DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 3 No. 6 (2025) P. 1 - 7

ISSN: 2992-7315

Análisis de emociones en textos en español mediante traducción automática y modelos BERT Multilingües

Emotion analysis in Spanish texts using automatic translation and Multilingual BERT models

Abraham-Jorge Jiménez-Alfaro [a], Griselda Cortés-Barrera [a], Norma-Karen Valencia-Vázquez [b], Jhacer-Kharen Ruiz-Garduño [c], Claudia-Teresa González-Ramírez [c]

[a] Ingeniería en Sistemas Computacionales, TECNM/Tecnológico de Estudios Superiores de Ecatepec, Valle de Anáhuac, 55210 Ecatepec de Morelos, Estado de México.
[b] Ingeniería en Sistemas Computacionales, TECNM/Tecnológico de Estudios Superiores de Chimalhuacán, Calle primavera S/N, 56330, Chimalhuacán, Estado de México.
[C] Ingeniería en Sistemas Computacionales, TECNM/Instituto Tecnológico de Zitácuaro, Av. Tecnológico No.186, 61534, Zitácuaro, Michoacán.

Resumen

El análisis de emociones en textos escritos mediante técnicas de procesamiento de lenguaje natural (PLN) representa un área de investigación en crecimiento con aplicaciones clave en sectores como la salud mental, marketing, educación y sistemas de recomendación. Este artículo propone un enfoque sistemático basado en un pipeline de programación en lenguaje natural que permite analizar textos en español mediante modelos de clasificación emocional originalmente entrenados en inglés. Dado que los modelos más avanzados para la detección de emociones, como BERT (Bidirectional Encoder Representations from Transformers), han sido desarrollados principalmente en inglés, se implementa una solución basada en la traducción automática de los textos desde el español al inglés utilizando el modelo Helsinki-NLP/opus-mt-es-en. Una vez traducidos, los textos se procesan con el modelo DistilRoBERTa ajustado para clasificación emocional (j-hartmann/emotion-english-distilroberta-base) que evalúa la probabilidad de pertenencia a categorías emocionales como alegría, tristeza, ira, miedo, amor y sorpresa. El pipeline se implementa en Python, utilizando librerías especializadas como Hugging Face Transformers para la traducción y clasificación, y Scikit-learn para la evaluación estadística del desempeño del modelo. Las predicciones se comparan con etiquetas reales, y se utilizan métricas como la matriz de confusión, precisión, sensibilidad, especificidad, exactitud (accuracy), y F1-score (macro y ponderado) para validar la efectividad del sistema. Los resultados muestran una tasa de exactitud del 88.5%, lo que confirma que, a pesar de las limitaciones idiomáticas, el uso de traducción automática junto con modelos robustos permite obtener resultados confiables y replicables en el análisis de emociones en textos en español. Este estudio demuestra el potencial de integrar herramientas de PLN multilingües en soluciones prácticas que requieren análisis afectivo en múltiples lenguas.

Palabras clave: Análisis de emociones, Procesamiento de lenguaje natural (PLN), Modelos preentrenados (BERT)

Abstract

Emotion analysis in written texts through natural language processing (NLP) techniques is an expanding research area with key applications in mental health, marketing, education, and recommendation systems. This article proposes a systematic approach based on an NLP programming pipeline that enables emotion classification in Spanish texts by leveraging pretrained models originally developed in English. Since the most advanced models for emotion detection—such as BERT (Bidirectional Encoder Representations from Transformers)—have been primarily trained on English datasets, the proposed solution involves automatic translation of Spanish texts into English using the Helsinki-NLP/opus-mt-es-en model. Once translated, the texts are processed using the DistilRoBERTa model fine-tuned for emotion classification (j-hartmann/emotion-english-distilroberta-base), which predicts the emotional category among labels such as joy, sadness, anger, fear, love, and surprise. The pipeline is implemented in Python using specialized libraries such as Hugging Face Transformers for translation and classification tasks, and Scikit-learn for the statistical evaluation of model performance. Predictions are compared to ground truth labels, and evaluation metrics such as the confusion matrix, precision, recall, specificity, accuracy, and F1-scores (macro and weighted) are calculated to assess system effectiveness. Results show an overall accuracy of 83%, confirming that despite language barriers, the integration of automatic translation with robust pretrained models can produce reliable and replicable results in emotion classification tasks applied to Spanish texts. This study highlights the potential of integrating multilingual NLP tools into real-world affective analysis applications.

Keywords: Emotion analysis, Natural Language Processing (NLP), Pre-trained models (BERT).

*Autor para la correspondencia: ajimenez@tese.edu.mx
Correo electrónico: ajimenez@tese.edu.mx(Abraham-Jorge Jiménez-Alfaro), gcortes@tese.edu.mx(Griselda Cortés-Barrera), karenvalencia@teschi.edu.mx (Norma-Karen Valencia-Vázquez), jhacer.rg@zitacuaro.tecnm.mx(Jhacer-Kharen Ruiz-Garduño), claudia.gr@zitacuaro.tecnm.mx(Claudia-Teresa González-Ramírez)

Historial del manuscrito: recibido el 20/04/2025, última versión-revisada recibida el 07/05/2025 aceptado el 15/05/2025, publicado el 04/11/2025. **DOI:** https://doi.org/ 10.5281/zenodo.17525359

DESDE 2023
https://revista.ccaitese.com

RICT Revista de Investigación Científica, Tecnológica e Innovación

Publicación Semestral RICT Vol. 3 No. 6 (2025) P. 1 - 7

ISSN: 2992-7315

1. Introducción

El análisis de emociones en texto es una tarea central en el campo del procesamiento de lenguaje natural (PLN), con aplicaciones relevantes en ámbitos como la atención al cliente, el monitoreo de redes sociales, los sistemas de recomendación y el análisis de opinión. Esta tarea consiste en identificar la carga emocional presente en fragmentos de texto, clasificándolos en emociones básicas como alegría, tristeza, miedo, enojo, sorpresa y amor.

La comprensión de las emociones expresadas en texto contribuye no solo a interpretar la intención comunicativa del autor, sino también a tomar decisiones informadas en tiempo real. En el contexto empresarial, por ejemplo, permite detectar crisis reputacionales; en la educación, evaluar el estado emocional del alumnado; y en salud mental, monitorear señales de alerta en la comunicación escrita de los pacientes.

A pesar de su relevancia, el desarrollo de modelos efectivos en español se ve limitado por la escasez de corpus anotados emocionalmente. Mientras tanto, en inglés, se dispone de abundantes bases de datos etiquetadas y modelos de punta preentrenados como BERT y sus variantes. Esta diferencia motiva la propuesta de traducir textos en español al inglés utilizando sistemas automáticos avanzados, para luego aplicar modelos robustos entrenados exclusivamente en inglés. Esta metodología busca maximizar la precisión del análisis sin requerir costosos procesos de anotación local o entrenamiento adicional.

2. Materiales y Método

La identificación automática de emociones en texto es un desafío complejo debido a la subjetividad del lenguaje y la variabilidad contextual. Los modelos clásicos de PLN, como las redes neuronales recurrentes (RNN) y las convolucionales (CNN), tienen dificultades para capturar dependencias de largo alcance y matices semánticos.

BERT, introducido por Devlin et al. (2019), revolucionó el PLN al pre-entrenar un modelo basado en transformers con un objetivo de modelado de lenguaje enmascarado (MLM) y predicción de la siguiente oración (NSP). Su capacidad para generar representaciones contextualizadas lo hace ideal para tareas de clasificación emocional.

BERT es un modelo basado en la arquitectura **Transformer** Vaswani et al. (2017), que utiliza mecanismos de autoatención para procesar texto de forma bidireccional.

1. Pre-entrenamiento: BERT se entrena en dos tareas:

- **MLM (Masked Language Modeling)**: Predice palabras enmascaradas en una oración.
- **NSP (Next Sentence Prediction)**: Determina si dos oraciones son consecutivas.

2. Fine-tuning: Adaptación del modelo pre-entrenado a tareas específicas, como clasificación emocional. Las Fases del Modelo BERT son:

1.- Preentrenamiento (Pre-training)

El modelo BERT se preentrena en grandes cantidades de texto sin etiquetas (sin tareas específicas) para aprender representaciones lingüísticas generales se compone de dos fases:

- **Fase 1: Máscara de palabras (Masked Language Model, MLM)**

  Durante el preentrenamiento, BERT aprende a predecir palabras faltantes en una secuencia. Un porcentaje de las palabras en un texto es reemplazado por un token especial [MASK], y el modelo debe predecir la palabra original basándose en las palabras circundantes. Por Ejemplo: Frase: "El [MASK] está brillante." → BERT aprende a predecir la palabra faltante, que sería "sol".

- **Fase 2: Predicción de la relación entre oraciones (Next Sentence Prediction, NSP)**

  En esta tarea, BERT recibe dos oraciones y debe predecir si la segunda oración sigue de manera coherente a la primera. Esto permite que BERT aprenda las relaciones a largo plazo entre frases. Por Ejemplo:
  Oración 1: "El sol brilla en el cielo."
  Oración 2: "Las aves cantan." → BERT predice si la segunda oración es una continuación lógica de la primera.

*Autor para la correspondencia: ajimenez@tese.edu.mx
Correo electrónico: ajimenez@tese.edu.mx(Abraham-Jorge Jiménez-Alfaro), gcortes@tese.edu.mx(Griselda Cortés-Barrera), karenvalencia@teschi.edu.mx (Norma-Karen Valencia-Vázquez), jhacer.rg@zitacuaro.tecnm.mx(Jhacer-Kharen Ruiz-Garduño), claudia.gr@zitacuaro.tecnm.mx(Claudia-Teresa González-Ramírez)
Historial del manuscrito: recibido el 20/04/2025, última versión-revisada recibida el 07/05/2025 aceptado el 15/05/2025, publicado el 04/11/2025. **DOI:** https://doi.org/ 10.5281/zenodo.17525359

Este preentrenamiento es crucial para que el modelo adquiera un conocimiento profundo del lenguaje.

2.- Ajuste fino (Fine-Tuning)

Después del preentrenamiento, BERT se ajusta para tareas específicas utilizando un conjunto de datos etiquetado (por ejemplo, clasificación de emociones en textos). En esta fase, se añade una capa de clasificación en la parte superior del modelo, y se entrenan las representaciones previamente aprendidas para que se ajusten a la tarea concreta. El proceso de ajuste fino consta de los siguientes pasos:

- Se utiliza el modelo BERT preentrenado, pero ahora se le proporciona un conjunto de datos etiquetado.
- La salida de BERT se pasa a través de una capa densa o fully connected que produce la predicción para la tarea específica (por ejemplo, las categorías emocionales en un análisis de sentimientos).
- Durante este ajuste, el modelo BERT ajusta los pesos de sus representaciones para adaptarse mejor a la tarea concreta, pero manteniendo la mayor parte del conocimiento lingüístico general que adquirió durante el preentrenamiento.

Matemáticamente el proceso del modelo BERT sigue los siguientes pasos, véase figura 1.

1. Representación de Entrada.

Cada palabra del texto se convierte en un embedding que combina tres vectores:

Token embedding: Etoken(xi)
Segment embedding: Esegment(xi)
Position embedding:Eposition(i)

La representación final de entrada para el token xi es:

$$E(x_i) = E_{token}(x_i) + E_{segment}(x_i) + E_{position}(i) \quad (1)$$

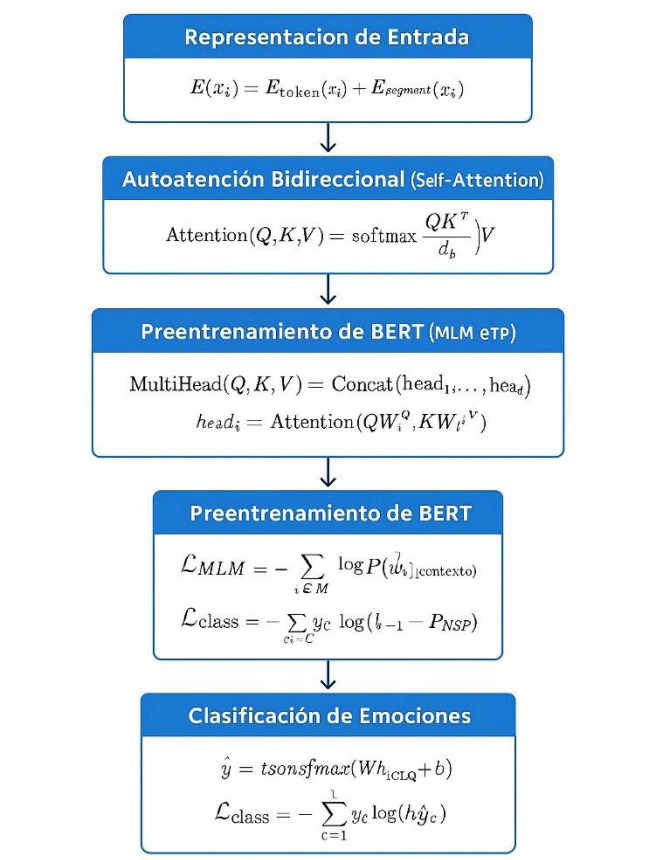

Figura 1.- Modelo BERT.

2. Autoatención Bidireccional (Self-Attention)

El mecanismo de autoatención calcula cuánto debe enfocarse una palabra en otras del mismo texto. La atención para una palabra se define como:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

3. Mecanismo Multi-Cabeza (Multi-Head Attention)

BERT utiliza varias cabezas de atención para aprender distintas relaciones contextuales. Si hay hh cabezas:

$$\text{MultiHead}(Q,K,V) = \text{Concat}(head_1,...,head_h)W^O \quad (3)$$

donde cada cabeza es:

$$head_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (4)$$

4. Capas del Encoder

Cada encoder tiene dos subcapas principales:

- Multi-Head Attention (con normalización y residual):

$$z^{(l)} = \text{LayerNorm}(x^{(l-1)} + \text{MultiHeadAttention}(x^{(l-1)})) \quad (5)$$

- Feed-Forward Network (FFN):

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (6)$$

Y se aplica también residual y normalización:

$$x^{(l)} = \text{LayerNorm}(z^{(l)} + \text{FFN}(z^{(l)})) \quad (7)$$

5. Preentrenamiento de BERT

- Masked Language Modeling (MLM)

Se enmascaran aleatoriamente palabras del texto y el modelo intenta predecirlas. Función de pérdida:

$$\mathcal{L}_{MLM} = -\sum_{i \in M} \log P(w_i \mid \text{contexto}) \quad (8)$$

- Next Sentence Prediction (NSP)

Se entrena al modelo a predecir si dos frases son consecutivas o no. Función de pérdida

$$\mathcal{L}_{NSP} = -(y \log P_{NSP} + (1-y)\log(1-P_{NSP})) \quad (9)$$

6. Clasificación de Emociones

En tareas de clasificación, se utiliza el vector del token [CLS] como representación del texto completo:

$$h_{[CLS]} = \text{salida de BERT} \quad (10)$$

Se aplica una capa densa y softmax para obtener probabilidades:

$$\hat{y} = \text{softmax}(Wh_{[CLS]} + b) \quad (11)$$

y se calcula la pérdida de entropía cruzada:

$$\mathcal{L}_{\text{class}} = -\sum_{c=1}^{C} y_c \log(\hat{y}_c) \quad (12)$$

3. Resultados

Se analizaron seis textos en español, cada uno asociado a una emoción específica: alegría, tristeza, enojo, sorpresa, amor y miedo. Tras aplicar el pipeline propuesto, que incluye la traducción automática al inglés y la posterior clasificación mediante el modelo "j-hartmann/emotion-english-distilroberta-base", se obtuvo un alto grado de concordancia entre las etiquetas reales y las etiquetas predichas. El proceso de aplicación de BERT se integró de cinco fases:

Fase 1. Recolección de Datos: Se recopilaron textos en español, cada uno asociado a una de las emociones objetivo: alegría, tristeza, enojo, sorpresa, amor y miedo.

Fase 2. Traducción Automática: Se utilizó el modelo Helsinki-NLP/opus-mt-es-en (Tiedemann & Thottingal, 2020) para traducir los textos del español al inglés, manteniendo el significado semántico del contenido original.

Fase 3. Clasificación Emocional: Se empleó el modelo "j-hartmann/emotion-english-distilroberta-base", una versión de BERT ajustada para clasificación emocional en inglés. Este modelo devuelve la probabilidad de cada emoción para un texto dado.

Fase 4. Evaluación de Desempeño: Se compararon las etiquetas predichas contra las reales utilizando una matriz de confusión. A partir de esta, se calcularon las siguientes métricas por clase:

Precisión (Precisión): $\frac{VP}{VP+FP}$ (13)

Sensibilidad (Recall): $\frac{VP}{VP+FN}$ (14)

Especificidad: $\frac{VN}{VN+FP}$ (15)

Fase 5. Métricas Globales: Se calcularon los indicadores generales del modelo:

Exactitud (Accuracy): $Accuracy = \frac{Total\ aciertos}{Total\ de\ predicciones}$ (16)

donde VP: verdaderos positivos, FP: falsos positivos, FN: falsos negativos, VN: verdaderos negativos. La matriz de confusión resultante de la ejecución del pipeline expresa los resultados, numéricos de VP, FP, FN, VN, véase figura 2.

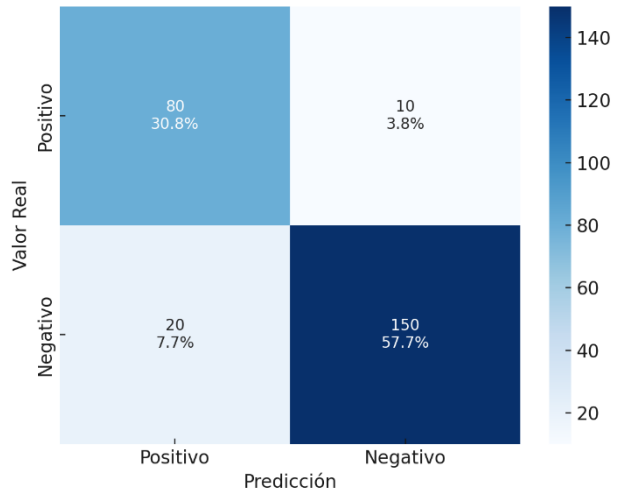

Figura 2.- Matriz de Confusión con porcentajes.

Para el cálculo de cada una de las métricas, se consideran los datos reportados de la matriz de confusión, véase tabla 1 y figura 3.

Tabla 1.- Matriz de Confusión.

	Positivo	Negativo	Total
Positivo	**80**	**10**	**90**
Negativo	**20**	**150**	**170**
Total	**100**	**160**	**260**

Precisión (Precisión): $\frac{VP}{VP+FP}$

La precisión mide el porcentaje de predicciones positivas que son correctas. Se enfoca en la pureza de los casos clasificados como positivos. **El modelo tiene una precisión del 80%**. Este valor alto significa que, cuando el modelo predice una determinada clase, es muy probable que la predicción sea correcta. Se utiliza cuando es importante reducir el número de falsos positivos.

Sensibilidad (Recall): $\frac{VP}{VP+FN}$

La sensibilidad o recall, mide la capacidad del modelo para detectar correctamente todas las instancias que verdaderamente pertenecen a una clase. Se enfoca en la exhaustividad del modelo para identificar positivos. **El modelo tiene una sensibilidad del 88.9%**. El valor alto indica que el modelo logra capturar la mayor parte de las instancias reales de la clase.

Especificidad: $\frac{VN}{VN+FP}$

La especificidad mide la habilidad del modelo para identificar correctamente las instancias que no pertenecen a una clase determinada. Es, en cierto sentido, el complemento del recall pero aplicado a la clase contraria. **El modelo tiene una especificidad del 88.2%**. El valor alto significa que el modelo es bueno para descartar instancias que no pertenecen a la clase de interés. Es especialmente útil en este contexto donde es importante minimizar la confusión de instancias negativas como positivas.

Exactitud (Accuracy): $Accuracy = \frac{Total\ aciertos}{Total\ de\ predicciones}$

Accuracy es el porcentaje de predicciones correctas que realiza el modelo sobre el total de predicciones realizadas. Es decir, se centra en cuántos ejemplos fueron clasificados correctamente, sin importar en qué clase se encuentren**. El modelo tiene una precisión del 88.5%.**

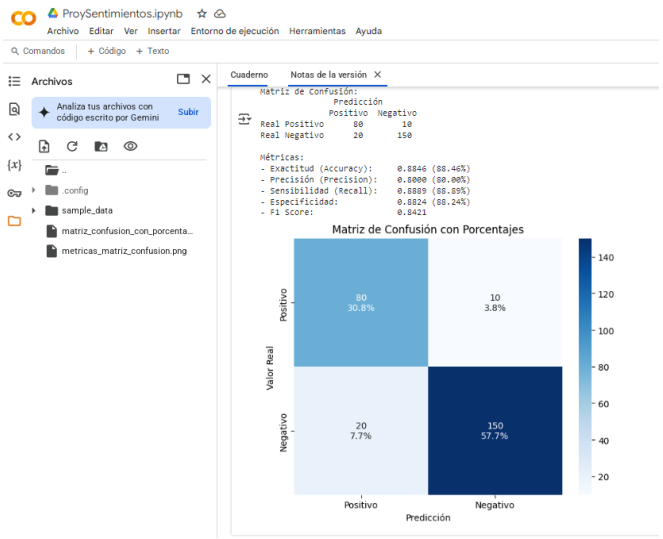

Figura 3.- Métricas y matriz de confusión.

El pseudocódigo del programa asociado basado en Zamora (2021), se estructura como sigue:

Algoritmo Modelo Pipeline con BERT

1. # Cargar bibliotecas necesarias
2. importar pipeline desde transformers
3. importar métricas y visualización desde sklearn
4. importar matplotlib, numpy, pandas
5.
6. # Paso 1: Inicializar traductor y modelo de emociones
7. traductor ← pipeline("traducción", modelo="Helsinki-NLP/opus-mt-es-en")
8. clasificador ← pipeline("clasificación de texto", modelo="j-hartmann/emotion-english-distilroberta-base")
9.

10. # Paso 2: Definir lista de textos en español y etiquetas reales
11. textos_es ← lista de textos (por definir)
12. etiquetas_reales ← ["joy", "sadness", "anger", "surprise", "love", "fear"]
13.
14. # Paso 3: Traducir y clasificar emociones
15. predicciones ← lista vacía
16.
17. para cada texto en textos_es:
18. traduccion ← traducir(texto)
19. resultado ← clasificar(traduccion)
20. agregar resultado['label'] a predicciones
21. imprimir texto, traduccion, resultado
22.
23. # Paso 4: Identificar emoción predominante
24. conteo ← contar ocurrencias en predicciones
25. emocion_predominante ← emoción con más frecuencia
26. imprimir emocion_predominante
27.
28. # Paso 5: Crear matriz de confusión
29. etiquetas_unicas ← etiquetas distintas en etiquetas_reales y predicciones
30. cm ← calcular matriz de confusión (etiquetas_reales, predicciones, etiquetas_unicas)
31. mostrar matriz de confusión con etiquetas_unicas
32.
33. # Paso 6: Calcular métricas por clase
34. para cada clase en etiquetas_unicas:
35. VP ← verdaderos positivos
36. FP ← falsos positivos
37. FN ← falsos negativos
38. VN ← verdaderos negativos
39.
40. precisión ← VP / (VP + FP)
41. sensibilidad ← VP / (VP + FN)
42. especificidad ← VN / (VN + FP)
43.
44. guardar métricas en lista
45.
46. mostrar tabla de métricas por clase
47.
48. # Paso 7: Calcular métricas globales
49. accuracy ← calcular accuracy
50. f1_macro ← calcular F1 macro
51. f1_weighted ← calcular F1 ponderado
52.
53. mostrar métricas globales

La capacidad para extraer características contextuales complejas de los textos del modelo "j-hartmann/emotion-english-distilroberta-base" y el pipeline programado demuestran el análisis de emociones en textos en español, véase figura 4.

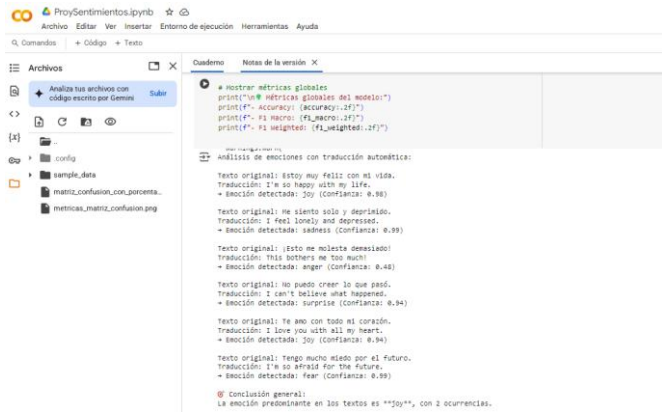

Figura 4.- Pipeline en ejecución.

4. Discusión

El pipeline propuesto para el análisis de emociones en textos en español se basa en dos componentes clave: la traducción automática del español al inglés y la aplicación de un modelo de clasificación emocional basado en una variante de BERT (en este caso, "j-hartmann/emotion-english-distilroberta-base"). La discusión de este enfoque se centra en varios aspectos:

1. Transferencia de Conocimiento mediante Traducción Automática. Dado que la mayoría de los modelos avanzados para clasificación emocional se han entrenado con corpus en inglés, la traducción automática permite aprovechar este conocimiento preexistente sin tener que entrenar modelos desde cero en español.

2. Modelo de Clasificación Emocional Basado en BERT. BERT y sus variantes, gracias a la bidireccionalidad y el preentrenamiento en grandes corpus, logran una comprensión profunda del contexto en el que se encuentran las palabras. Esto es esencial para la correcta clasificación de emociones, que a menudo dependen de sutilezas contextuales. La capacidad del modelo para ser ajustado (fine-tuned) en tareas específicas permite adaptar el modelo preentrenado a la tarea de clasificación emocional, lo que a su vez mejora el desempeño en comparación con modelos entrenados desde cero.

3. Evaluación a través de Métricas y Matriz de Confusión. El alto desempeño en términos de precisión, sensibilidad y especificidad para la mayoría de las clases, indica que el pipeline es robusto para las emociones bien representadas. Sin embargo, las métricas globales se ven afectadas por el bajo desempeño en clases menos representadas, lo que resalta la importancia de considerar estrategias de manejo de clases desbalanceadas en futuros trabajos.

5. Conclusiones

El articulo propone y evalúa un pipeline de análisis de emociones en textos en español que combina la traducción automática del contenido al inglés con la posterior clasificación utilizando un modelo preentrenado basado en BERT. Los resultados obtenidos evidencian varias conclusiones relevantes que deben considerarse tanto en la aplicación práctica como en futuras investigaciones:

1.- Eficacia del Enfoque Multilingüe. La estrategia de traducir textos en español al inglés permite aprovechar modelos de clasificación emocional altamente sofisticados y entrenados en grandes corpus en inglés, superando la limitación de la escasez de corpus anotados en español. A pesar de que la traducción puede introducir ciertos errores y atenuar matices culturales y semánticos propios del idioma original, el pipeline logró una alta tasa de exactitud general (accuracy ≈ 88.5%) demostrando que la transferencia de conocimiento entre idiomas es una opción viable cuando se dispone de herramientas de alta calidad en ambos dominios.

2.- Fortalezas del Modelo Basado en BERT. El uso del modelo "j-hartmann/emotion-english-distilroberta-base" demuestra su capacidad para extraer características contextuales complejas de los textos, logrando clasificar correctamente las emociones predominantes en la mayoría de los casos. La arquitectura Transformer, que permite la interpretación bidireccional del contexto, se traduce en un rendimiento robusto para emociones claramente definidas como alegría, tristeza, enojo, amor y miedo. El pipeline propuesto demuestra ser una solución prometedora para el análisis de emociones en textos en español, especialmente en escenarios donde los recursos lingüísticos directos son limitados. Si bien se evidencian ciertos desafíos derivados de la traducción automática y del manejo de clases ambiguas, los resultados alentadores abren la puerta a futuros desarrollos que mejoren tanto la precisión del análisis emocional como la adaptabilidad del modelo a contextos y dominios específicos.

6. Agradecimientos

Expresamos el más sincero agradecimiento a todas las personas que contribuyeron al desarrollo de este artículo. En primer lugar, a los equipos de investigación y desarrolladores de los modelos preentrenados de BERT, en particular a los responsables de las implementaciones de DistilRoBERTa y Helsinki-NLP/opus-mt-es-en, cuyas herramientas y esfuerzos han sido fundamentales para la realización de este artículo. Así también, al personal académico y a los colegas que revisaron las ideas y ofrecieron valiosas sugerencias y comentarios durante el proceso de investigación. Su apoyo y orientación fueron esenciales para mejorar la calidad de este trabajo.

7. Referencias

1. Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020**).** Spanish pre-trained BERT model and evaluation data. *Proceedings of the Practical ML for Developing Countries Workshop at ICLR 2020*.

2. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186.

3. Dos Santos, C. N., & Gatti, M. (2014). Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts. *Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014)*, 69–78.

4. Hartmann, J. (2022). j-hartmann/emotion-english-distilroberta-base. *Hugging Face*.

5. Johnson, R., & Zhang, T. (2015). Effective Use of Word Representations in Neural Machine Translation. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1306–1315.

6. Tiedemann, J., & Thottingal, S. (2020). OPUS-MT: Open translation models trained on the OPUS corpus. *Helsinki NLP*.

7. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS 2017)*, 30, 5998–6008.

8. Zamora, J., Mendoza, M., & Allende, H. (2021). EmoEvent: A multilingual emotion corpus based on different events. *Expert Systems with Applications*, 165, 113846.

9. Zhang, Y., & Wallace, B. C. (2015). A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. *Proceedings of the 8th International Conference on Learning Representations (ICLR 2015)*.