Polo del Conocimiento

Pol. Con. (Edición núm. 85) Vol. 8, No 8
Agosto 2023, pp. 2552-2565
ISSN: 2550 - 682X
DOI: 10.23857/pc.v8i8.5987

Identificando cambios de autor en un texto mediante codificación de embeddings de tokens iniciales de las capas de atención.

Identifying author changes in a text by encoding embeddings of initial tokens from the attention layers.

Identificar alterações de autor em um texto codificando incorporações de tokens iniciais das camadas de atenção.

César Espín-Riofrio [I]
cesar.espinr@ug.edu.ec
https://orcid.org/0000-0001-8864-756X

Fernando Alvear-Ferrín [II]
fernando.alvearf@ug.edu.ec
https://orcid.org/0009-0009-4813-2532

Bolívar Pazmiño-Bermúdez [III]
bolivar.pazminob@ug.edu.ec
https://orcid.org/0009-0008-3102-1740

Tania Peralta-Guaraca [IV]
tania.peraltag@ug.edu.ec
https://orcid.org/0000-0002-4879-6824

Rocío Carchi-Encalada [V]
rocio.carchie@ug.edu.ec
https://orcid.org/0009-0009-6343-2939

Correspondencia: https://orcid.org/0000-0001-8864-756X

Ciencias de la Computación
Artículo de Investigación

*** Recibido:** 23 de junio de 2023 ***Aceptado:** 12 de julio de 2023 *** Publicado:** 30 de agosto de 2023

I. Magister, Universidad de Guayaquil, Ecuador
II. Universidad de Guayaquil, Ecuador
III. Universidad de Guayaquil, Ecuador
IV. Magister, Universidad de Guayaquil, Ecuador
V. Máster, Universidad de Guayaquil, Ecuador

Resumen

La determinación de autoría es una herramienta esencial en la detección de plagio y atribución errónea de autor en diversas áreas. En este trabajo, se aborda la problemática de determinar cambios de autor en un texto. Tradicionalmente, muchas investigaciones utilizan la salida final de codificación de las capas de atención en tareas de clasificación de textos. Proponemos extraer los embeddings de codificación de los tokens iniciales de las capas de atención de modelos Transformer pre entrenados basados en BERT, aplicando aprendizaje por transferencia para realizar un ajuste fino del modelo y luego proceder a la predicción. Los modelos mDeBERTa y DeBERTa se seleccionan para la experimentación. El enfoque se valida utilizando un dataset de las campañas PAN 2023 para determinar cambios de autor, que contiene pares de textos en inglés de distintos dominios. Este estudio tiene una importancia significativa en diversas disciplinas que requieran la verificación de autoría. Si bien los resultados obtenidos no fueron los esperados, el método propuesto es un prometedor punto de partida para futuras investigaciones sobre el tema.

Palabras Clave: Cambios de autor; Procesamiento de Lenguaje Natural; Modelos Transformers; Embeddings de tokens iniciales.

Abstract

The determination of authorship is an essential tool in the detection of plagiarism and erroneous author attribution in various areas. In this paper, the problem of determining author changes in a text is addressed. Traditionally, many investigations use the final encoding output of attentional layers in text classification tasks. We propose to extract the encoding embeddings of the initial tokens from the attention layers of pre-trained BERT-based Transformer models, applying transfer learning to fine tune the model and then proceed to prediction. The mDeBERTa and DeBERTa models are selected for experimentation. The approach is validated using a dataset from the PAN 2023 campaigns to determine author changes, which contains pairs of texts in English from different domains. This study has significant importance in various disciplines that require verification of authorship. Although the results obtained were not as expected, the proposed method is a promising starting point for future research on the subject.

Keywords: Author changes; Natural Language Processing; Transformer models; Initial token embeddings.

Resumo

A determinação da autoria é uma ferramenta essencial na detecção de plágio e atribuição errônea de autores em diversas áreas. Neste artigo, é abordado o problema de determinar mudanças de autor em um texto. Tradicionalmente, muitas investigações utilizam a saída final de codificação de camadas de atenção em tarefas de classificação de texto. Propomos extrair os embeddings de codificação dos tokens iniciais das camadas de atenção de modelos Transformer pré-treinados baseados em BERT, aplicando aprendizagem de transferência para ajustar o modelo e então prosseguir para a previsão. Os modelos mDeBERTa e DeBERTa são selecionados para experimentação. A abordagem é validada utilizando um conjunto de dados das campanhas PAN 2023 para determinar mudanças de autor, que contém pares de textos em inglês de diferentes domínios. Este estudo tem importância significativa em diversas disciplinas que exigem verificação de autoria. Embora os resultados obtidos não tenham sido os esperados, o método proposto é um ponto de partida promissor para futuras pesquisas sobre o tema.

Palavras-chave: Mudanças de autor; Processamento de linguagem natural; Modelos de transformadores; Incorporações de token iniciais.

Introducción

En la era digital, con la proliferación de información en línea, la atribución de autoría se ha vuelto un campo de relevancia para la detección de cambio de autores, la desinformación y el contenido generado automáticamente. Esto lo hace una herramienta clave para verificar la autenticidad de la información y protegerse contra la manipulación y fraude, lo que da relevancia al presente artículo que se enfoca en la verificación de autoría de textos de diferentes tipos usando modelos de lenguaje basados en Transformers.

El Procesamiento del Lenguaje Natural (PLN) se encuentra inmerso en el reconocimiento de discursos, entendimiento del lenguaje, establece como objetivo principal que las computadoras entiendan el lenguaje y lo procesen de la misma forma que los humanos (Beltrán & Rodríguez Mojica, 2021).

La verificación de autoría ha tomado mucho poder, a través de esta es posible determinar si un texto pertenece a un autor, han disminuido las formas de plagio, y también se ha aplicado en otras áreas como la seguridad ya que es posible detectar y obtener información de personas que expresan violencia u odio en Internet. La verificación de autoría aplica diversas técnicas y métodos para

determinar a qué autor pertenece un texto, recoge las características de estos autores, su elección de palabras, su forma de escribir oraciones, la aplicación de signos de puntuación. La clasificación de textos se basa en insertar de manera correcta a un texto dentro de una categoría, bajo las diversas características que presente. (Minaee et al., 2021).

Existen varios estudios que abordan la detección de cambio de autor usando diversas técnicas de clasificación, entre los cuales podemos destacar a (Barlas & Stamatatos, 2020) donde proponen verificar cambio de autoría en un texto utilizando Multi Neural Network (MNN) combinado con modelos pre entrenados BERT, ELMo, ULMFiT y GPT-2, donde los resultados demuestran que BERT y ELMo contienen los enfoques más estables. (Fabien et al., n.d.) para la identificación de autor proponen utilizar el modelo BertAA, el cual es basado en el modelo BERT y que contiene un ajuste fino añadiendo la aplicación de rasgos estilométricos, donde analizan 3 conjuntos de datos mediante los cuales se analizará el rendimiento del modelo. Los resultados demuestran que BertAA es útil para resolver este tipo de tareas donde mejora la precisión añadiendo las características estilométricas. (Avram, 2023) para la identificación de cambio autoría, usó un modelo Transformer basado en BERT, en un dataset en lengua rumana a pesar de que se encontraba desbalanceado porque eran textos de épocas distintas y el número de autores desigual, se presentaron resultados favorables aplicando los métodos Support Vector Machine (SVM), Decision Trees (DT), Multi Expression Programming (MEP), Artificial Neural Networks (ANN) y k-Nearest Neighbour, donde luego de evaluar al modelo BERT presentó un 85.9% de precisión en las métricas usadas.

El concepto de Transformer se hace presente en 2017 por medio del artículo Attention Is All You Need (Vaswani et al., 2017), este se basa en un mecanismo de atención conectando la entrada y salida de una red neuronal de manera que el desempeño y rendimiento sea mejor (Beltrán & Rodríguez Mojica, 2021).

A finales del 2018 los científicos del laboratorio del lenguaje de IA de Google presentaron un modelo lingüístico BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al., 2018), su propósito es permitir un aprendizaje profundo representado de forma bidireccional para ser usado en los modelos de aprendizaje automático. BERT presenta 2 arquitecturas: BERT-base y BERT-large, el primero cuenta con 12 bloques codificadores y cada uno tiene 12 capas de atención y 768 capas ocultas, el segundo posee 24 bloques codificadores con 16 capas de atención cada uno y 1024 capas ocultas. Dentro de estas capas BERT posee una jerarquía de información lingüística, en las capas inferiores tiene rasgos superficiales, en las capas intermedias rasgos

sintácticos y en las capas superiores rasgos semánticos (Singh, 2022). El modelo DeBERTa (Decoding-enhanced BERT with disentangled attention) cuenta con un mecanismo de atención desarrollado y un decodificador mejorado esto lo hace más eficiente al momento de entrenar. El modelo mDeBERTa es una versión multilingüe del anterior, este cuenta con 12 capas que permite insertar 190M de parámetros en las capas de embedding (Xia et al., n.d.)

El enfoque propuesto en la presente investigación involucra la obtención de los tokens iniciales de información de capas de atención en modelos basados en BERT, es una alternativa distinta y poco investigada a los métodos actuales de verificación de autoría. La utilización de la información de capas de atención de BERT puede potenciar a mejorar la precisión de la verificación de autoría y permitir la verificación de textos de diferentes tipos de discurso. Los resultados podrían tener un impacto positivo en campos como el forense digital, la literatura y la seguridad de la información, y podrían ser utilizados por organizaciones gubernamentales, empresas y entidades académicas para mejorar la autenticidad y la integridad de la información.

Método

Este trabajo está sustentado por medio de una investigación bibliográfica, indicada para el análisis de distintos artículos científicos de alta relevancia que permitieron conocer el estado del arte y a su vez los diversos métodos empleados en investigaciones similares. Se hace uso de una metodología experimental dentro de la cual se realizan pruebas en la extracción de los embeddings de los tokens iniciales usando modelos Transformers basados en BERT. Empleando además el método cuantitativo que por medio de diversas métricas evalúa el rendimiento del modelo planteado dentro de sus dos etapas, la de entrenamiento y la de prueba.

En figura 1 se plantea el método propuesto y sus etapas, donde el dataset de entrenamiento pasa por un pre procesamiento de los datos y extracción de embeddings iniciales en todos los modelos pre entrenados, por otra parte, el dataset de pruebas se tokeniza y de esa forma ambos datasets quedan listo para el entrenamiento y ajuste. Una vez entrenado el modelo es guardado y llamado para realizar las predicciones y posterior evaluación de rendimiento.

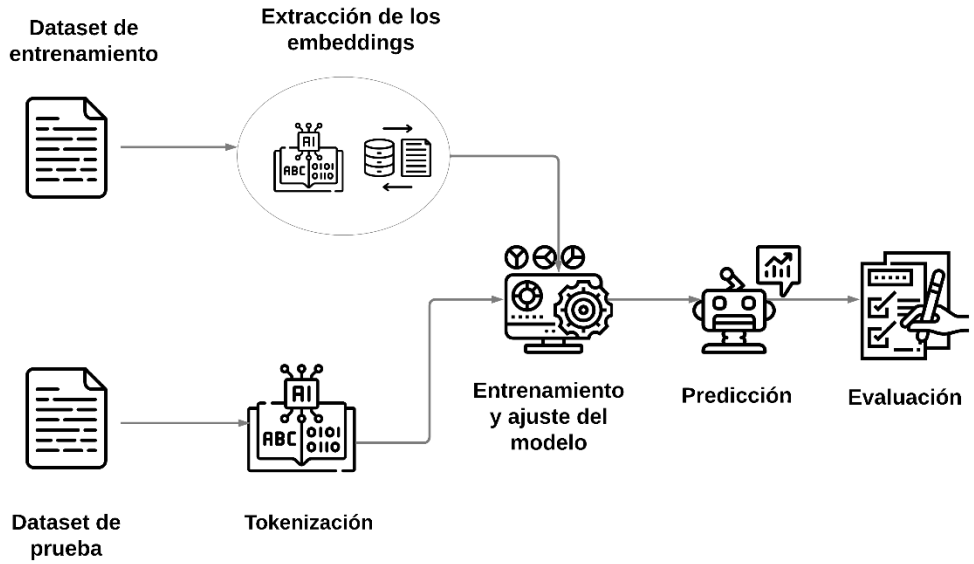

Figura 1 Método implementado en la investigación

Dataset

El dataset utilizado corresponde a una de las campañas PAN 2023 de CLEF para determinar cambios de autor. Se encuentra compuesto por correos, ensayos, entrevistas y transcripciones de discursos en idioma inglés, donde su nivel de formalidad varía entre los distintos tipos de texto. Cuenta con 8836 pares de datos que contienen dos textos de dominio distintos, como se muestra en la tabla 1.

Tabla 1

Cantidad de muestra de los dataset

Dataset	Cantidad de Muestras	Valores
Entrenamiento	[interview, email]:	4564
	[essay, email]:	1454
	[email, speech_transcription]:	1036
	[essay, interview]:	884
	[speech_transcription, interview]:	642
	[essay, speech_transcription]:	256
	Total:	**8836**

El dataset tiene un campo que nos indica si el texto fue escrito o no por el mismo autor. Con el valor 'True' indica que el texto es escrito por el mismo autor, por otra parte, el valor 'False' determina que el texto no es escrito por el mismo autor.

	id	discourse_types	pair	same
0	a0e656a3-955e-44f6-b5b3-dfd6e360a962	[email, speech_transcription]	[I think I have decided with regards to the <t...	True
1	c42efa75-f418-4fac-80ca-4d0982704d25	[email, speech_transcription]	[Dear <addr10_NN>, <nl><nl>If I was to go with...	True
2	a3a6745f-af62-4889-af49-93ec88594142	[interview, email]	[So, erm, so I'm, erm, I'm actually from <cou...	True
3	244235f3-4647-4ed3-b078-60a2ea2850b7	[email, speech_transcription]	[Hi <addr2_FN><nl><nl>This is my picture, that...	False
4	d410b3f0-b331-48ab-a524-c29ed1d487a2	[interview, email]	[Oh, okay. Erm, well, I think of myself as kin...	False

Figura 2 Muestra del dataset de origen.

Pre procesamiento de datos

Al tener valores "True" y "False", se dificulta el entrenamiento del modelo, para ello lo codificamos para que los valores sean 1 y 0 respectivamente para la correcta comprensión del algoritmo, como se puede apreciar en la siguiente figura 3. Por lo tanto, si el valor es 1 significa que el texto tiene el mismo autor, por el contrario, si el valor es 0 significa que el autor no es el mismo, es decir existe un cambio de autor.

	id	discourse_types	pair	same
219	ac9529bb-c85e-4f06-b2b0-36e59aa70dac	[interview, email]	[I think I would like to try sushi. I'm not a ...	0
801	35e48194-46c9-4c1d-8c10-3c2a9154393d	[interview, email]	[Restaurant. Erm, I don't have any fancy resta...	1
366	8a51b7b5-d920-4986-bed8-c8ebdcd248fa	[essay, email]	[Method<nl>Participants: <nl>A total of 60 par...	1
1027	8a6ed13f-dc6d-4b20-a4a8-c8e40dea514b	[essay, email]	[This essay aims to explore the e business con...	0
571	59f9e517-6ca1-4846-b128-38f70cd98fe9	[interview, email]	[Okay. Erm, so normally, every single day I h...	1

Figura 3 Datos preprocesados.

Tokenización

Se debe implementar los tokenizadores adecuados para cada modelo ya que estos necesitan procesar la información de manera numérica. La cantidad máxima de tokens con la que trabajan los modelos son 512, se realiza una segmentación de datos para evitar la pérdida de información y se los concatena para conservar la misma cantidad de registros.

Tabla 2

Tokenizadores utilizados

Modelo	Tokenizador
BERT	BertTokenizer
DeBERTa	DebertaTokenizer
mDeBERTa	AutoTokenizer

Hiperparámetros

Los hiperparámetros determinados para el entrenamiento del modelo son: función de activación, learning rate, batch size y dropout. Mediante la librería Optuna se realizaron diversas pruebas para obtener los mejores hiperparámetros en base a distintos valores o rangos propuestos.

Tabla 3

Valores para optimizar los Hiperparámetros

Hiperparámetro	Rango
Función de activación	Tanh, ReLU, GELU
Learning rate	3e-5 – 5e-5
Dropout	0.2 – 0.5
Epoch	1 – 5
Batch size	8, 16

Para potenciar la determinación de hiperparámetros, usamos la característica “EarlyStopping” la cual permite hacer una parada temprana cuando se alcanzan los criterios de maximización definidos y no se tienen variaciones relevantes en las diversas ejecuciones que realiza Optuna. Para la presente investigación, se define la parada temprana para que tome acción si en las últimas 4 ejecuciones no existen variaciones en los resultados de la variable a maximizar ‘F1’. Con esto podemos lograr tener ejecuciones más ágiles y evitar realizar intentos innecesarios que pueden aumentar el tiempo de ejecución y uso de recursos.

Tabla 4

Mejores hiperparámetros obtenidos Optuna

modelo	Función activación	Learning rate	Dropout	Epoch	Batch size
BERT	GELU	3.8e-5	0.2313	3	8
DeBERTa	Tanh	3e-5	0.2169	2	8
mDeBERTa	Tanh	4e-5	0.3303	3	16

Ajuste y entrenamiento del modelo

A los modelos base pre entrenados se les realiza un ajuste fino (fine tuning) añadiendo 2 dos funciones lineales, dropout, la función de activación y la función CrossEntropyLoss que calcula la pérdida durante el entrenamiento.

```
for index in range(t_cls_tensors.shape[0]):
  tind_cls_tensors = t_cls_tensors[index].transpose(1,0)
  pooled_layers = torch.mm(tind_cls_tensors, self.Fusion).squeeze()
  x = self.lin1(pooled_layers)
  x = self.dropout(x)
  x = self.activacion(x)
  logit = self.lin2(x)
  logits.append(logit) #guarda en cada capa
  loss = None

  if labels is not None:
    loss = self.loss_func(logit, labels[index].float())
    loss =loss.mean()
    losses.append(loss)  # guarda el loss de cada capa

logits = torch.stack(logits,dim=0)
loss = None
if labels is not None:
  loss = torch.mean(torch.stack(losses))

return SequenceClassifierOutput(loss=loss,logits=logits)
```

Figura 4 Ajuste fino de los modelos preentrenados.

Para el entrenamiento de los modelos por medio del dataset para entrenamiento, se aplican los mejores hiperparámetros capturados en la ejecución de Optuna para potenciar el entrenamiento y precisión en las métricas de evaluación.

Pol. Con. (Edición núm. 85) Vol. 8, No 8, Agosto 2023, pp. 2552-2565, ISSN: 2550 - 682X

```
trainer = MyTrainer(
    model=model,
    compute_metrics=compute_metrics,
    args=training_args,
    train_dataset=train_set,
    eval_dataset=test_set)

print("Training")
result = trainer.train()
evaluation_result = trainer.evaluate()
```

Figura 5 Entrenamiento de los modelos.

Predicción

Una vez llamado el modelo, tokenizado el dataset de prueba con el tokenizer del modelo entrenado para ser usado como entrada, se ejecuta el método model.predict() cargado para realizar predicciones y poder realizar evaluaciones de rendimiento.

.

Resultados

Para evaluar los modelos tomamos las métricas F1, accuracy, Brier y la matriz de confusión. Donde el F1 demuestra qué tan preciso son los resultados de las predicciones, Brier detalla los valores de pérdida que existan en los resultados, el accuracy va a medir la exactitud de los algoritmos y la matriz de confusión es la encargada de determinar la cantidad de predicciones correctas e incorrectas. Luego de obtener las predicciones y métricas de evaluación para ambos modelos, se procede a realizar el análisis y la valoración de los diversos resultados para evaluación de entrenamiento y predicción del modelo. Tabla 5 muestra la evaluación durante el entrenamiento.

Tabla 5

Métricas de evaluación en entrenamiento de los modelos

Modelo	Accuracy	F1	Brier	Precision
BERT	0.5253	0.5252	0.5253	0.5254
mDeBERTa	0.5099	0.3377	0.5099	0.2549
DeBERTa	0.5173	0.51731	0.5173	0.5175

Aunque los resultados obtenidos no han sido los esperados, se analiza mediante el Accuracy que los valores obtenidos para determinar el mejor rendimiento en entrenamiento, se puede evidenciar

que los modelos BERT y DeBERTa tienen las mejores marcas con un accuracy de 52.53% y 51.73% respectivamente en la determinación de autoría de textos en idioma inglés, mientras que mDeBERTa siendo el menos efectivo obtuvo un 50.99%.

Una vez finalizada la etapa de entrenamiento, se realizan pruebas de predicción con el dataset de prueba con los cuales obtuvimos el siguiente rendimiento:

Tabla 6

Métricas de evaluación en predicción de los modelos

Modelo	Accuracy	F1	Brier	Precision	Tiempo
BERT	0.5173	0.5173	0.5173	0.5174	02:43h
mDeBERTa	0.4984	0.3326	0.4985	0.2492	01:44h
DeBERTa	0.5203	0.5202	0.5204	0.5203	03:20h

Por medio de la evaluación F1 se determina el modelo más preciso en las predicciones para determinar cambio de autor en un texto. Los modelos BERT y DeBERTa obtuvieron las mejores marcas con 51.73% y 52.02% respectivamente. La métrica Brier muestra el porcentaje de pérdida, al obtener una puntuación menor indica que tan bueno es el rendimiento del modelo en evitar pérdida de información en la predicción, en este caso, el modelo mDeBERTa tiene un mejor rendimiento en porcentaje menor de pérdida a comparación con BERT y DeBERTa. Dentro de los resultados se debe tomar en cuenta el tiempo que le toma a los modelos entrenar y predecir, como se aprecia en la tabla, al modelo mDeBERTa le tomó menos la ejecución de las predicciones, sin embargo, fue el modelo más impreciso en la determinación de cambio de autor según el resto de las métricas. Como adicional, se muestra la matriz de confusión de los modelos, la cual detalla la información expuesta en la tabla anterior acorde a las predicciones realizadas, mostrando de manera gráfica los resultados verdaderos positivos y verdaderos negativos.

Pol. Con. (Edición núm. 85) Vol. 8, No 8, Agosto 2023, pp. 2552-2565, ISSN: 2550 - 682X

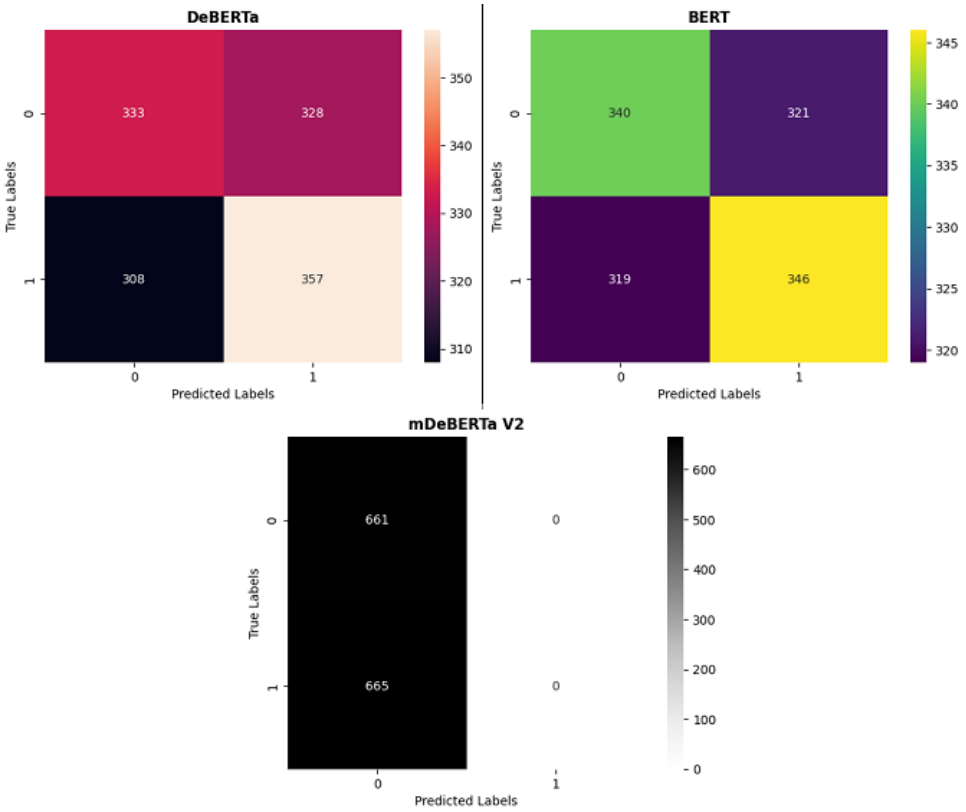

Figura 6 Matrices de confusión de la predicción

Discusión

Con base a las métricas de evaluación seleccionadas para el presente trabajo, se realiza una comparación de los resultados obtenidos para determinar el desempeño de cada uno, demostrando que BERT tuvo un rendimiento del 51.73% mientras que DeBERTa obtuvo un 52.03% y por último mDeBERTa con un 33.26% en la predicción de cambio de autor en un texto, por otro lado, este último modelo muestra un tiempo de ejecución de 01:44h y valor de pérdida del 49.84%, ambos valores menores que los obtenidos en los modelos BERT y DeBERTa.

El dataset utilizado cuenta con 8836 datos los cuales tienen entradas con textos largos que superan los 512 tokens admitidos por los modelos utilizados, lo que puede afectar el tratamiento de la data, el rendimiento del entrenamiento y la predicción de los valores, a lo que se sugiere utilizar métodos apropiados para el tratamiento de textos largos.

La extracción de embeddings de los tokens iniciales de las capas de atención es un enfoque novedoso y poco implementado al momento de determinar el cambio de autor en un texto, este

enfoque puede llegar a tener efectividad en este tipo de tareas, debido a que se extrae la características semánticas y sintácticas de los textos que pueden ayudar al aprendizaje del modelo.

Conclusiones

Para esta investigación se planteó un método de aprendizaje automático que permita determinar el cambio de autor en un texto mediante la codificación de capas de atención de modelos basados en BERT, donde se demuestra que los modelos BERT y DeBERTa logran rendimientos similares en cuanto a la predicción de cambio de autor en textos de idioma inglés, teniendo una ligera ventaja en el modelo DeBERTa con un 52.03% evidenciando así que las experimentaciones realizadas siembran una base interesante para el tipo de tarea elegido con el enfoque de extracción de embeddings de tokens iniciales en las capas de atención.

Para trabajos futuros se recomienda experimentar con otros modelos y diferentes métodos para la identificación de autor como la extracción de características estilométricas de los textos con la finalidad de obtener un mejor rendimiento de los modelos.

Si bien los resultados no fueron los esperados en la determinación de cambio de autor en un texto, el método propuesto de extraer la información sintáctica y semántica de los embbedings iniciales de las capas de atención, abre paso a nuevas investigaciones que puedan adoptar el enfoque a la determinación de cambio de autor en un texto.

Referencias

Avram, S.-M. (2023). *BERT-based Authorship Attribution on the Romanian Dataset called ROST*. *January*, 1–18. http://arxiv.org/abs/2301.12500

Barlas, G., & Stamatatos, E. (2020). Cross-domain authorship attribution using pre-trained language models. *IFIP Advances in Information and Communication Technology*, *583 IFIP*, 255–266. https://doi.org/10.1007/978-3-030-49161-1_22/FIGURES/2

Beltrán, N. C., & Rodríguez Mojica, E. C. (2021). Procesamiento del lenguaje natural (PLN) - GPT-3.: Aplicación en la Ingeniería de Software. *Tecnología Investigación y Academia*, *8*(1), 18–37. https://revistas.udistrital.edu.co/index.php/tia/article/view/17323

Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Naacl-Hlt 2019*, *Mlm*, 4171–4186.

Fabien, M., Villatoro-Tello, E., Motlicek, P., & Parida, S. (n.d.). *BertAA: BERT fine-tuning for Authorship Attribution*. 127–137.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep Learning-Based Text Classification. *ACM Computing Surveys*, *54*(3). https://doi.org/10.1145/3439726

Singh, R. (2022). *Utilizing Transformer Representations Efficiently | Kaggle*. https://www.kaggle.com/code/rhtsingh/utilizing-transformer-representations-efficiently/notebook

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*, *2017-Decem*(Nips), 5999–6009.

Xia, F., Li, B., Weng, Y., He, S., Sun, B., Li, S., Liu, K., & Zhao, J. (n.d.). *LingJing at SemEval-2022 Task 3: Applying DeBERTa to Lexical-level Presupposed Relation Taxonomy with Knowledge Transfer*. 239–246. Retrieved June 12, 2023, from https://sites.google.com/view/semeval2022-pretens/