

# Descifrando el lenguaje emocional en Twitter: Un análisis predictivo basado en aprendizaje automático

Neivys Luz González Gómez

**Abstracto**— El análisis de sentimientos es crucial para predecir las tendencias y preferencias de los usuarios en diversos ámbitos de la vida cotidiana. En este estudio se evaluaron las técnicas más comunes para el análisis de sentimientos, tanto de aprendizaje automático como de aprendizaje profundo, mediante una estrategia metodológica que cubrió el pre-procesamiento de datos, la construcción de modelos predictivos y su evaluación. Se evaluaron dos modelos de aprendizaje automático, SGDClassifier y BERT, en la tarea de clasificación de sentimientos en diferentes conjuntos de datos etiquetados. Los resultados indican que el modelo BERT supera al SGDClassifier en todos los conjuntos de datos en términos de métricas de rendimiento de clasificación de sentimientos. En particular, el modelo BERT mostró un mejor rendimiento en el conjunto de datos etiquetados con sentimientos, alcanzando valores de *f1*, precisión y recall superiores al 60%. Además, se encontró que el balanceo de los datos puede ser una estrategia efectiva para mejorar el rendimiento del modelo en la clasificación de sentimientos con polaridad de Vader. Sin embargo, se observó que el modelo BERT aún tiene dificultades en la detección de ciertas emociones, lo que sugiere la necesidad de investigaciones adicionales para mejorar su rendimiento en estas áreas. Estos hallazgos tienen implicaciones importantes para el desarrollo de herramientas de análisis de sentimientos más precisas y efectivas en diversas áreas, como la mercadotecnia, la política y la psicología.

**Palabras claves** — análisis de sentimientos, machine learning, aprendizaje profundo, SGDClassifier, BERT, clasificación de sentimientos, polaridad de Vader, detección de emociones.

## I. INTRODUCTION

En la actualidad, Twitter se ha convertido en una plataforma clave para la interacción social y la difusión de noticias y opiniones. Además de ser una herramienta útil para la comunicación, las redes sociales como Twitter son una fuente rica de información que puede ser utilizada para analizar y comprender el comportamiento humano y la forma en que se expresan las emociones. El análisis emocional en las redes sociales ha crecido rápidamente en los últimos años gracias a la tecnología predictiva y los algoritmos de aprendizaje automático que mejoran constantemente en su capacidad para identificar patrones en el lenguaje emocional de los usuarios de Twitter.

Las emociones tienen un impacto significativo en el comportamiento de los usuarios en línea, y los algoritmos de aprendizaje automático son capaces de identificar patrones en

el lenguaje que indican emociones como la felicidad, la tristeza o el enfado. Con la enorme cantidad de datos disponibles en Twitter, estos algoritmos pueden aprender y mejorar constantemente, lo que ha revolucionado la forma en que entendemos las redes sociales y ha abierto un mundo de posibilidades para los expertos en marketing, analistas de datos y otras organizaciones relacionadas con los negocios, la psicología, la atención médica, la política, la seguridad y otros campos.

Los modelos de clasificación inteligentes han demostrado su capacidad de predicción de sentimientos en textos, para determinar la percepción de los usuarios sobre aspectos de la vida cotidiana. La información extraída por el análisis de sentimientos se puede usar como conocimiento para análisis posteriores, y en general, se quiere predecir los resultados de preferencias o tendencias de un tópico particular a partir del sentimiento.

En el análisis de emociones en textos cortos, como los tweets, es necesario utilizar técnicas adecuadas para pre-procesar los datos, ya que estos pueden contener errores tipográficos, abreviaturas y lenguaje coloquial. Para poder evaluar las técnicas más comunes para detectar sentimientos en tweets, se debe realizar un análisis exhaustivo y riguroso, utilizando una gran cantidad de datos etiquetados en diferentes emociones.

El objetivo principal de este artículo es evaluar diversas técnicas de modelos inteligentes de Machine Learning y Deep Learning en un conjunto de datos de 40.000 registros de tweets etiquetados en 13 emociones diferentes. Los modelos que se evaluarán en este artículo son algoritmos de aprendizaje automático (Machine Learning) y de aprendizaje profundo (Deep Learning) que se han utilizado ampliamente en la detección de sentimientos en textos cortos, como los tweets.

El algoritmo de Máquinas de Vectores de Soporte (SVC) es una técnica de clasificación que separa los datos en diferentes categorías utilizando un hiperplano. El Clasificador de Descenso de Gradiente Estocástico (SGDClassifier) es un algoritmo de aprendizaje automático utilizado para clasificación binaria y multiclase. Los clasificadores Naive Bayes, MultinomialNB y BernoulliNB, son métodos estadísticos basados en el Teorema de Bayes para la clasificación de textos. El modelo de Gradient Boosting Extreme (XGBClassifier) es un algoritmo de aprendizaje

automático basado en árboles de decisión, que utiliza un conjunto de modelos débiles para construir un modelo fuerte. Finalmente, el modelo pre-entrenado de BERT es un modelo de lenguaje basado en redes neuronales transformadoras (transformer neural networks) que se ha demostrado que funciona muy bien en la detección de sentimientos en textos.

Para lograr este objetivo, se llevará a cabo un proceso de pre-procesamiento de datos adecuado para los tweets, que incluirá la eliminación de caracteres especiales, signos de puntuación y stop words, la lematización y la eliminación de palabras irrelevantes. A continuación, se aplicarán técnicas de vectorización para convertir los tweets en representaciones numéricas que puedan ser utilizadas por los modelos de aprendizaje automático y profundo. Una vez preprocesados los datos, se llevará a cabo la evaluación de los modelos utilizando una variedad de métricas, como la precisión, la recall, el F1-score y la matriz de confusión.

En conclusión, este artículo aborda una problemática actual en el campo del análisis emocional en textos cortos, como los tweets, y tiene como objetivo evaluar las técnicas más comunes de Machine Learning y Deep Learning en un conjunto de datos etiquetados en diferentes emociones. Los resultados de este estudio podrán ser utilizados por profesionales en áreas como el la psicología, marketing, y la investigación de opinión pública para obtener información valiosa sobre el comportamiento y las emociones de los usuarios en Twitter o en otro tipo de red social.

## II. ESTADO DEL ARTE

El análisis de emociones en textos cortos, como los tweets, ha cobrado una gran importancia en los últimos años debido a la gran cantidad de información emocional que estos pueden proporcionar. Diversos estudios han evaluado diferentes técnicas de modelos de inteligencia artificial (IA) para la detección y análisis de emociones en tweets, destacando los algoritmos de aprendizaje automático (Machine Learning) y de aprendizaje profundo (Deep Learning).

En la literatura existente, se pueden identificar dos categorías principales: la primera basada en métodos tradicionales de Machine Learning y la segunda en métodos de Deep Learning. En la primera categoría, se han utilizado diversos algoritmos como el de Máquinas de Vectores de Soporte (SVC), el Clasificador de Descenso de Gradiente Estocástico (SGDClassifier), los clasificadores Naive Bayes, MultinomialNB y BernoulliNB, el modelo de Gradient Boosting Extreme (XGBClassifier) y el modelo pre-entrenado de BERT para clasificar emociones en tweets. Estudios como los de Asgas et al. (2018), Suhasini et al. (2019), Bhagat et al. (2019) y Gaind et al. (2018) han evaluado el rendimiento de estos algoritmos y han encontrado que algunos, como la regresión logística y el clasificador de Máquinas de Vectores de Soporte, pueden tener un desempeño superior en la detección de emociones en tweets.

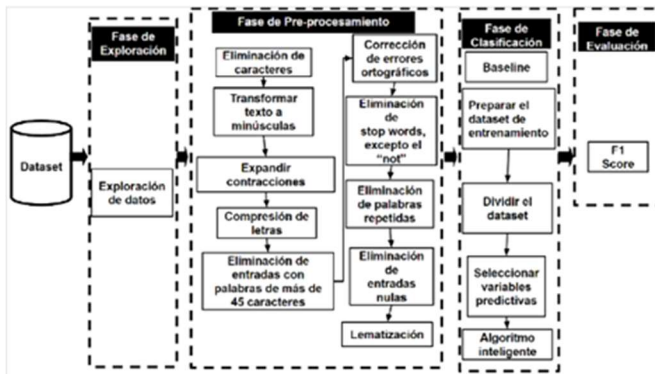
En la segunda categoría, basada en Deep Learning, se han utilizado modelos como long short-term memory (LSTM) y Nested LSTM para clasificar textos en varias emociones. Haryadi et al. (2019) encontraron que el mejor desempeño se logró con Nested LSTM, mientras que Munikar et al. (2020) lograron la mejor precisión utilizando el modelo pre-entrenado de BERT y afinándolo para la clasificación de análisis de sentimientos de granulado fino. Huang et al. (2021) propusieron un modelo híbrido de Deep Learning y Machine Learning para la detección de emociones en tweets, combinando la capacidad de aprendizaje profundo de BERT con la eficiencia de Naive Bayes.

En cuanto a las estrategias de pre-procesamiento, en Angiani et al. (2016) se especificaron diferentes filtros para limpiar los datos originales (en este caso tweets) y determinar cuál combinación de filtros produce mejores resultados en el aprendizaje. También se utilizan técnicas de stemming para evitar variaciones de palabras, como la eliminación de URLs, palabras de detención (stop words) y números, que afectan mínimamente el rendimiento de los clasificadores, como concluyen Jianqiang & Xiaolin (2017).

En resumen, la detección y análisis de emociones en tweets se ha estudiado con diversas técnicas y modelos de IA, y se han identificado algunos algoritmos y modelos que pueden mejorar el rendimiento en esta tarea. Además, las estrategias de pre-procesamiento, como la eliminación de URLs, palabras de detención y números, han demostrado ser efectivas en mejorar la precisión de los clasificadores. Para este estudio, se llevará a cabo una comparación entre un modelo de machine learning y el modelo de BERT, que se ha destacado como una de las técnicas más prometedoras en el procesamiento del lenguaje natural.

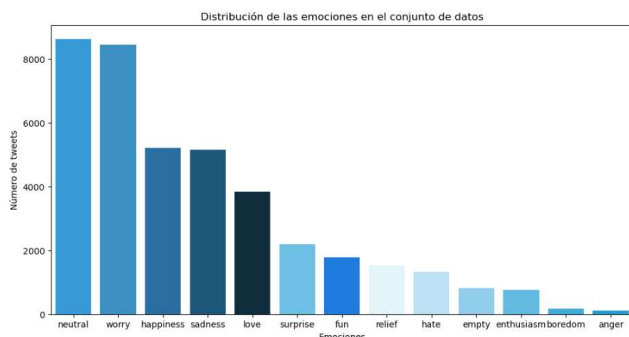
## III. METODOLOGIA

Se propone una metodología para comparar diversas técnicas de aprendizaje, la cual abarca todas las fases del proceso desde la extracción del conjunto de datos hasta la evaluación comparativa. En la Figura 1, se muestra el esquema general de nuestra estrategia metodológica, sus fases y las actividades que se llevan a cabo en cada una. Las dos primeras fases utilizan técnicas de Procesamiento de Lenguaje Natural (NLP) para normalizar el texto, lo cual garantiza una representación adecuada para los algoritmos de aprendizaje.



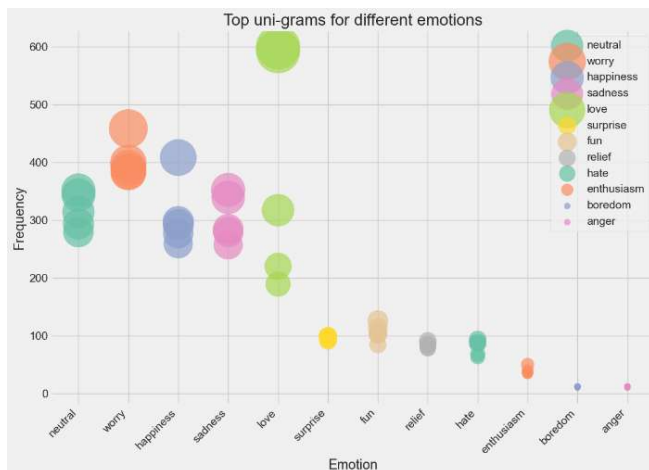
### Figura 1. Flujo de trabajo

**En la Etapa 1**, se realiza la exploración de datos, que es el primer paso para comprender el conjunto de datos utilizado. Se aplican diversas técnicas y visualizaciones para obtener información relevante sobre los datos, como la distribución de las palabras utilizadas, el número de tweets positivos, negativos y neutros, y la diversidad y equilibrio del conjunto de datos.



**Figura 2. Distribución de las emociones. Dataset base**

Esta etapa es esencial para comprender los datos en profundidad, y aunque puede llevar bastante tiempo, es crucial para obtener una visión general del conjunto de datos y para identificar posibles problemas o sesgos que puedan afectar la calidad de los resultados. La figura 3 muestra un ejemplo de visualización que se puede utilizar en esta fase para obtener información sobre la distribución de las palabras utilizadas.



**Figura 3. Uni-grams de las diferentes emociones**

**Las fases de pre-procesamiento y clasificación** son cruciales para el análisis de sentimientos en texto no estructurado. Una vez que se ha realizado la limpieza del conjunto de datos, se han generado 3 conjuntos de datos para su análisis, clasificando las emociones en: emociones catalogadas del conjunto de datos original, sentimientos negativos, positivos y neutros producto de la transformación de las emociones en sentimientos y análisis de sentimientos aplicando la librería de Vader. Mejorar estas fases puede ayudar a obtener resultados más precisos y confiables. En la fase de pre-procesamiento, es importante eliminar los caracteres innecesarios, convertir el texto a minúsculas, eliminar entradas nulas y lematizar el texto. Estas acciones pueden ayudar a reducir el ruido y mejorar la uniformidad del texto. En la fase de clasificación, se debe establecer una línea base para identificar el peor rendimiento que puede obtener un modelo y compararlo con modelos más avanzados. Además, es necesario transformar el conjunto de datos de entrenamiento y dividir el conjunto de datos en conjuntos de entrenamiento y prueba para obtener una estimación precisa del error. Para la transformación del texto a vectores numéricos, se pueden utilizar diferentes técnicas, como CountVectorizer y TfidfVectorizer. En el caso de modelos de deep learning pre-entrenados como BERT, se requiere una tokenización específica del modelo para representar el texto como vectores numéricos. Mejorar estas fases puede ayudar a obtener mejores resultados en el análisis de sentimientos en texto no estructurado.

La división del conjunto de datos es una etapa crítica para garantizar la validez del modelo de aprendizaje. Para lograr esto, es necesario separar los datos de entrenamiento de los de prueba. Por lo general, se destina el 70-75% del conjunto de datos para el entrenamiento, mientras que el 25-30% restante se convierte en un subconjunto independiente destinado a las pruebas. Aunque la división del conjunto puede ser arbitraria, se recomienda que el subconjunto de entrenamiento sea mayor que el subconjunto de pruebas, ya que esto permite mejorar el modelo de clasificación y obtener una estimación más precisa del error (Dobbin & Simon, 2011). En el caso del modelo pre-entrenado de BERT, se utiliza una porción del conjunto de entrenamiento (15%) para validar el modelo y garantizar su rendimiento.

**Selección de los algoritmos a evaluar:** Para obtener el mejor rendimiento de un algoritmo de aprendizaje automático, es necesario ajustar ciertos parámetros clave, como la cantidad de neuronas en una red neuronal. Es importante tener en cuenta que no todos los algoritmos tienen la misma flexibilidad en términos de hiperparámetros. En el estudio, utilizamos una combinación de algoritmos clásicos de aprendizaje automático para la clasificación, así como el modelo de deep learning BERT para lograr un mejor desempeño en la detección de emociones en los tweets. Para cada algoritmo, se realizaron pruebas exhaustivas para encontrar los parámetros óptimos

que maximizaran el rendimiento del modelo.

El algoritmo de **Máquinas de Vectores de Soporte (SVC)** es una técnica de aprendizaje supervisado utilizada para clasificación y regresión. En el caso de clasificación, su objetivo es separar los datos en diferentes categorías utilizando un hiperplano. El SVC encuentra el hiperplano que maximiza la distancia entre las dos clases más cercanas y se puede usar con diferentes kernels para adaptarse a diferentes tipos de datos. Es un algoritmo muy popular y eficiente para conjuntos de datos de tamaño pequeño y mediano.

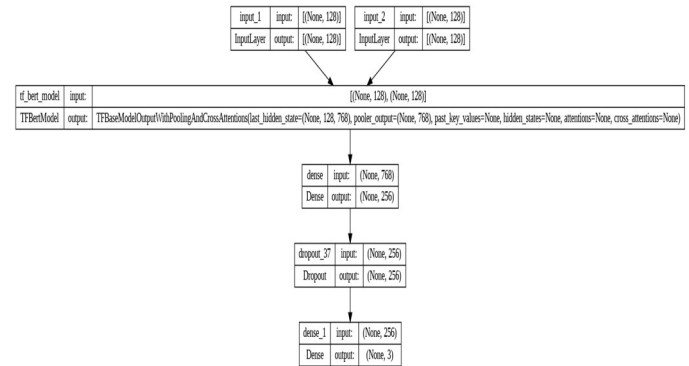
**El Clasificador de Descenso de Gradiente Estocástico (SGDClassifier)** es un algoritmo de aprendizaje automático utilizado para clasificación binaria y multiclase. Utiliza el descenso de gradiente estocástico para optimizar los pesos del modelo y ajustarlos a los datos. Es un algoritmo rápido y escalable, lo que lo hace ideal para conjuntos de datos grandes. También tiene la ventaja de ser fácil de implementar y de ajustar sus parámetros. El Clasificador de Descenso de Gradiente Estocástico (SGDClassifier) es un algoritmo de aprendizaje automático que se utilizó para clasificar las emociones de los tweets debido a su excelente rendimiento y tiempos de respuesta. Utiliza el descenso de gradiente estocástico para optimizar los pesos del modelo y ajustarlos a los datos. Esto hace que sea rápido y escalable, lo que lo hace ideal para conjuntos de datos grandes. Además, tiene la ventaja de ser fácil de implementar y ajustar sus parámetros, lo que lo hace una excelente opción para proyectos de aprendizaje automático.

**Los clasificadores Naive Bayes, MultinomialNB y BernoulliNB**, son métodos estadísticos basados en el Teorema de Bayes para la clasificación de textos. Estos algoritmos asumen que las características son independientes entre sí, lo que hace que el cálculo de la probabilidad sea más simple. El clasificador Naive Bayes se utiliza comúnmente en problemas de clasificación de texto, como la detección de spam o la clasificación de documentos.

**El modelo de Gradient Boosting Extreme (XGBClassifier)** es un algoritmo de aprendizaje automático basado en árboles de decisión, que utiliza un conjunto de modelos débiles para construir un modelo fuerte. Es un algoritmo de clasificación altamente efectivo que utiliza el boosting para mejorar el rendimiento del modelo. XGBClassifier tiene la capacidad de trabajar con diferentes tipos de datos, incluyendo datos categóricos y numéricos, y es adecuado para grandes conjuntos de datos.

Finalmente, **el modelo pre-entrenado de BERT** es un modelo de lenguaje basado en redes neuronales transformadoras (transformer neural networks) que se ha demostrado que funciona muy bien en la detección de sentimientos en textos. BERT (Bidirectional Encoder Representation from Transformers) es un modelo de lenguaje pre-entrenado que se ha entrenado en grandes conjuntos de

datos de texto para aprender representaciones de palabras y frases en un corpus. Luego, se puede utilizar este modelo pre-entrenado en tareas de clasificación de texto, como la detección de sentimientos, con una fina sintonización de los pesos del modelo en un conjunto de datos específico. En el caso de los modelos de deep learning, BERT se utilizó como modelo de referencia debido a su capacidad para aprender representaciones semánticas complejas en texto, lo que lo hace ideal para tareas de procesamiento de lenguaje natural como la detección de sentimientos. La figura 4, muestra la arquitectura del modelo pre-entrenado de BERT y cómo se utiliza para el aprendizaje de lenguaje natural en dos fases: pre-entrenamiento y ajuste fino.



**Figura 4. Arquitectura del modelo pre-entrenado de BERT**

**Etapas 4:** En esta etapa, es crucial seleccionar una métrica adecuada para evaluar el rendimiento de cada algoritmo y comparar su eficacia en la tarea que se quiere resolver. Cabe destacar que no todas las métricas son relevantes para todos los problemas de Machine Learning, por lo que es importante elegir la métrica adecuada en función del problema. La Matriz de Confusión es una herramienta valiosa para evaluar el desempeño de los algoritmos de clasificación en el aprendizaje supervisado. Esta matriz proporciona varias métricas útiles, como la precisión, que mide qué tan a menudo un modelo predice una clase como positiva y acierta. La precisión es especialmente útil cuando los falsos positivos son costosos. Para comparar la eficiencia de los algoritmos de clasificación en términos de precisión, se puede utilizar la métrica F1, que representa la relación entre la precisión y la exhaustividad (recall). La métrica F1 es especialmente útil cuando tanto los falsos positivos como los falsos negativos son importantes. La medida de exhaustividad o recall, por su parte, es valiosa cuando el costo de los falsos negativos es elevado.

#### IV. RESULTADOS

En el presente estudio, se compararon dos modelos de análisis de sentimientos: SGDClassifier y BERT, aplicados a diferentes conjuntos de datos etiquetados con diferentes niveles de granularidad. La Tabla 1 muestra una comparación de las métricas de rendimiento de ambos modelos. Los resultados indican que el modelo BERT obtiene un mejor rendimiento que SGDClassifier en todos los casos.

Tabla 1. Resultados de los modelos

Dataset/Métricas	Modelos					
	SGDClassifier			BERT		
	f1	Precision	Recall	f1	Precision	Recall
Base-Label emociones	0.18	0.31	0.19	0.33	0.33	0.34
Label Sentimientos	0.53	0.54	0.54	0.61	0.61	0.62
Label sentimiento Polarity Vader	0.81	0.81	0.81	0.93	0.93	0.93
Label Sentimientos - Balanceo	-	-	-	0.57	0.58	0.57
Label sentimiento Polarity Vader - Balanceo	-	-	-	0.89	0.89	0.89

En el conjunto de datos etiquetados con emociones básicas, se observó que el modelo SGDClassifier obtiene un bajo rendimiento en todas las métricas, mientras que BERT logra una mejora significativa en f1, precisión y recall. Para el conjunto de datos etiquetados con sentimientos, BERT logra un rendimiento superior a SGDClassifier en todas las métricas, alcanzando valores de f1, precisión y recall superiores al 60%.

En el conjunto de datos etiquetados con sentimientos polarizados utilizando la librería de Vader, ambos modelos logran un alto rendimiento en todas las métricas, siendo el modelo BERT el que obtiene valores de f1, precisión y recall más altos. Además, se evaluó el rendimiento de ambos modelos en conjuntos de datos balanceados para el etiquetado de sentimientos.

En cuanto al modelo SGDClassifier en el conjunto de datos de emociones básicas, la Figura 5 muestra la evaluación utilizando diferentes métricas. Se puede observar que el modelo no es muy efectivo para clasificar las emociones de los textos de los tweets, y se necesita mejorar el modelo o recopilar más datos para mejorar la precisión y recall.

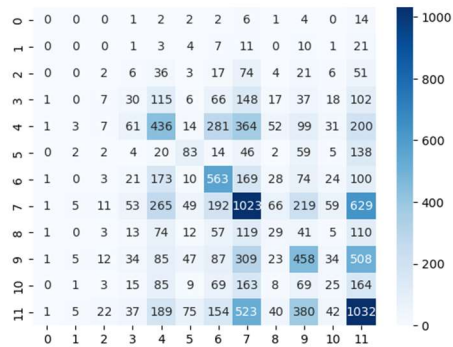


Figura 5. Matriz de confusión del modelo base con SGD.

La Figura 6 presenta la curva ROC del modelo SGDClassifier en el conjunto de datos de emociones básicas. El valor de 0.48 sugiere que el modelo tiene una capacidad limitada para distinguir entre las diferentes clases.

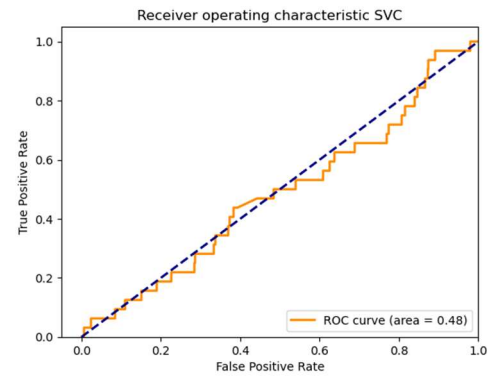


Figura 6. La curva ROC del modelo base SGD.

En el caso del modelo SGDClassifier con clasificación de los sentimientos con polaridad de Vader, se obtuvo mejoras en el rendimiento y clasificación, aunque el balanceo de la data no obtuvo una gran mejora en la homogeneidad. La Figura 7 muestra la matriz de confusión donde se observa el rendimiento del modelo.

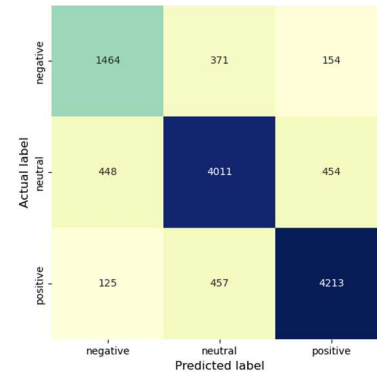


Figura 7. Matriz de confusión aplicando análisis de sentimiento de Vader.

En la figura 8. La curva ROC del modelo tiene un valor de 0.92, lo que sugiere que el modelo tiene una capacidad muy alta para distinguir entre las diferentes clases.

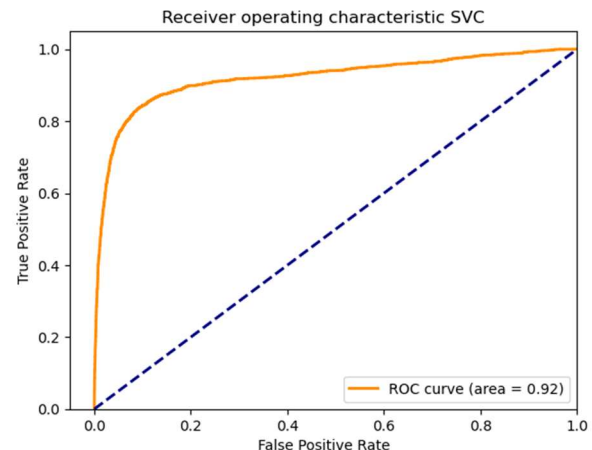


Figura 8. La curva ROC del modelo aplicando análisis de sentimiento de Vader con SGD.

En el caso del modelo pre-entrenado BERT, se observó una mejora en el modelaje con las emociones del dataset base u original. La Figura 9 muestra la matriz de confusión y se puede observar que el modelo tuvo un rendimiento razonable en algunas emociones, como "love" y "happiness", pero tuvo un rendimiento bastante pobre en otras, como "anger" y "boredom".

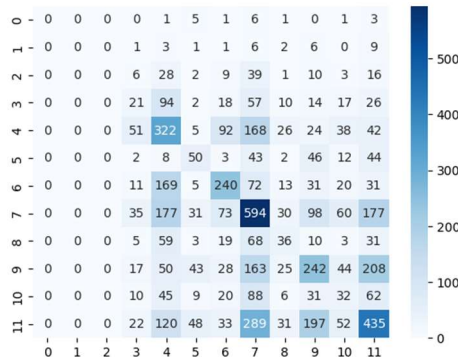


Figura 9. Matriz de confusión del modelo base con Bert.

Esto puede indicar que ciertas emociones son más difíciles de detectar que otras, y que puede ser necesario ajustar el modelo para mejorar su rendimiento en estas emociones.

Para el modelo de Bert con clasificación de los sentimientos con polaridad de Vader se observa una gran mejora, sobre todo con los datos balanceados. En la figura 10, con la dataset no balanceada se puede ver que el modelo tiene un buen desempeño en general, con la mayoría de las predicciones en la diagonal principal. Sin embargo, parece haber un poco más de dificultad en la clasificación de la clase "Neutral", donde hay un mayor número de falsos negativos (FN) y falsos positivos (FP) en comparación con las otras clases.

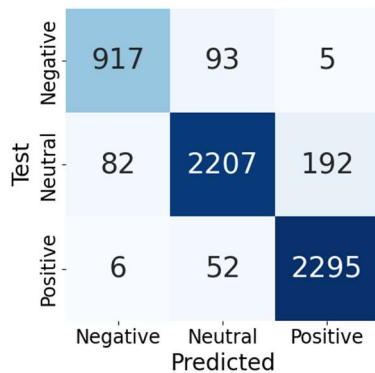


Figura 10. Matriz de confusión aplicando análisis de sentimiento de Vader con Bert, sin balancear muestra.

En cambio con la muestra balanceada, se observa El modelo tiene un desempeño similar en las categorías de sentimiento "negativo" y "positivo", con una precisión de 0.85 y 0.92 respectivamente. En cambio, el desempeño en la categoría "neutral" es ligeramente menor, con una precisión de 0.85.

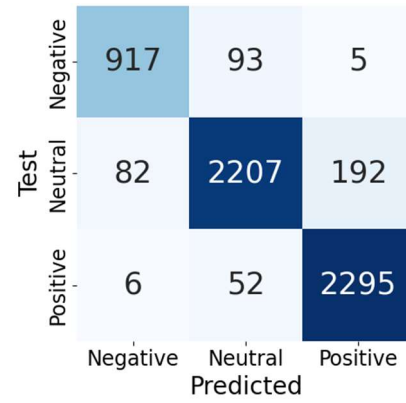


Figura 11. Matriz de confusión aplicando análisis de sentimiento de Vader con Bert, balanceada muestra.

Con respecto a la gráfica visual sobre el rendimiento del modelo durante el entrenamiento y la validación y muestran que el modelo no está sobre ajustando.

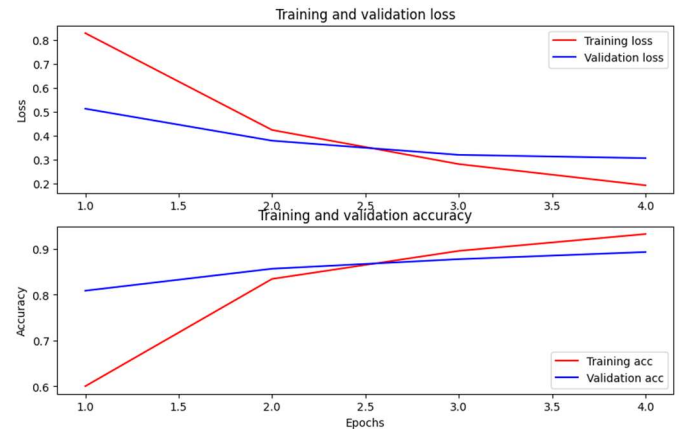


Figura 12. El rendimiento del modelo de Bert. Muestra balanceada.

En resumen, los resultados indican que el modelo BERT supera al SGDClassifier en todos los conjuntos de datos analizados en términos de métricas de rendimiento de clasificación de sentimientos. En particular, el modelo BERT mostró un mejor rendimiento en el conjunto de datos etiquetados con sentimientos, alcanzando valores de f1, precisión y recall superiores al 60%. Además, los resultados sugieren que el balanceo de los datos puede ser una estrategia efectiva para mejorar el rendimiento del modelo en la clasificación de sentimientos con polaridad de Vader. Sin embargo, el modelo BERT todavía tiene dificultades en la detección de ciertas emociones, como "anger" y "boredom", lo que indica que se necesitan más investigaciones para mejorar el rendimiento del modelo en estas áreas.

## V. CONCLUSION

En conclusión, el modelo BERT supera significativamente al modelo SGDClassifier en la tarea de análisis de sentimientos en diferentes conjuntos de datos etiquetados con diferentes niveles de granularidad. Estos resultados sugieren que los

modelos basados en deep learning, como BERT, pueden ser una opción más efectiva para el análisis de sentimientos en texto no estructurado.

Aun cuando realizar análisis de sentimientos en textos muy cortos (que por lo general no cumplen con las debidas estructuras gramaticales y reglas sintácticas del lenguaje) representa un gran reto, en general, los resultados demuestran que los modelos basados en Deep Learning combinados con técnicas de NLP resultan ser adecuados en este contexto.

## REFERENCES

- [1] D. Haryadi and G. P. Kusuma, "Emotion detection in text using nested long short-term memory," 11480 (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 10, pp. 6, 2019.
- [2] K. Sailunaz, M. Dhaliwal, J. Rokne and R. Alhajj, "Emotion detection from text and speech: A survey," Social Network Analysis and Mining, vol. 8, pp. 1–26, 2018.
- [3] Z. Jianqiang, G. Xiaolin and Z. Xuejun, "Deep convolution neural networks for twitter sentiment analysis," IEEE Access, vol. 6, pp. 23253–23260, 2018.
- [4] M. Munikar, S. Shakya and A. Shrestha, "Fine-grained sentiment classification using BERT," IEEE, vol. 1, pp. 1–5, 2019.
- [5] Abirami, A. M., & Gayathri, V. (2017). A survey on sentiment analysis methods and approach. 2016 Eighth International Conference on Advanced Computing (ICoAC), 72–76. <https://doi.org/10.1109/ICoAC.2017.7951748>.
- [6] Al-Ayyoub, M., Khamaiseh, A. A., Jararweh, Y., & Al-Kabi, M. N. (2019). A comprehensive survey of arabic sentiment analysis. Information Processing & Management, 56(2), 320–342. <https://doi.org/10.1016/j.ipm.2018.07.006>.
- [7] Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., & Manicardi, S. (2016). A comparison between preprocessing techniques for sentiment analysis in Twitter. CEUR Workshop Proceedings, 1748, 1–11. <https://ceur-ws.org/Vol-1748/paper-06.pdf>.
- [8] Baby, C. J., Khan, F. A., & Swathi, J. N. (2017). Home automation using IoT and a chatbot using natural language processing.
- [9] 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), 1–6. <https://doi.org/10.1109/IPACT.2017.8245185> Breck, E., & Cardie, C. (2017).
- [10] Opinion Mining and Sentiment Analysis. In The Oxford Handbook of Computational Linguistics 2nd edition. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199573691.013.43>.
- [11] Cambria, E. (2016). Affective Computing and Sentiment Analysis. IEEE Intelligent Systems, 31(2), 102–107. <https://doi.org/10.1109/MIS.2016.31>
- [12] Chaturvedi, I., Cambria, E., Welsch, R. E., & Herrera, F. (2018). Distinguishing between facts and opinions.
- [13] <https://github.com/twintproject/twint>
- [14] <https://data.world/crowdfunder/sentiment-analysis-in-text>
- [15] [https://huggingface.co/docs/transformers/model\\_doc/bert](https://huggingface.co/docs/transformers/model_doc/bert)
- [16] Wong, Y. K., Lee, S. S., & Hung, E. (2017). Emotion Detection in Twitter Using Deep Learning Approaches. In Proceedings of the 2017 International Joint Conference on Neural Networks.