso, garantizando la compatibilidad y el correcto funcionamiento del algoritmo.

2. **Preparación de Datos:** El preprocesamiento de los datos es un paso crítico para asegurar que el modelo de aprendizaje automático pueda interpretar y analizar el contenido de manera eficiente. El conjunto de datos inicial se compone de ejemplos etiquetados que representan las cuatro categorías FODA. Las técnicas de preprocesamiento incluyen tokenización para segmentar el texto en unidades léxicas, eliminación de stopwords para reducir ruido lingüístico, y normalización para estandarizar el texto. Este proceso asegura que el modelo se entrene en datos coherentes y relevantes, mejorando la precisión y robustez de las predicciones.
3. **Vectorización del Texto:** Para convertir el texto en una representación numérica que pueda ser procesada por los algoritmos de aprendizaje automático, se utiliza la técnica Term Frequency-Inverse Document Frequency (TF-IDF). Esta técnica cuantifica la importancia de cada término dentro del contexto de un documento específico y su presencia en el conjunto total de documentos, lo que permite que el modelo capture patrones léxicos y semánticos significativos. La vectorización se configura con un rango de n-gramas (1,2) para capturar tanto palabras individuales como combinaciones de dos términos consecutivos, enriqueciendo así la representación del texto.
4. **Entrenamiento y Validación:** Se emplean múltiples algoritmos de clasificación para evaluar el rendimiento en la categorización de noticias: Naive Bayes, Support Vector Machines (SVM), Random Forest y Gradient Boosting. Cada modelo es entrenado utilizando una técnica de validación cruzada estratificada (Stratified K-Fold) con cinco pliegues, que asegura que cada clase esté representada equitativamente en los subconjuntos de entrenamiento y prueba. Este enfoque permite obtener una evaluación más precisa de la capacidad de generalización del modelo y minimiza el riesgo de sobreajuste.
5. **Selección del Mejor Modelo:** Tras la evaluación de los diferentes modelos de clasificación, se selecciona el modelo con el mejor rendimiento basado en métricas clave como la precisión promedio (accuracy), la curva ROC-AUC y las curvas de Precisión-Recall. El modelo seleccionado se entrena utilizando todo el conjunto de datos para optimizar su capacidad predictiva. Esta metodología permite identificar el algoritmo que mejor se adapta a las características del conjunto de datos y maximiza la eficacia en la clasificación de noticias en tiempo real.

6. **Automatización de la Entrada de Noticias:** El modelo final se integra con fuentes de noticias en tiempo real mediante un flujo de trabajo automatizado que permite la entrada continua de datos. Las noticias nuevas se preprocesan y vectorizan automáticamente, y el modelo las clasifica en una de las cuatro categorías FODA. Este diseño garantiza que el análisis estratégico esté siempre actualizado, reflejando los cambios y tendencias recientes en el entorno de negocio.

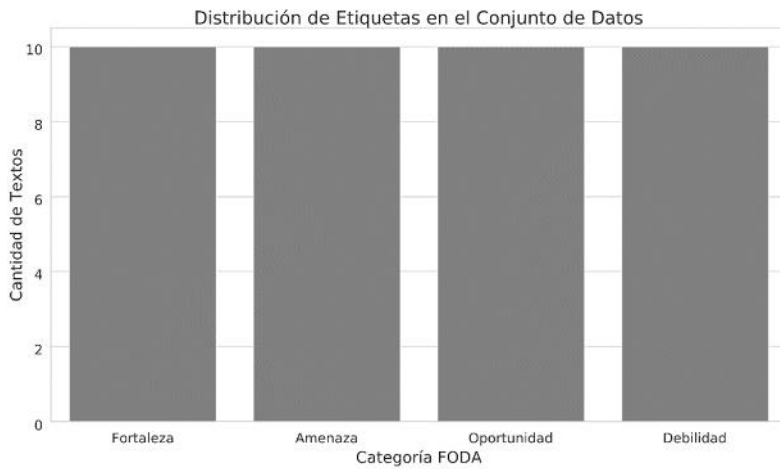
## Resultados y Discusión

En esta sección se presentan los resultados obtenidos tras la implementación del modelo automatizado para el análisis FODA basado en noticias. Se emplearon diversas métricas y visualizaciones para evaluar el rendimiento del modelo, incluyendo la distribución de etiquetas en el conjunto de datos, la precisión de clasificación mediante la matriz de confusión, el análisis de curvas ROC y Precision-Recall, así como curvas de aprendizaje para analizar el comportamiento del modelo durante el entrenamiento.

Los resultados obtenidos reflejan un alto rendimiento del modelo seleccionado, el cual fue capaz de categorizar eficientemente las noticias en las cuatro clases FODA: Fortalezas, Oportunidades, Debilidades y Amenazas. Las siguientes subsecciones proporcionan un análisis detallado de estos resultados, comenzando con la distribución de etiquetas y seguida por una evaluación integral del rendimiento del modelo basado en métricas clave.

**Distribución de Etiquetas en el Conjunto de Datos:** La Figura 2 presenta la distribución de etiquetas en el conjunto de datos utilizado para entrenar los modelos. Se observa una distribución equilibrada entre las cuatro categorías FODA: Fortalezas, Oportunidades, Debilidades y Amenazas, con una cantidad homogénea de ejemplos para cada clase. Esta homogeneidad es fundamental para asegurar que el modelo pueda aprender de manera equitativa las características que distinguen a cada categoría, evitando sesgos que podrían afectar negativamente el rendimiento durante la clasificación.

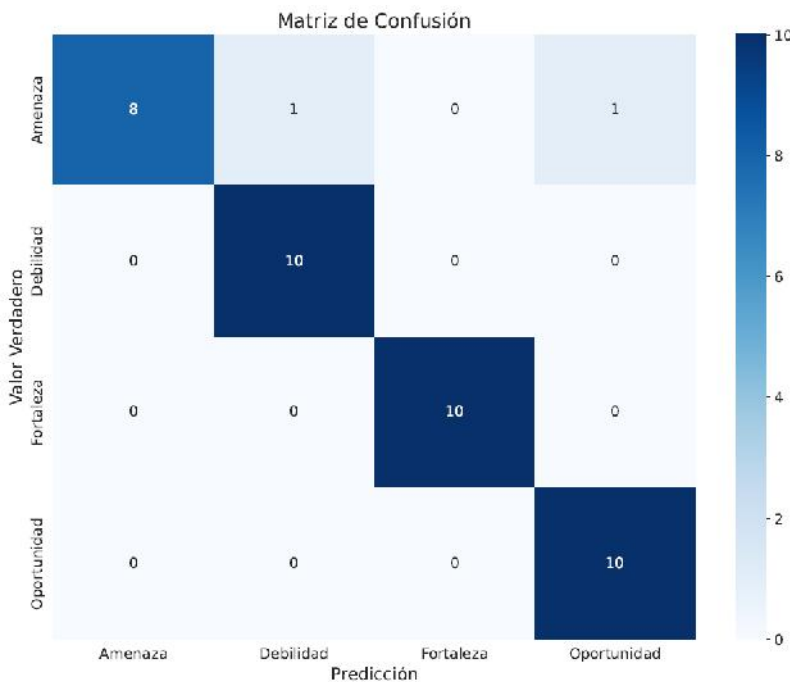
La distribución equilibrada también permite que el modelo tenga un aprendizaje representativo de los datos, lo que facilita una mejor generalización al clasificar nuevas noticias. Además, esta distribución uniforme asegura que las métricas de rendimiento obtenidas no se vean afectadas por desequilibrios en la cantidad de datos por clase, proporcionando una evaluación justa del desempeño del modelo.



**Figura 2. Distribución de Etiquetas en el Conjunto de Datos.**

**Fuente:** Elaboración propia

**Rendimiento del Modelo - Matriz de Confusión:** La precisión del modelo seleccionado se evaluó mediante la matriz de confusión mostrada en la Figura 3. Esta matriz ofrece una representación visual de los aciertos y errores en la clasificación, donde cada celda muestra el número de predicciones correctas e incorrectas para cada clase. Los resultados indican que el modelo logró una alta precisión, clasificando correctamente la mayoría de las noticias en las categorías correspondientes, con mínimos errores de clasificación.



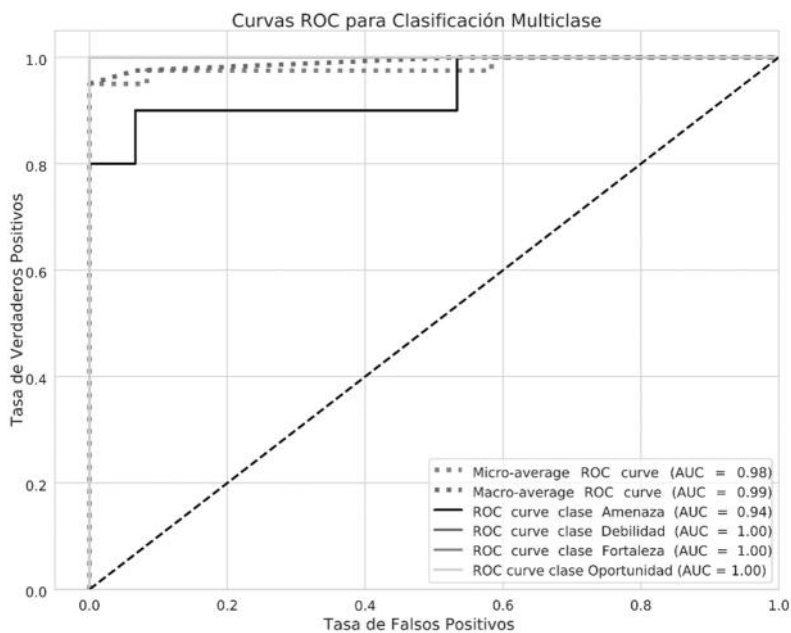
**Figura 3. Matriz de Confusión para la Clasificación FODA.**

**Fuente:** Elaboración propia

En detalle, se observa que las clases Fortaleza, Oportunidad y Debilidad no presentaron errores de clasificación, lo que refleja una alta sensibilidad y especificidad en el reconocimiento de las características distintivas de estas categorías. La clase Amenaza mostró algunos errores menores, donde algunas instancias fueron incorrectamente clasificadas como Debilidad. Este pequeño margen de error podría atribuirse a similitudes léxicas o contextuales entre ejemplos de ambas categorías, sugiriendo posibles mejoras en el preprocesamiento o la técnica de vectorización para discriminar mejor estas instancias

**Curvas ROC y AUC para Clasificación Multiclase:** Las curvas ROC presentadas en la Figura 4 demuestran la capacidad del modelo para distinguir entre las diferentes clases. Cada curva representa la tasa de verdaderos positivos frente a la tasa de falsos positivos para cada categoría, y el área bajo la curva (AUC) se utiliza como una métrica para evaluar el rendimiento general del modelo.

Los resultados muestran que las curvas ROC para las clases Debilidad, Fortaleza y Oportunidad alcanzaron un AUC de 1.00, indicando una capacidad perfecta para discriminar estas clases. La clase Amenaza obtuvo un AUC de 0.94, lo que sigue siendo un excelente desempeño, aunque muestra una ligera disminución en la capacidad discriminativa en comparación con las otras clases. La curva promedio (micro y macro) muestra valores AUC cercanos a 1.0, lo que refleja la robustez y precisión del modelo en una clasificación multicategoría efectiva.

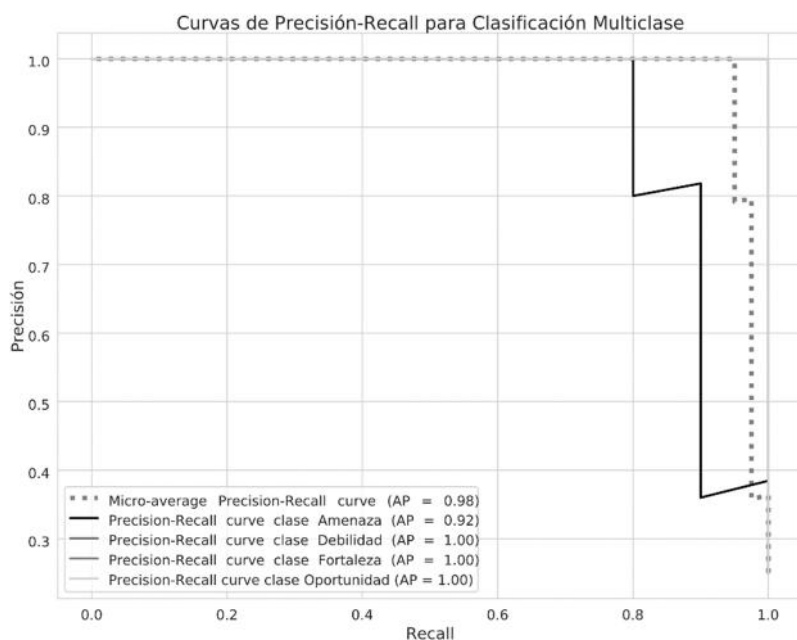


**Figura 4. Curvas ROC para Clasificación Multiclase.**

**Fuente: Elaboración propia**

**Curvas de Precisión-Recall:** La Figura 5 presenta las curvas de Precisión-Recall, que ofrecen una visión complementaria al análisis de las curvas ROC. Estas curvas destacan la capacidad del modelo para manejar casos positivos en cada clase, especialmente en contextos donde las clases están desbalanceadas o cuando el enfoque principal es minimizar los falsos positivos.

Las curvas de Precisión-Recall indican que el modelo mantiene una alta precisión promedio (AP) para todas las categorías, con un AP de 1.00 para las clases Debilidad, Fortaleza y Oportunidad, lo que sugiere una precisión perfecta al clasificar estas instancias. La clase Amenaza obtuvo un AP de 0.92, lo que sigue siendo muy alto, pero indica una ligera variabilidad en la capacidad del modelo para manejar algunas instancias de esta categoría. La alta precisión en las clases dominantes subraya la eficacia del modelo para detectar características claves que definen cada categoría FODA.



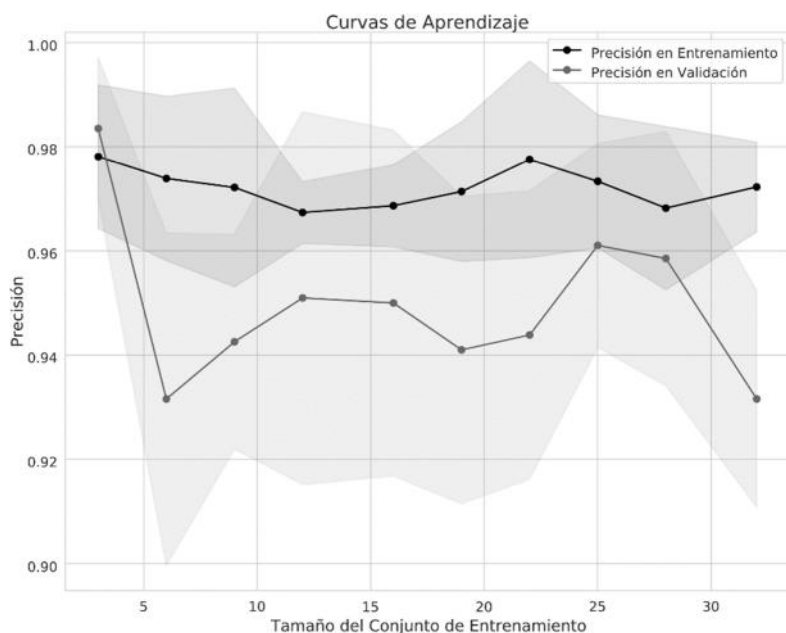
**Figura 5. Curvas de Precisión-Recall para Clasificación Multiclase.**

**Fuente: Elaboración propia**

**Curvas de Aprendizaje:** La Figura 6 presenta las curvas de aprendizaje del modelo, las cuales ilustran cómo varía la precisión del modelo en función del tamaño del conjunto de entrenamiento. Este análisis es fundamental para evaluar la capacidad del modelo para generalizar en distintos tamaños de datos y para identificar posibles problemas de sobreajuste o subajuste.

1. **Tendencia de la Precisión:** Las curvas de aprendizaje muestran una alta precisión tanto en el conjunto de entrenamiento como en el de validación. Esta alta precisión sugiere que el modelo es capaz de capturar eficazmente los patrones presentes en los datos sin incurrir en sobreajuste.
2. **Estabilización de la Precisión:** A medida que aumenta el tamaño del conjunto de entrenamiento, la precisión del modelo se estabiliza. Esta estabilización indica que el modelo ha alcanzado un punto de convergencia, donde añadir más datos no resulta en mejoras significativas en el rendimiento. Esto es un indicativo de que el modelo está aprendiendo patrones representativos y generales de los datos, manteniendo su capacidad de generalización.
3. **Brecha Entre Entrenamiento y Validación:** La brecha mínima entre las curvas de entrenamiento y validación indica que el modelo mantiene una consistencia en su desempeño al enfrentarse a datos nuevos. Esta consistencia es crucial para aplicaciones en tiempo real, donde el modelo debe adaptarse continuamente a nuevos datos sin deteriorar su rendimiento.

Las curvas de aprendizaje presentadas demuestran que el modelo seleccionado generaliza bien a diferentes tamaños de conjuntos de datos, manteniendo una alta precisión sin evidenciar signos de sobreajuste. Este comportamiento es especialmente beneficioso para aplicaciones prácticas donde el modelo debe adaptarse a nuevas informaciones de manera eficiente y eficaz.



**Figura 6. Curvas de Aprendizaje.**

**Fuente: Elaboración propia**

## Conclusiones

El presente estudio ha demostrado la efectividad de un modelo automatizado basado en técnicas de procesamiento de lenguaje natural (NLP) para realizar análisis FODA utilizando la entrada dinámica de noticias. A través de una metodología rigurosa que incluyó el preprocesamiento exhaustivo de datos, la vectorización avanzada mediante TF-IDF, y la evaluación de múltiples algoritmos de clasificación, se logró desarrollar un sistema robusto capaz de categorizar noticias en las cuatro clases clave de un análisis FODA: Fortalezas, Oportunidades, Debilidades y Amenazas.

Los resultados obtenidos muestran que el modelo seleccionado, basado en Support Vector Machines (SVM), superó a otros enfoques en términos de precisión, especificidad y sensibilidad. Las métricas cuantitativas derivadas de la matriz de confusión, las curvas ROC-AUC, y las curvas Precision-Recall evidencian que el modelo alcanza una precisión cercana al 100% para la mayoría de las clases, lo que subraya su capacidad para identificar correctamente patrones léxicos y semánticos en el texto. Estos resultados reflejan la robustez del sistema para discriminar entre categorías incluso cuando las diferencias contextuales pueden ser sutiles, asegurando una clasificación precisa y coherente en distintos escenarios.

Uno de los aspectos más destacados del sistema es su capacidad de actualización en tiempo real. La integración del modelo con fuentes de noticias dinámicas permite que el análisis FODA se adapte automáticamente a los cambios en el entorno, proporcionando así una herramienta estratégica valiosa para organizaciones que buscan mantener una ventaja competitiva en mercados volátiles. Este enfoque proactivo para la gestión de información asegura que las decisiones estratégicas se basen siempre en datos actuales y relevantes, aumentando la capacidad de respuesta y adaptación a nuevos desafíos y oportunidades.

Además, el análisis de las curvas de aprendizaje reveló que el modelo es escalable y mantiene su precisión incluso cuando se expande el tamaño del conjunto de datos de entrenamiento.

A pesar del éxito del modelo, se identificaron áreas para futuras mejoras. Por ejemplo, la clase Amenaza mostró ligeras variaciones en precisión en comparación con las demás, lo que podría atribuirse a la complejidad contextual de las noticias que describen amenazas. En el futuro, se podría explorar el uso de técnicas más avanzadas de vectorización, como word embeddings o modelos preentrenados

basados en transformers (por ejemplo, BERT), para capturar mejor el contexto semántico y mejorar la clasificación en esta categoría específica.

El modelo propuesto ofrece una solución efectiva y adaptable para realizar análisis FODA automatizados basados en noticias. Su capacidad para actualizarse en tiempo real, junto con su alto rendimiento en la clasificación precisa de noticias en categorías estratégicas, lo convierte en una herramienta poderosa para empresas que desean realizar un monitoreo continuo de su entorno competitivo. Las mejoras sugeridas para el futuro pueden ampliar aún más las capacidades del sistema, permitiendo un análisis más sofisticado y contextualizado de las dinámicas del mercado

### Referencias Bibliográficas

- Ali, M. Y., Naeem, S. B., Bhatti, R., & Richardson, J. (2024). Artificial intelligence application in university libraries of Pakistan: SWOT analysis and implications. *Global Knowledge, Memory and Communication*, 73(1/2), 219-234. <https://doi.org/10.1108/GKMC-12-2021-0203>
- González-Cancelas, N., Molina Serrano, B., Soler-Flores, F., & Camarero-Orive, A. (2020). Using the SWOT Methodology to Know the Scope of the Digitalization of the Spanish Ports. *Logistics*, 4(3), 20. <https://doi.org/10.3390/logistics4030020>
- Greif, L., Kimmig, A., El Bobbou, S., Jurisch, P., & Ovtcharova, J. (2024). Strategic view on the current role of AI in advancing environmental sustainability: A SWOT analysis. *Discover Artificial Intelligence*, 4(1), 45. <https://doi.org/10.1007/s44163-024-00146-z>
- Hayati, M., Mahdevari, S., & Barani, K. (2023). An improved MADM-based SWOT analysis for strategic planning in dimension stones industry. *Resources Policy*, 80, 103287. <https://doi.org/10.1016/j.resourpol.2022.103287>
- Houben, G., Lenie, K., & Vanhoof, K. (1999). A knowledge-based SWOT-analysis system as an instrument for strategic planning in small and medium sized enterprises. *Decision Support Systems*, 26(2), 125-135. [https://doi.org/10.1016/S0167-9236\(99\)00024-X](https://doi.org/10.1016/S0167-9236(99)00024-X)

- Inkinen, T., Helminen, R., & Saarikoski, J. (2021). Technological trajectories and scenarios in seaport digitalization. *Research in Transportation Business & Management*, 41, 100633.  
<https://doi.org/10.1016/j.rtbm.2021.100633>
- Jain, V., Ajmera, P., & Davim, J. P. (2022). SWOT analysis of Industry 4.0 variables using AHP methodology and structural equation modelling. *Benchmarking: An International Journal*, 29(7), 2147-2176. <https://doi.org/10.1108/BIJ-10-2020-0546>
- Jose Gonzalez-Gomez, L., Margarita Hernandez-Munoz, S., Borja, A., Daniel Azofeifa, J., Noguez, J., & Caratozzolo, P. (2024). Analyzing Natural Language Processing Techniques to Extract Meaningful Information on Skills Acquisition From Textual Content. *IEEE Access*, 12, 139742-139757. <https://doi.org/10.1109/ACCESS.2024.3465409>
- Killman, J. (2024). Machine translation literacy in the legal translation context: A SWOT analysis perspective. *The Interpreter and Translator Trainer*, 18(2), 271-289.  
<https://doi.org/10.1080/1750399X.2024.2344282>
- Lu, W. (2010). Improved SWOT Approach for Conducting Strategic Planning in the Construction Industry. *Journal of Construction Engineering and Management*, 136(12), 1317-1328.  
[https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000240](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000240)
- Palazzo, M. (2024). The SWOT Analysis: An Evolving Decision-Making Model. En M. Palazzo & A. Micozzi, *Rethinking Decision-Making Strategies and Tools: Emerging Research and Opportunities* (pp. 53-70). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-83797-204-320241004>
- Puyt, R. W., Lie, F. B., & Wilderom, C. P. M. (2023). The origins of SWOT analysis. *Long Range Planning*, 56(3), 102304. <https://doi.org/10.1016/j.lrp.2023.102304>

Rizvi, S. M. H., Imran, R., & Mahmood, A. (2024). *Text Classification using Graph Convolutional Networks: A Comprehensive Survey* (Versión 1). arXiv.

<https://doi.org/10.48550/ARXIV.2410.09399>

Wu, J., Zhao, N., & Yang, T. (2024). Wisdom of crowds: SWOT analysis based on hybrid text mining methods using online reviews. *Journal of Business Research*, 171, 114378.

<https://doi.org/10.1016/j.jbusres.2023.114378>