UNIVERSIDAD NACIONAL DE SAN MARTIN

RCSI
https://revistas.unsm.edu.pe/index.php/rcsi
e-ISSN: 2709-992X

Artículo Original
Original Article
Jul-Dic, 2025

Un marco de atención dual contrastiva para mejorar la extracción de relaciones entre eventos adversos de medicamentos

A contrastive dual attention framework for enhancing adverse drug event relations extraction

Kashtriya, Poonam[1]*

Singh, Pardeep[1]

[1]National Institute of Technology, Hamirpur, India

Received: 15 Feb. 2025 | **Accepted:** 20 May. 2025 | **Published:** 20 Jul. 2025

Corresponding author*: Poonam_phdcse@nith.ac.in

How to cite this article: Kashtriya, K. & Singh, P. (2025). Un marco de atención dual contrastiva para mejorar la extracción de relaciones entre eventos adversos de medicamentos. *Revista Científica de Sistemas e Informática*, *5*(2), e968. https://doi.org/10.51252/rcsi.v5i2.968

RESUMEN

La extracción precisa de relaciones entre fármacos y eventos adversos a medicamentos (ADE) es fundamental para mejorar la seguridad del paciente. Sin embargo, los enfoques actuales tienen dificultades para captar relaciones complejas debido a limitaciones en la representación contextual. En el conjunto de datos n2c2, las instancias ADE-Fármaco (1107) son considerablemente menos numerosas que otras como Fuerza-Fármaco (6702) o Razón-Fármaco (5169), lo que introduce un fuerte desequilibrio que complica su identificación. Se emplea un modelo basado en codificadores de transformadores para generar representaciones contextuales, incorporando un mecanismo de atención dual que enfoca tanto en las entidades como en su entorno clínico. A través del aprendizaje contrastivo, se refina la representación de los pares de entidades, diferenciando con mayor precisión las relaciones correctas de las incorrectas. En las evaluaciones experimentales, se alcanzó un F1 general del 93,31 % y un 83,31 % en la relación ADE-Fármaco, superando a métodos previos. La combinación de codificación contextual, atención especializada y discriminación contrastiva permite afrontar con mayor eficacia los desafíos derivados del desequilibrio de clases y de la complejidad semántica del lenguaje clínico.

Palabras clave: eventos adversos de medicamentos; PNL biomédica; aprendizaje contrastivo; atención dual; aprendizaje automático; procesamiento del lenguaje natural

ABSTRACT

Accurate extraction of relationships between drugs and adverse drug events (ADEs) is essential for improving patient safety. However, current approaches struggle to capture complex relationships due to limitations in contextual representation. In the n2c2 dataset, ADE-Drug instances (1107) are significantly fewer than others such as Strength-Drug (6702) or Reason-Drug (5169), creating a strong class imbalance that hinders identification. A model based on transformer encoders is used to generate contextual embeddings, incorporating a dual attention mechanism that focuses on both the entities and their clinical context. Contrastive learning refines the representation of entity pairs, enabling more precise differentiation between correct and incorrect relationships. Experimental evaluations show a general F1 score of 93.31% and 83.31% for the ADE-Drug relation, outperforming previous methods. The combination of contextual encoding, specialized attention, and contrastive discrimination effectively addresses the challenges of class imbalance and the semantic complexity of clinical language.

Keywords: adverse drug events; biomedical NLP; contrastive learning; dual attention; machine learning; natural language processing

© Los autores. Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia Creative Commons Atribución, que permite el uso, distribución y reproducción sin restricciones en cualquier medio, siempre que se cite correctamente la obra original.

1. INTRODUCCIÓN

Los EAM se definen como respuestas dañinas o involuntarias a los medicamentos. Representan un desafío significativo para la seguridad del paciente y los sistemas de salud a nivel mundial. Los EAM representan una proporción sustancial de hospitalizaciones, reingresos y costos de atención médica, y estudios estiman que contribuyen a más de 100,000 muertes anuales solo en Estados Unidos (IOM, 2007). La detección temprana y la mitigación de los EAM son cruciales para reducir su impacto, lo que requiere un monitoreo integral de la seguridad de los medicamentos (OMS, 2008).

Los textos clínicos, como las historias clínicas electrónicas (HCE), los resúmenes de alta y las notas médicas, sirven como valiosos repositorios para la detección de eventos adversos de la atención (EAA). Estos registros suelen documentar detalles esenciales sobre el historial médico del paciente, los medicamentos prescritos y cualquier reacción adversa. Las bases de datos clínicas a gran escala como i2b2 (Murphy et al., 2010) y MIMIC-III (Johnson et al., 2016) han facilitado avances en la extracción de información médica relevante. Sin embargo, estos textos siguen siendo difíciles de analizar debido a su naturaleza no estructurada, su lenguaje específico y su terminología dependiente del contexto (Roberts et al., 2009).

La extracción de relaciones (RE), la tarea de identificar relaciones entre entidades, es un paso clave para aprovechar los textos clínicos para la detección de EAM. En el contexto de los EAM, la extracción de relaciones se centra en vincular los fármacos con sus eventos adversos asociados. Por ejemplo, en la frase "El paciente experimentó mareos después de tomar metformina", la tarea consiste en establecer la relación "EAM causado por el fármaco" entre "Metformina" y "mareos". Esta capacidad es esencial para tareas posteriores como la farmacovigilancia, donde la identificación precisa de los riesgos relacionados con los fármacos puede fundamentar las acciones regulatorias, y para los sistemas de apoyo a la toma de decisiones clínicas que ayudan a prevenir resultados adversos.

A pesar de los avances en PLN, los enfoques existentes para la extracción de relaciones ADE presentan limitaciones significativas. Los métodos tradicionales suelen tener dificultades para captar las sutiles relaciones entre fármacos y eventos adversos, especialmente en casos con asociaciones indirectas o estructuras oracionales complejas. Además, los modelos actuales suelen carecer de la capacidad de integrar eficazmente tanto el contexto local (p. ej., las interacciones entre fármacos y entidades ADE) como el contexto global (p. ej., la narrativa clínica general), lo que resulta en un rendimiento deficiente en situaciones reales. Para abordar los desafíos persistentes en la identificación precisa de las relaciones ADE-fármaco dentro de narrativas clínicas complejas, exploramos una nueva dirección de modelado que va más allá del aprendizaje de representación tradicional. Al alinear el enfoque del modelo con la estructura intrínseca del texto clínico, buscamos captar características sutiles ADE-fármaco que a menudo se pasan por alto. Este trabajo destaca la necesidad crucial de una extracción precisa e interpretable de las relaciones ADE, un área que sigue siendo fundamental para mejorar la seguridad clínica y los sistemas de farmacovigilancia.

En este artículo, proponemos un innovador marco integral de extracción de relaciones ADE que aprovecha tanto el aprendizaje contrastivo como un mecanismo de atención dual. Este mecanismo se centra por separado en las interacciones de pares de entidades y en el contexto clínico más amplio para captar mejor las relaciones ADE-fármaco con matices. El aprendizaje contrastivo

refina aún más el espacio de incrustación al acercar los pares ADE-fármaco y alejar los pares no relacionados, lo que mejora el rendimiento de la clasificación. Para abordar el desequilibrio de clases y la variabilidad lingüística, ampliamos los datos con un DistilGPT-2 optimizado, generando ejemplos sintéticos específicos del dominio. Evaluaciones exhaustivas del conjunto de datos n2c2 2018 (Henry et al., 2020) demuestran que nuestro método presenta una mejora significativa con respecto a los parámetros existentes para la extracción de la relación ADE-fármaco.

2. TRABAJOS RELACIONADOS

La RE en PNL clínica busca identificar relaciones significativas entre entidades en textos clínicos no estructurados, como HCE, resúmenes de alta y notas clínicas. Esto es esencial para aplicaciones como la farmacovigilancia, los sistemas de apoyo a la toma de decisiones clínicas y la monitorización de ADE. La RE implica vincular entidades con relaciones semánticas específicas, como "Medicamento-Causado-ADE" o "Medicamento-Trata-Afección", para apoyar a los profesionales sanitarios y automatizar la extracción de información útil, reduciendo así la necesidad de revisión manual de datos. Sin embargo, existen numerosos desafíos, como la naturaleza no estructurada de los datos clínicos, la ambigüedad y variabilidad del lenguaje, la complejidad de las estructuras oracionales y la necesidad de comprender la terminología médica específica del dominio. Los enfoques existentes abarcan desde métodos basados en reglas y aprendizaje automático (ML) hasta aprendizaje profundo (DL) y grandes modelos de lenguaje (LLM), incluyendo modelos de lenguaje preentrenados como BERT y ClinicalBERT (Devlin et al., 2018; Alsentzer et al., 2019; Lee et al., 2019; Sun et al., 2019). La investigación que aprovecha modelos adaptados al dominio como BioBERT (Gurulingappa et al., 2012) y ERNIE (Liu et al., 2016) ha demostrado avances significativos. Se han explorado estrategias basadas en reglas (Brown et al., 2020), redes neuronales convolucionales (CNN) para la extracción de interacciones fármaco-fármaco (Vig, 2019a) y métodos basados en LLM (Zhou et al., 2016), lo que indica un amplio espectro de técnicas para la monitorización de ADE.

Recientemente se han publicado dos revisiones detalladas sobre ADE. Modi et al. (2024) realizaron una revisión exhaustiva que categorizó las técnicas de extracción de ADE en enfoques basados en reglas, aprendizaje automático (ML), aprendizaje a distancia (DL), híbridos y LLM. Su trabajo enfatiza las tareas duales de reconocimiento de entidades nombradas (NER) y RE, analizando su rendimiento en marcos de aprendizaje de canalización, tareas conjuntas y multitarea, a la vez que identifican desafíos como inconsistencias en las anotaciones y conjuntos de datos desequilibrados. Li et al. (2024), por otro lado, se centraron en la aplicación de métodos de ML y DL específicamente en conjuntos de datos de referencia como los desafíos n2c2 de 2018 y MADE 1.0. Demostraron el rendimiento superior de los modelos basados en transformadores, incluyendo BERT y sus variantes, particularmente en configuraciones de conjunto, y destacaron áreas clave de mejora, como el manejo de desajustes de límites y conjuntos de datos desequilibrados en tareas de NER y RE. Ambos estudios subrayan la importancia de enfoques innovadores para mejorar la detección de ADE y la farmacovigilancia, con énfasis distintivo en la amplitud metodológica y las optimizaciones específicas del modelo.

Recientemente se han aplicado diversos enfoques a la tarea de ADE RE. Kim y Meystre (2020) propusieron un sistema de PLN que emplea un método de conjunto apilado que integra varios clasificadores secuenciales, como Campos Aleatorios Condicionales (CRF), redes de Memoria

Bidireccional a Largo Plazo (Bi-LSTM) y algoritmos de predicción estructurada basados en SEARN, para tareas de extracción de conceptos y clasificación de relaciones. Su trabajo destacó la eficacia de combinar múltiples modelos de predicción estructurada con máquinas de vectores de soporte (SVM) para extraer entidades relacionadas con medicamentos y ADE de textos clínicos. De forma similar, Christopoulou et al. (2020) introdujeron un enfoque de aprendizaje profundo de conjunto que abordó por separado la extracción de relaciones intra-oracionales e inter-oracionales utilizando Bi-LSTM con mecanismos de atención y redes Transformer, respectivamente. Su modelo abordó eficazmente las dependencias tanto en contextos cortos como largos, lo que permitió una extracción de relaciones robusta sin depender de herramientas sintácticas externas. Yang et al. (2020) propusieron un enfoque híbrido que incorpora redes neuronales convolucionales recurrentes (RCNN) para NER y exploraron modelos de aprendizaje automático, como SVM, Random Forests y Gradient Boosting, para la clasificación de relaciones. Su trabajo también integró la incorporación de conocimientos médicos para mejorar la identificación de relaciones entre medicamentos y ADE en narrativas clínicas. Wei et al. (2020) desarrollaron una canalización basada en aprendizaje profundo donde se utilizaron modelos Bi-LSTM-CRF para NER y se exploraron arquitecturas CNN-RNN para la clasificación de relaciones. Además, propusieron un modelo de aprendizaje conjunto para identificar simultáneamente entidades y sus relaciones, mitigando los problemas de propagación de errores inherentes a los enfoques tradicionales de canalización. Por el contrario, El-allaly et al. (2021) abordaron desafíos complejos como las relaciones multi-cabeza y las menciones discontinuas mediante la introducción de un método basado en el aprendizaje por transferencia multitarea (MTTLADE). Su sistema aprovechó modelos de lenguaje preentrenados basados en Transformers y un enfoque de etiquetado de secuencias de doble tarea para extraer simultáneamente menciones de ADE y sus relaciones, ofreciendo un marco unificado para abordar tareas complejas de extracción de ADE. A pesar de estos avances, persisten desafíos en la comprensión contextual, la generalización y la explicabilidad. Estudios previos han demostrado que la integración del conocimiento del dominio (Y. Zhang et al., 2020), la mejora de la interpretabilidad mediante la visualización de la atención (Vig, 2019b) y el aprovechamiento de incrustaciones avanzadas (Ji et al., 2019; Peng et al., 2018) pueden mejorar el rendimiento. Sin embargo, la captura de claves sutiles de entidad-contexto en las narrativas clínicas sigue siendo un problema importante.

El aprendizaje contrastivo, una técnica auto supervisada, ha recibido gran atención en el ámbito del procesamiento del lenguaje natural (PLN) y la visión artificial por su capacidad para aprender representaciones robustas y significativas mediante el entrenamiento de modelos para distinguir entre pares similares (positivos) y diferentes (negativos) (Zhang et al., 2019; Thapa et al., 2022), mejorando así la calidad de las incrustaciones para tareas posteriores como la clasificación, la agrupación en clústeres y la RE. En esencia, el aprendizaje contrastivo opera comparando muestras dentro de un espacio de incrustación compartido, lo que motiva al modelo a maximizar la similitud de los pares positivos (p. ej., entidades o contextos relacionados) y minimizar la de los pares negativos (p. ej., entidades no relacionadas), garantizando así que las incrustaciones capturen las relaciones semánticas subyacentes. En el contexto de la PNL, este enfoque es particularmente eficaz para tareas que involucran relaciones entre entidades, como la extracción de ADE (Thapa et al., 2022); por ejemplo, en una oración clínica como "El paciente experimentó mareos después de tomar metformina", el modelo aprende a asociar el medicamento "Metformina"

con el ADE "mareo", al tiempo que lo distingue de entidades no relacionadas, como otros medicamentos o afecciones.

Los conceptos clave del aprendizaje contrastivo incluyen los pares positivos y negativos, donde los pares positivos comparten una relación predefinida, como los pares fármaco-ADE, y los pares negativos carecen de esta relación, lo que permite al modelo diferenciarse eficazmente; el espacio de incrustación, donde los pares de entrada se proyectan en un espacio de alta dimensión con los pares positivos posicionados más cerca y los negativos más lejos; y la función de pérdida, típicamente una pérdida contrastiva como InfoNCE, que optimiza la similitud de los pares positivos y penaliza los negativos. En aplicaciones clínicas de PLN, el aprendizaje contrastivo ha mostrado resultados prometedores en RE al refinar las incrustaciones para pares de entidades, mejorar la precisión en la identificación de relaciones entre fármacos, ADE y otras entidades clínicas, mejorar la representación contextual al centrarse en las partes más relevantes del texto y aprovechar el aumento de datos para simular diversos contextos clínicos, mejorando así la generalización del modelo. Las ventajas del aprendizaje contrastivo en PLN clínico incluyen la eficiencia del etiquetado, ya que puede utilizar grandes cantidades de datos sin etiquetar, lo que lo hace ideal para dominios de la salud donde los datos etiquetados son limitados; Robustez, al modelar explícitamente similitudes y disimilitudes, lo que resulta en incrustaciones más resistentes a entradas ruidosas o ambiguas; y generalización, ya que las representaciones aprendidas suelen ser transferibles entre tareas, lo que reduce la necesidad de ajustes específicos para cada tarea. Específicamente, en la extracción de ADE, el aprendizaje contrastivo refina las incrustaciones de pares fármaco-ADE al contrastarlos con pares sin ADE. Por ejemplo, al usar un par positivo como ("Metformina", "mareo") con la relación "Fármaco-Causa-ADE" y un par negativo como ("Metformina", "dolor de cabeza") sin relación causal, se mejora la precisión de la clasificación de relaciones y la interpretabilidad al estructurar el espacio de incrustación de forma semánticamente significativa. Técnicas como SimCSE (Gao et al., 2021) han demostrado mejoras en las incrustaciones a nivel de oración, mientras que los enfoques adaptados al dominio (Peng et al., 2020) aprovechan texto médico sin etiquetar para mejorar la generalización. Mediante la ampliación de datos (Kumar et al., 2020), el aprendizaje contrastivo puede superar el desequilibrio de clases y mejorar la discriminación de relaciones sutiles. El funcionamiento general del aprendizaje contrastivo se muestra en la Figura 1.

El mecanismo de atención (Devlin et al., 2018; Alsentzer et al., 2019) ha revolucionado el PLN al permitir que los modelos se centren en las partes más relevantes del texto de entrada, lo cual es especialmente esencial para tareas de RE donde el contexto y las interacciones entre entidades son cruciales. La autoatención, introducida en la arquitectura Transformer, evalúa las relaciones entre cada token de una secuencia para crear representaciones contextualizadas, lo que permite a los modelos capturar dependencias de largo alcance y priorizar dinámicamente segmentos de entrada importantes. Por ejemplo, en la frase «El paciente experimentó mareos después de tomar metformina», la autoatención identifica la relación causal entre «mareos» y «metformina» centrándose en palabras clave como «experimentó» y «después». Partiendo de esta base, la atención dual (Zhang et al., 2020) mejora la capacidad del modelo al incorporar dos mecanismos de atención distintos pero complementarios: la atención de pares de entidades y la atención contextual. La Atención de Pares de Entidades se centra específicamente en las interacciones entre pares predefinidos de entidades, como un fármaco y un evento adverso, capturando así relaciones precisas y específicas de cada entidad con alta fidelidad. Simultáneamente, la Atención Contextual

examina el contexto textual más amplio que rodea a estas entidades, incluyendo síntomas, diagnósticos o cronogramas de tratamiento relacionados, para recopilar información adicional que podría influir en la relación entre ellas. Este enfoque dual permite al modelo equilibrar la granularidad de las interacciones de las entidades con la riqueza del contexto circundante, lo que resulta en una comprensión más completa de los datos. En aplicaciones clínicas de PLN, como la detección de ADE, los mecanismos de atención dual son particularmente ventajosos, ya que permiten aislar e interpretar eficazmente relaciones complejas dentro de narrativas clínicas extensas y complejas.

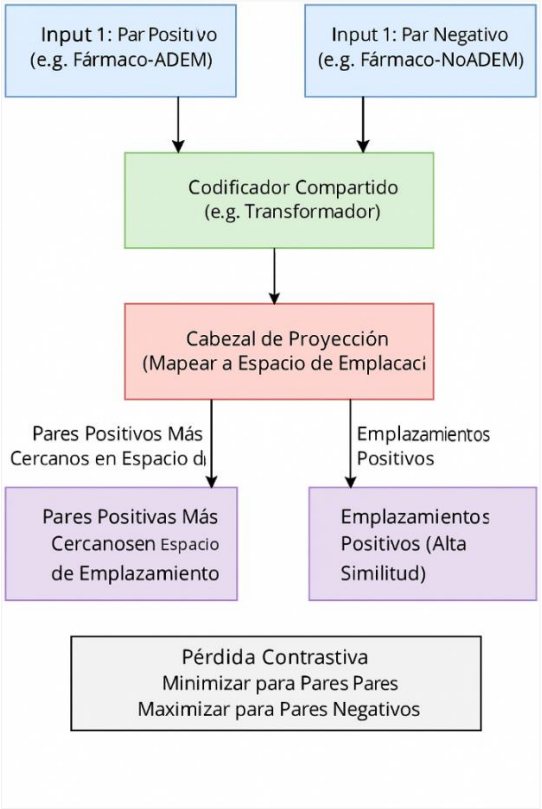

Figura 1. Aprendizaje contrastivo

Los modelos basados en transformadores, como ClinicalBERT y BioBERT, aprovechan la autoatención para extraer relaciones fármaco-ADE. La integración de la atención dual refina aún más este proceso al centrarse por separado en las interacciones entre entidades y la información contextual, mejorando así el rendimiento del modelo. En escenarios que involucran múltiples entidades dentro de una misma oración, la atención dual facilita la desambiguación de relaciones al aislar pares de entidades relevantes y considerar su contexto, lo que reduce la probabilidad de asociaciones incorrectas. Las ventajas de la atención dual van más allá de un mejor rendimiento; también mejoran la interpretabilidad al proporcionar información clara sobre qué pares de entidades y elementos contextuales influyeron en las decisiones del modelo, lo que fomenta una mayor confianza y fiabilidad en entornos clínicos. Los mecanismos de atención dual ofrecen escalabilidad, lo que permite a los modelos procesar y analizar eficientemente grandes cantidades de datos clínicos complejos sin sacrificar el rendimiento.

Al captar indicios contextuales tanto locales como globales (Johnson et al., 2020), la atención dual garantiza que incluso las relaciones más sutiles y matizadas dentro del texto se reconozcan y utilicen, lo que favorece resultados de RE más robustos y fiables. Contribuye a la robustez de los

modelos al permitirles adaptarse a diversos niveles de complejidad contextual y densidad de entidades, comunes en la documentación clínica del mundo real. La evolución continua promete mejorar aún más la precisión, la eficiencia y la interpretabilidad de los mecanismos de atención dual, garantizando su continua relevancia y eficacia para abordar los complejos desafíos inherentes a la extracción de relaciones significativas de los datos biomédicos. La Figura 2 describe el mecanismo de atención dual.

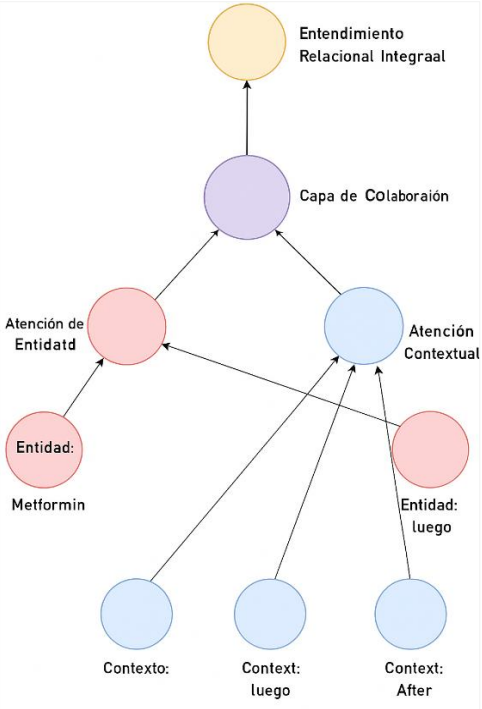

Figura 2. Ejemplo de mecanismo de atención dual

3. METODOLOGÍA

3.1. Enfoque propuesto

Este enfoque integra mecanismos de atención dual con aprendizaje contrastivo para mejorar la extracción de eventos adversos a medicamentos (ADE) en tareas de PLN clínico, al refinar simultáneamente las incrustaciones semánticas y enfocarse en interacciones significativas entre entidades y contexto. Comenzamos aplicando atención dual: una capa de atención captura específicamente las relaciones entre pares de entidades (atención a pares de entidades) y otra capa contextualiza estas entidades dentro de su entorno textual más amplio (atención contextual). Las salidas de estos dos módulos de atención complementarios generan incrustaciones distintas pero sinérgicas. Las incrustaciones de pares de entidades representan la conexión intrínseca entre un fármaco (por ejemplo, "Metformina") y un posible ADE (por ejemplo, "mareos"), mientras que las incrustaciones contextuales codifican la narrativa circundante (por ejemplo, "experimentados después de"). Para unificar estas señales y mejorar su calidad semántica, primero concatenamos las incrustaciones del par de entidades y del contexto, y luego alimentamos este vector combinado a una pequeña red de proyección compuesta por capas completamente conectadas con activaciones no lineales. Este paso de proyección no solo reduce la dimensionalidad y estabiliza el entrenamiento (mediante normalización), sino que también integra señales relacionales y contextuales en un espacio de incrustación cohesivo. Dentro de este espacio unificado, los objetivos de aprendizaje contrastivo (por ejemplo, InfoNCE) fomentan que los pares positivos

instancias relacionadas conocidas como (Metformina, mareos) se agrupen, mientras que los pares no relacionados se separen, asegurando así que las incrustaciones derivadas de nuestros mecanismos de atención dual sean tanto específicas para la tarea como estructuralmente robustas.

Optimizamos conjuntamente los parámetros del modelo utilizando una función de pérdida híbrida compuesta por una pérdida de clasificación, derivada del objetivo de extracción de relaciones, y una pérdida contrastiva que mejora continuamente la geometría del espacio de incrustación. Durante el entrenamiento, cada oración clínica de entrada, por ejemplo: "El paciente experimentó mareos después de tomar Metformina", atraviesa las canalizaciones de atención dual para generar incrustaciones fusionadas y proyectadas, las cuales luego se refinan mediante objetivos contrastivos. Esto da como resultado un modelo cohesivo que sobresale en la detección de ADE mediante una identificación precisa de relaciones impulsada por atención, y una mejor calidad semántica de las incrustaciones obtenidas a través del aprendizaje contrastivo. Este enfoque evalúa implícitamente si existe alguna relación significativa entre las entidades identificadas antes de clasificar la naturaleza de dicha relación. Esta funcionalidad se integra sin interrupciones en el modelo mediante el uso combinado de aprendizaje contrastivo y atención dual. El aprendizaje contrastivo estructura el espacio de incrustación de tal manera que los pares de entidades con relaciones válidas (por ejemplo, Fármaco-Causa-ADE) se acerquen entre sí, mientras que los pares no relacionados se distancien. Esto garantiza que el modelo distinga inherentemente entre pares con relevancia relacional y aquellos sin ninguna. Además, el mecanismo de atención dual amplifica este proceso de filtrado al enfocarse tanto en la interacción entre el par de entidades (atención a pares de entidades) como en el contexto clínico circundante (atención contextual), capturando señales sutiles que indican la presencia de una relación. Las formulaciones matemáticas se presentan a continuación y la representación gráfica de nuestro enfoque se muestra en la Figura 3.

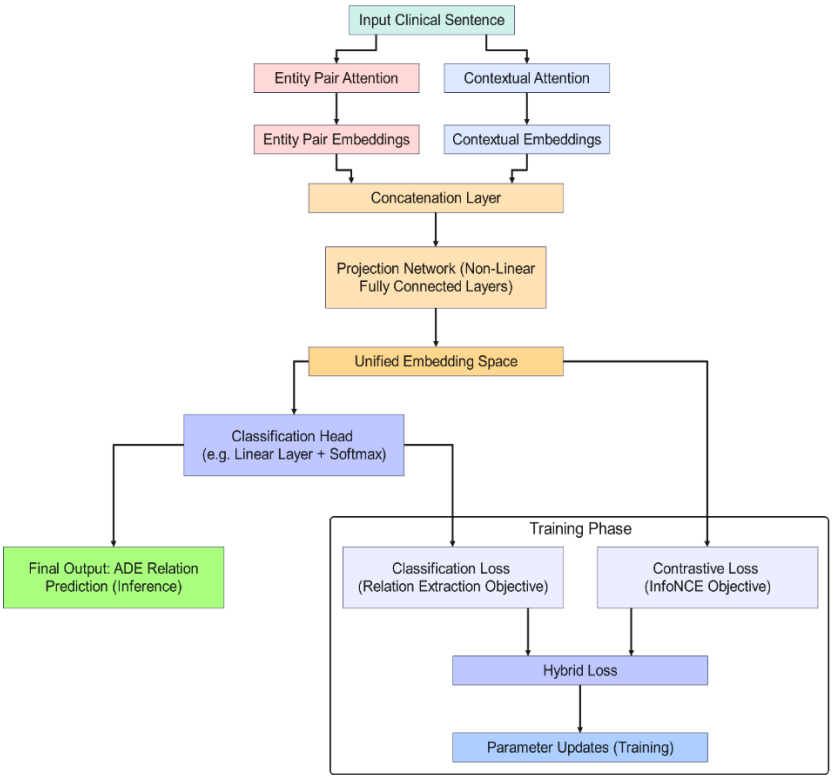

Figura 3. Flujo de trabajo del enfoque propuesto

- Mecanismos de Atención Dual

Sean las incrustaciones de una oración clínica de entrada con tokens, cada uno representado como un vector de dimensión α. Supongamos que hemos identificado dos entidades de interés (por ejemplo, un fármaco y un posible ADE) dentro de la oración.

- Atención del par de entidades

Defina los parámetros $\mathbf{W}_Q^e, \mathbf{W}_K^e, \mathbf{W}_V^e$ para la atención del par de entidades, donde $\mathbf{x}_e$ representa la incrustación combinada del par de entidades. Proyectamos la representación de la entidad en el espacio de consulta y la oración completa en el espacio clave-valor:

$$Q_e = \mathbf{W}_Q^e \mathbf{x}_e \tag{1}$$

$$K_e = \mathbf{W}_K^e \mathbf{X} \tag{2}$$

$$V_e = \mathbf{W}_V^e \mathbf{X} \tag{3}$$

Salida de atención del par de entidades $\mathbf{H}_{entity}$ se calcula como:

$$\mathbf{H}_{entity} = \text{Attention}(Q_e, K_e, V_e) = \text{softmax}\left(\frac{Q_e K_e^{\bullet}}{\sqrt{d_k}}\right) V_e \tag{4}$$

Aquí, d_k está la dimensión de las teclas.

- Atención contextual

De manera similar, defina parámetros $\mathbf{W}_Q^c, \mathbf{W}_K^c, \mathbf{W}_V^c$ para la atención contextual:

$$Q_c = \mathbf{W}_Q^c \mathbf{X} \tag{5}$$

$$K_c = \mathbf{W}_K^c \mathbf{X} \tag{6}$$

$$V_c = \mathbf{W}_V^c \mathbf{X} \tag{7}$$

La salida de atención contextual $\mathbf{H}_{context}$ es:

$$\mathbf{H}_{context} = \text{Attention}(Q_c, K_c, V_c) = \text{softmax}\left(\frac{Q_c K_c^{\bullet}}{\sqrt{d_k}}\right) V_c \tag{8}$$

Después de obtener estos resultados, normalmente agrupamos o seleccionamos las partes relevantes de $\mathbf{H}_{entity}$ and $\mathbf{H}_{context}$ para formar incrustaciones de tamaño fijo. Dejar $\mathbf{h}_{entity}$ and $\mathbf{h}_{context}$ sean estas incrustaciones agrupadas resultantes.

- Fusión y proyección de incrustaciones

Concatenamos la incrustación del par de entidades y la incrustación contextual:

$$\mathbf{z} = [\mathbf{h}_{entity}; \mathbf{h}_{context}] \in \square^{2d} \tag{9}$$

Luego proyectarnos $\mathbf{z}$ en un espacio unificado utilizando una pequeña red de proyección (por ejemplo, una capa completamente conectada con activación no lineal):

$$\mathbf{z}' = \sigma(\mathbf{W}_{proj}\mathbf{z} + \mathbf{b}_{proj}) \tag{10}$$

Opcionalmente, aplicamos normalización de capas o normalización L2 para estabilidad:

$$\mathbf{z}'' = \text{LayerNorm}(\mathbf{z}') \tag{11}$$

La resultante $\mathbf{z}''$ integra señales relacionales y contextuales.

- Aprendizaje contrastivo

Para el aprendizaje contrastivo, supongamos que tenemos un conjunto de pares positivos y negativos. Dado un par positivo $(\mathbf{z}_i'', \mathbf{z}_j'')$ y negativos $\{\mathbf{z}_k''\}_{k \neq j}$, dado un par positivo es:

$$\mathrm{L}_{\text{contrastive}} = -\log\left(\frac{\exp\left(\frac{\mathbf{z}_i'' \cdot \mathbf{z}_j''}{\tau}\right)}{\sum_k \exp\left(\frac{\mathbf{z}_i'' \cdot \mathbf{z}_k''}{\tau}\right)}\right) \tag{12}$$

Donde $\tau > 0$ es un parámetro de temperatura y denota el producto escalar.

- Pérdida de clasificación para la extracción de relaciones

Supongamos que hay mapas de cabezas de clasificación $\mathbf{z}''$ a los logits $\mathbf{y}$, y dejar $\mathbf{y}^*$ sea la verdadera distribución de etiquetas. La pérdida de entropía cruzada es:

$$\mathrm{L}_{cls} = -\sum_c y_c^* \log(\hat{y}_c) \tag{13}$$

- Pérdida híbrida

Combinamos las pérdidas de clasificación y contraste en una única pérdida híbrida:

$$\mathrm{L}_{\text{hybrid}} = \alpha \mathrm{L}_{cls} + (1-\alpha)\mathrm{L}_{\text{contrastive}} \tag{14}$$

Donde $\alpha \in [0,1]$ es un factor de ponderación.

Minimizar $\mathrm{L}_{\text{hybrid}}$ durante el entrenamiento, se alienta al modelo a sobresalir tanto en la tarea de extracción de relaciones como en la formación de incrustaciones semánticamente significativas.

3.2. Preprocesamiento del conjunto de datos

Para este estudio, utilizamos el conjunto de datos Track 2 de la tarea compartida n2c2 de 2018, un referente ampliamente reconocido para la extracción de relaciones de ADE. Este conjunto de datos proporciona etiquetas de referencia que se centran en los nombres de los fármacos, los ADE y sus relaciones, lo que garantiza anotaciones de alta calidad para el entrenamiento y la evaluación. El conjunto de datos consta de 505 notas clínicas, con 303 notas asignadas al conjunto de entrenamiento, de las cuales el 78,20 % contiene relaciones de ADE, y 202 notas en el conjunto de prueba, de las cuales el 76,73 % contiene relaciones de ADE. Para mejorar el rendimiento del modelo, realizamos preprocesamiento y aumento de datos. Los pasos de preprocesamiento incluyeron la limpieza y tokenización del texto para garantizar la compatibilidad con nuestro marco, mientras que se emplearon técnicas de aumento para mitigar el desequilibrio de clases y enriquecer la diversidad de las muestras de entrenamiento, mejorando así la robustez del modelo en la extracción de relaciones de ADE.

- Aumento de datos

Las técnicas tradicionales de aumento de datos, como el reemplazo de sinónimos, la inserción aleatoria o la eliminación, a menudo no preservan la semántica específica del dominio ni la validez clínica requeridas en los conjuntos de datos biomédicos. Estos métodos pueden introducir variaciones sin sentido o clínicamente irrelevantes, lo que afecta negativamente el rendimiento

del modelo en tareas como la extracción de ADE. Para abordar estos desafíos y superar los desequilibrios de clase en el conjunto de datos, empleamos el aumento de datos contextual utilizando el modelo distilGPT2 preentrenado (Kim et al., 2024; Aguilar-Canto et al., 2023), una versión destilada de GPT-2 conocida por su eficiencia y capacidad generativa. DistilGPT2 se ajustó en un corpus biomédico para alinearse mejor con el contexto clínico. A diferencia de los métodos de aumento tradicionales (p. ej., el reemplazo de sinónimos), que a menudo introducen cambios médicamente imprecisos o sin sentido, nuestro aumento basado en DistilGPT2 aprovecha la conciencia contextual para generar alternativas plausibles y clínicamente válidas. Los tokens clave, como los que representan fármacos, ADE u otras palabras contextuales relevantes, se enmascararon o reemplazaron selectivamente según los resultados generativos del modelo. Por ejemplo, en la oración "El paciente experimentó mareos tras tomar metformina", distilGPT2 generó variaciones como "El paciente experimentó náuseas tras tomar metformina" o "El paciente experimentó vértigo tras tomar metformina" sustituyendo el término ADE (mareos) por alternativas contextualmente apropiadas. De igual forma, creó oraciones alternativas como "El paciente experimentó mareos tras tomar ibuprofeno" variando el término farmacológico (metformina). La capacidad de DistilGPT2 para generar variaciones plausibles se debió a su preentrenamiento con conjuntos de datos amplios y diversos, lo que le permitió mantener la coherencia semántica a la vez que introducía mejoras novedosas y relevantes. Cada frase aumentada generada se revisó y validó manualmente para garantizar su relevancia clínica e integridad, preservando así la calidad del conjunto de datos aumentados. Este enfoque no solo amplió el conjunto de datos sino que también mejoró el equilibrio de clases, facilitando un entrenamiento sólido para tareas posteriores como la clasificación de relaciones y manteniendo los altos estándares requeridos para las aplicaciones biomédicas.

- Candidatos a Relación

Para preparar el conjunto de datos para la extracción de relaciones de ADE, seguimos un proceso sistemático que garantiza la identificación y representación precisas de las relaciones entre entidades. Primero, generamos pares de entidades a partir de las entidades identificadas en cada oración para analizar sus posibles relaciones. Por ejemplo, a partir de la oración "El paciente experimentó mareos después de tomar metformina", se creó el par de entidades (Metformina, mareos). A continuación, etiquetamos estos pares de entidades para diferenciar entre aquellos con relaciones válidas de "Fármaco-Causado-ADE" (pares positivos) y aquellos sin relación causal o pares no relacionados (pares negativos). El conjunto de datos etiquetado se organizó en un formato estructurado con columnas como ID de Oración, Fármaco, ADE y Etiqueta, lo que permitió un entrenamiento y una evaluación eficientes. Para guiar el mecanismo de atención dual, marcamos las entidades en las oraciones con tokens especiales (p. ej., <DROGA> y <ADE>), lo que garantiza que el modelo pueda centrarse en estos términos críticos. Por ejemplo, la oración original «El paciente experimentó mareos después de tomar metformina» se transformó en «El paciente experimentó <ADE>mareos</ADE> después de tomar <DRUG>Metformina</DRUG>». Preparamos los datos para el aprendizaje contrastivo creando pares de oraciones, como se muestra en la Tabla 1. Los pares positivos consistieron en oraciones con las mismas entidades y relaciones válidas, como «El paciente experimentó <ADE>mareos</ADE> después de tomar <DRUG>Metformina</DRUG>» y «La metformina causó <ADE>mareos</ADE>». Los pares negativos incluyeron oraciones con entidades no relacionadas o sin una relación válida, como aquellas que presentan diferentes fármacos y ADE. Preparamos los datos para el mecanismo de

atención dual tokenizando las oraciones marcadas e incorporando indicadores de entidad, especificando las posiciones de las entidades en la secuencia de entrada. Este formato de entrada integral permitió al modelo centrarse simultáneamente en las interacciones de pares de entidades y su contexto textual más amplio, sentando las bases para la extracción de relaciones de ADE sensible al contexto.

Tabla 1. Preparación de los datos

ID	Sentencia	Medicamento	ADE	Etiqueta de clasificación	Sentencia emparejada	Etiqueta contrastante
1	The patient experienced <ADE>dizziness</ADE> after taking <DRUG>Metformin</DRUG>.	Metformina	mareos	Drug-Caused-ADE	Metformin caused <ADE>dizziness</ADE>	Positivo
2	The patient experienced <ADE>nausea</ADE> after taking <DRUG>Metformin</DRUG>.	Metformina	náuseas	Not Drug-Caused-ADE	The patient reported <ADE>headache</ADE> after taking <DRUG>Metformin</DRUG>.	Negativo
3	The patient experienced <ADE>dizziness</ADE> after taking <DRUG>Lisinopril</DRUG>.	Lisinopril	mareos	Drug-Caused-ADE	After prescribing <DRUG>Lisinopril</DRUG>, the patient experienced <ADE>dizziness</ADE>.	Positivo
4	The patient reported <ADE>headache</ADE> after taking <DRUG>Metformin</DRUG>.	Metformina	dolor de cabeza	Not Drug-Caused-ADE	Metformin caused <ADE>dizziness</ADE>	Negativo

4. EXPERIMENTACIÓN Y RESULTADOS

En nuestra configuración experimental, utilizamos una configuración de hardware compuesta por un procesador Intel i7, 32 GB de RAM y una GPU NVIDIA (RTX) con 12 GB de VRAM. Modificamos ClinicalBERT para incorporar mecanismos de atención dual y aprendizaje contrastivo para una mejor extracción de ADE. Se añadieron capas de atención dual para capturar tanto las interacciones entre entidades (Atención de Par de Entidades) como la información contextual más amplia que rodea a estas entidades (Atención Contextual).

Las salidas de estos módulos de atención se fusionan y se transmiten a través de una red de proyección, integrando señales relacionales y contextuales en un espacio de incrustación unificado. Nuestro proceso de preprocesamiento comenzó segmentando el texto clínico de entrada en párrafos de hasta cuatro oraciones, garantizando que el contexto presentado al modelo fuera lo suficientemente rico como para capturar relaciones matizadas y computacionalmente manejable.

Fundamentalmente, a cada oración dentro de estos párrafos seleccionados se le asignó una etiqueta contrastiva (Positiva o Negativa) que determinó cómo formamos pares para el objetivo de aprendizaje contrastivo. Nuestro criterio de selección para construir estos párrafos de cuatro oraciones se basó en la consistencia de las etiquetas contrastivas: si la primera oración tenía una etiqueta positiva y la segunda

también la exhibía, pero la tercera introducía una etiqueta negativa, esta inconsistencia nos llevó a truncar el párrafo en la segunda oración.

Las oraciones restantes (más allá de la segunda en este escenario) no se consideraron, y sus espacios correspondientes se rellenaron para mantener una forma de entrada fija. Al adherirnos a esta regla, garantizamos que cada entrada al modelo presentara una señal relacional consistente en términos de alineación contrastiva, aprovechando al mismo tiempo hasta cuatro oraciones de contexto cuando estuvieran disponibles.

Los mecanismos de atención dual de nuestro modelo se centraron entonces en las entidades de interés y sus claves contextuales en estas oraciones seleccionadas, produciendo incrustaciones centradas en la entidad y conscientes del contexto. Estas incrustaciones se refinaron posteriormente mediante un objetivo de aprendizaje contrastivo, que acercó los ejemplos positivos y separó los negativos en el espacio de incrustación. Los pseudocódigos que representan este flujo de trabajo se muestran en la página siguiente. El algoritmo 1 muestra la fase de entrenamiento mientras que el algoritmo 2 muestra la fase de inferencia.

Algoritmo 1: Inicialización y entrenamiento del modelo

Result: A trained Dual Attention model for ADE detection
Initialize: ClinicalBERT model;
Add Dual Attention Layers:
- Entity Pair Attention Layer
- Contextual Attention Layer

Add Projection Network (Fully Connected Layers with Non-linear Activations);
Add Classification Head (Linear Layer + Softmax);
Define loss functions:
- Classification Loss (e.g., CrossEntropy Loss)
- Contrastive Loss (e.g., InfoNCE)

Preprocessing Function: preprocess_input(clinical_text):
Split text into paragraphs of up to 4 sentences;
foreach *paragraph* **do**
    Assign contrastive labels (Positive/Negative) to sentences;
    if *label consistency* then
        Keep paragraph;
    else
        Truncate at inconsistent sentence;
    end
    Pad remaining slots to maintain fixed input shape;
end
return processed paragraphs and labels;

Function: dual_attention(input_sentence):
Compute Entity Pair Embeddings using Entity Pair Attention;
Compute Contextual Embeddings using Contextual Attention;
Concatenate the two embeddings;
return concatenated embeddings;

Function: project_embeddings(embeddings):
Pass embeddings through Fully Connected Layers with Non-linear Activations;
return unified embedding space;

Training Function: train_model(clinical_data):
foreach *epoch* **do**
    foreach *batch in clinical_data* **do**
        inputs, labels = preprocess_input(batch);
        attention_output = dual_attention(inputs);
        projected_embeddings = project_embeddings(attention_output);

```
        classification_logits = ClassificationHead(projected_embeddings);
        classification_loss = ComputeLoss(classification_logits, labels);
        contrastive_loss = ComputeContrastiveLoss(projected_embeddings, labels);
        hybrid_loss = classification_loss + contrastive_loss;
        Update model parameters using hybrid_loss;
    end
end
```

Algoritmo 2: Fase de inferencia

```
Result: Predict ADE relationships from unseen clinical text
Inference Function: infer_relation(input_sentence):
processed_input = preprocess_input(input_sentence);
attention_output = dual_attention(processed_input);
projected_embeddings = project_embeddings(attention_output);
classification_output = ClassificationHead(projected_embeddings);
return classification_output;

Main Function: main():
clinical_data = LoadDataset();
train_model(clinical_data);
relation_prediction = infer_relation("The patient experienced dizziness after taking Metformin.");
Print "Predicted ADE Relation:", relation_prediction;
```

4.1. Fase de interferencia

- NER contextualizado

En la fase de inferencia, realizamos una extracción de relaciones completa de extremo a extremo, comenzando con el procesamiento de texto clínico sin procesar mediante NER. Para garantizar una NER contextualizada de alta calidad, entrenamos un modelo BioBERT independiente en nuestro conjunto de datos aumentados específicamente para esta tarea. Este modelo identificó eficazmente entidades relevantes como fármacos y eventos adversos dentro de las narrativas clínicas, sentando las bases para la posterior clasificación de relaciones. Nuestro modelo NER basado en BioBERT alcanza un rendimiento robusto con una puntuación F1 promedio micro del 96,4 %, lo que subraya su fiabilidad en la extracción de representaciones de entidades precisas y contextualizadas, cruciales para la extracción de relaciones posterior.

- Métricas de Evaluación

Nos hemos centrado en la extracción de relaciones de extremo a extremo, una tarea más compleja y completa en comparación con la extracción de relaciones tradicional. A diferencia de los métodos convencionales que asumen entidades preidentificadas como entrada, la extracción de relaciones de extremo a extremo abarca tanto el reconocimiento de entidades como la identificación de relaciones. Esto es fundamental para tareas clínicas de PLN como la detección de ADE, ya que entidades como fármacos y eventos adversos deben identificarse con precisión antes de extraer sus relaciones. Al abordar tanto el reconocimiento de entidades como la clasificación de relaciones, nuestro enfoque se ajusta estrechamente a los escenarios clínicos del mundo real, donde las entidades preidentificadas rara vez están disponibles, lo que lo convierte en una solución más práctica.

Al evaluar modelos de clasificación, especialmente en situaciones de desequilibrio de clases, la precisión, la recuperación y la puntuación F1 son métricas fundamentales. La precisión cuantifica la proporción de casos positivos correctamente identificados entre todos los casos predichos como positivos. La

recuperación, por otro lado, refleja la capacidad del modelo para recuperar casos positivos reales del conjunto de datos. Para capturar el equilibrio entre estas dos medidas, se utiliza la puntuación F1, que representa su media armónica. Estas métricas se definen de la siguiente manera:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{15}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{16}$$

$$\text{F1 Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{17}$$

Para evaluar el rendimiento de nuestro modelo propuesto, empleamos la puntuación F1 Lenient, que ofrece una métrica de evaluación más flexible en comparación con la puntuación F1 Strict. La puntuación F1 Strict requiere una coincidencia exacta de todos los componentes, incluyendo la amplitud de las entidades, los tipos de entidad y las etiquetas de relación, lo que a menudo puede penalizar desviaciones menores, como una ligera desalineación de los límites de las entidades. En conjuntos de datos clínicos como n2c2 ADE 2018, donde la amplitud de las entidades puede ser ambigua o solaparse, la puntuación F1 Lenient resulta ser una métrica más robusta al permitir coincidencias parciales. Por ejemplo, si el tipo de relación se predice correctamente, pero existe una ligera desalineación en la amplitud de las entidades, la puntuación F1 Lenient aún lo considera un éxito parcial, lo que proporciona una representación más precisa de las capacidades del modelo.

Estudios recientes sobre extracción de relaciones, especialmente en el ámbito clínico, han empleado predominantemente el promedio F1 Micro o el promedio F1 Macro para la evaluación. La puntuación F1 Macro calcula la puntuación F1 de forma independiente para cada clase y luego obtiene el promedio no ponderado. Este enfoque trata a todas las clases por igual, independientemente de su frecuencia. Por otro lado, la puntuación Micro-F1 agrega las contribuciones de todas las clases sumando los verdaderos positivos, falsos positivos y falsos negativos de cada clase antes de calcular la puntuación F1. En consecuencia, el micro promedio pondera las clases según su soporte (es decir, el número de instancias verdaderas), lo que favorece el rendimiento en las clases frecuentes. Macro-F1 y Micro-F1 se calculan de la siguiente manera:

$$\text{F1}_{\text{Macro}} = \frac{1}{N} \sum_{i=1}^{N} \text{F1}_i \tag{18}$$

Donde N es el número de clase y F1_i es el F1 de la puntuación de la clase i.

$$\text{F1}_{\text{Micro}} = \frac{2 \cdot \sum_{i=1}^{N} TP_i}{2 \cdot \sum_{i=1}^{N} TP_i + \sum_{i=1}^{N} FP_i + \sum_{i=1}^{N} FN_i} \tag{19}$$

Donde TP_i, FP_i, FN_i son los conteos para cada clase i.

Si bien el Promedio Macro F1 destaca el rendimiento de las clases minoritarias, en ocasiones puede sobreestimar el efecto de las clases poco comunes o con ruido, lo que genera evaluaciones sesgadas. Dado el desequilibrio de clases inherente a los conjuntos de datos de ADE, donde las relaciones ADE-Fármaco son menores en comparación con otros, hemos elegido el Promedio Micro F1 como nuestra métrica de

evaluación principal. Proporciona una evaluación equilibrada del rendimiento general del modelo, considerando la distribución de frecuencias naturales de las categorías de relación.

Además del Promedio Micro F1 general, la Precisión y la Recall, también hemos evaluado el rendimiento de nuestro modelo en categorías de relación individuales utilizando F1, la Precisión y la Recall. Este análisis es esencial para evaluar la eficacia de nuestro enfoque propuesto, específicamente en las relaciones ADE-Fármaco, que son el enfoque principal de esta tarea. Si bien la mayoría de los trabajos previos solo han reportado el Promedio Micro F1, algunos estudios han presentado el rendimiento por categoría, ofreciendo una visión más profunda de las capacidades de sus modelos.

Para garantizar una comparación exhaustiva y justa, hemos incluido tanto el Promedio Micro F1 como las métricas por categoría en nuestra evaluación. Los resultados, presentados en la subsección Resultados, resaltan la comparación de nuestro enfoque para identificar relaciones ADE-Fármaco mientras mantenemos un buen desempeño en otras categorías de relaciones.

RESULTADOS

Para garantizar una evaluación exhaustiva y justa, comparamos el rendimiento de nuestro enfoque con investigaciones previas que reportaron el Micro Promedio F1 Lenient para la tarea de extracción de relaciones de extremo a extremo. También comparamos el rendimiento de cada categoría de relación con las investigaciones que lo han reportado individualmente.

La Tabla 2 presenta un resumen del rendimiento alcanzado por investigaciones previas, junto con las diversas técnicas y modelos que emplearon. Estos trabajos sirven como puntos de referencia críticos para la comparación, permitiéndonos contextualizar las mejoras logradas por nuestro método propuesto. Entre estos estudios, El-Allaly et al. reportaron el Micro Promedio F1 general más alto, del 91,25%, lo que demuestra un sólido rendimiento en el conjunto de datos n2c2 ADE 2018.

Sin embargo, al analizar el rendimiento del modelo para la categoría específica de relación ADE-Fármaco, la puntuación F1 correspondiente fue del 64,62%, lo que indica margen de mejora en la captura de esta relación crítica. En particular, una variante del enfoque de El-Allaly et al. que utilizó el modelo BlueBERT logró una puntuación F1 ligeramente mejor del 66,85 % para la categoría ADE-Droga, pero su promedio F1 micro general cayó al 90,34 %, lo que refleja un equilibrio entre el desempeño general y la mejora específica de la categoría.

Tabla 2. Comparación de rendimiento en términos de Lenient Micro Average F1

Estudios	Enfoques	P	R	F1
Kim & Meystre (2020)	Un método de conjunto apilado combina CRF, Bi-LSTM y SEARN para extraer medicamentos y ADE, con SVM utilizado para la clasificación de relaciones en textos clínicos.	92.52	81.88	86.88
Christopoulou et al. (2020)	Un enfoque de aprendizaje profundo de conjunto integra Bi-LSTM con mecanismos de atención para relaciones intra-oraciones y redes Transformer para dependencias entre oraciones.	92.64	83.18	87.65
Yang et al. (2020)	Un sistema híbrido utiliza RCNN para el reconocimiento de entidades con nombre y modelos de aprendizaje automático, como el aumento de gradiente y SVM, para la clasificación de relaciones, con integraciones de conocimiento médico.	91.87	85.93	88.80
Wei et al. (2020)	Se comparan el aprendizaje profundo y los métodos tradicionales utilizando Bi-LSTM-CRF para NER, CNN-RNN para la clasificación de relaciones y un modelo conjunto para la extracción simultánea de entidades y relaciones.	-	-	89.05

Estudios	Enfoques	P	R	F1
El-allaly et al. (2021).	Un método de aprendizaje por transferencia de múltiples tareas (MTTLADE) combina el aprendizaje de múltiples tareas y modelos de lenguaje basados en Transformer para manejar relaciones multi-cabeza y discontinuas en la extracción de ADE.	91,93	90,57	91,25
Enfoque propuesto	Integra mecanismos de atención dual para capturar tanto las relaciones entre entidades como su contexto más amplio. Combina incrustaciones basadas en la atención con aprendizaje contrastivo, refinando el espacio de incrustación para agrupar las relaciones significativas y separar las irrelevantes, mejorando así la detección de ADE y la clasificación de relaciones.	94,22	92,42	93,31

Para abordar esta brecha y mejorar el rendimiento, específicamente en las relaciones ADE-fármaco, nuestro enfoque propuesto integra mecanismos de atención dual con aprendizaje contrastivo para refinar las incrustaciones de entidades y optimizar la comprensión relacional. Como se muestra en la Tabla 3, que compara las puntuaciones F1 individuales para cada categoría de relación con estudios previos, nuestro enfoque logra una mejora significativa en el rendimiento.

En concreto, mejoramos la puntuación F1 de ADE-fármaco hasta un impresionante 83,31 %, lo que representa un aumento sustancial de más de 16,46 puntos porcentuales en comparación con el mejor rendimiento de ADE-fármaco reportado por El-allaly et al. (2021). Además, nuestro método alcanzó un micropromedio general de F1 del 93,31 %, superando la puntuación más alta anterior del 91,25 %.

Esta mejora destaca no solo la solidez de nuestro enfoque para identificar con precisión las relaciones de entidades en todas las categorías, sino también su capacidad para enfocarse y destacar en la extracción de la relación ADE-fármaco, que fue el objetivo principal de nuestro trabajo. El mapa de calor del rendimiento de nuestro enfoque y la comparación del rendimiento con enfoques anteriores se muestran en la Figura 4 y la Figura 5 respectivamente.

Tabla 3. Desempeño de la categoría de relación individual

Relación	Métrica	El-allaly et al. (2021)	Wei et al. (2020)	Kim & Meystre (2020)	Enfoque propuesto
ADE-Drug	P	66,28	-	48,46	**84,39**
	R	63,03	-	15,01	**82,27**
	F1	64,62	47,55	22,92	**83,31**
Strength-Drug	P	97,81	-	94,55	98,65
	R	93,87	-	95,71	94,13
	F1	95,80	97,20	95,13	96,33
Dosage-Drug	P	93,77	-	92,12	94,19
	R	94,88	-	91,58	95,25
	F1	94,32	93,53	91,85	94,72
Duration-Drug	P	83,01	-	83,14	87,23
	R	81,46	-	65,96	85,52
	F1	82,23	78,61	73,56	86,37
Frequency-Drug	P	96,81	-	95,09	97,46
	R	93,81	-	93,58	94,68
	F1	94,95	95,82	94,33	96,05
Form-Drug	P	96,36	-	96,96	97,74
	R	94,40	-	92,52	95,51
	F1	95,37	95,16	94,69	96,61
Route-Drug	P	95,06	-	93,87	96,02

Relación	Métrica	El-allaly et al. (2021)	Wei et al. (2020)	Kim & Meystre (2020)	Enfoque propuesto
	R	92,72	-	92,44	94,14
	F1	93,88	94,15	93,15	95,07
Reason-Drug	P	76,25	-	74,97	81,27
	R	79,91	-	34,87	82,53
	F1	78,04	57,92	47,6	81,89
Micro Average	P	91,93	-	92,52	94,22
	R	90,57	-	81,88	92,42
	F1	91,25	89,05	86,88	**93,31**

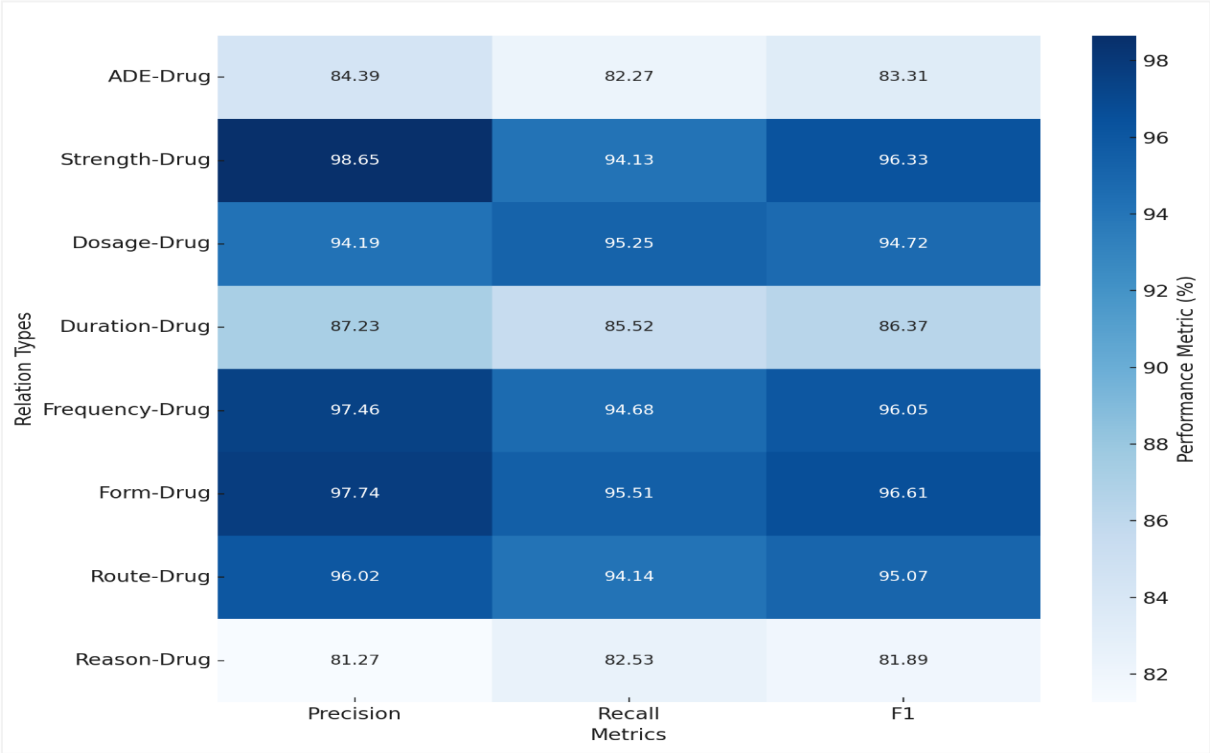

Figura 4. Mapa de calor del rendimiento del enfoque propuesto en varios tipos de relaciones

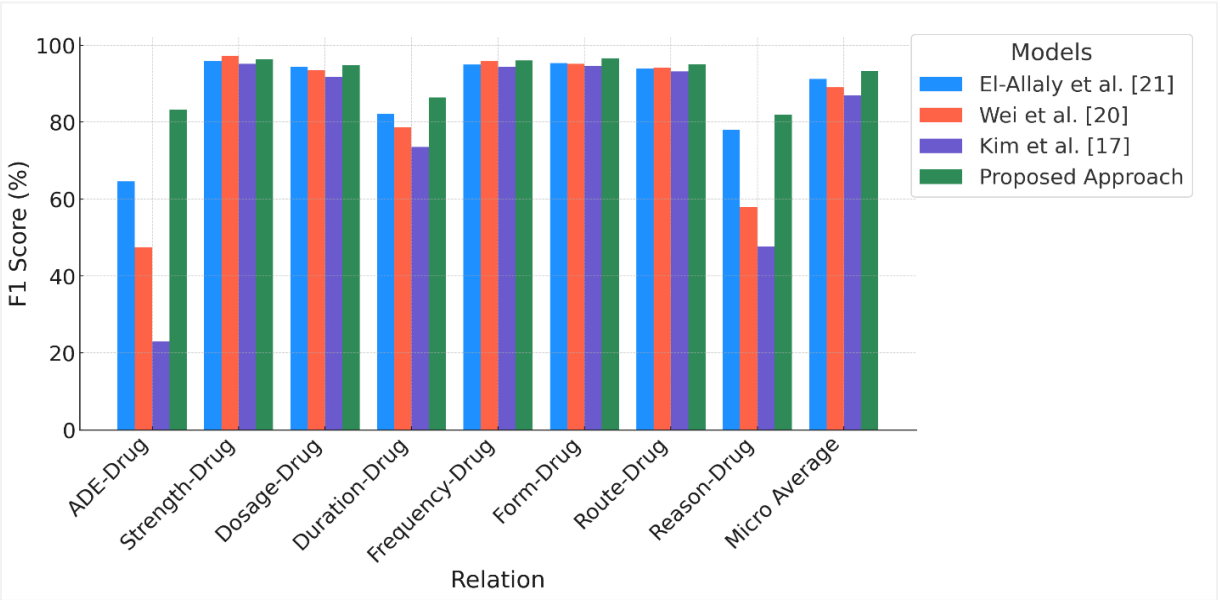

Figura 5. Comparación de rendimiento

Los resultados demuestran que nuestro enfoque aborda eficazmente el reto de la extracción de relaciones ADE en PNL clínica, aprovechando mecanismos de atención dual para capturar interacciones entre entidades e información contextual, a la vez que emplea aprendizaje contrastivo para estructurar el espacio de incrustación. Nuestra hipótesis es que el mecanismo de atención dual permite al modelo atender por separado las menciones de fármacos y ADE, así como el contexto circundante, capturando tanto las características de interacción locales como las claves semánticas globales. Esto podría permitir al modelo resolver casos ambiguos que los modelos anteriores pasaban por alto. Mientras tanto, el objetivo de aprendizaje contrastivo fomenta explícitamente que los pares correctos ADE-Fármaco estén más cerca en el espacio de incrustación y que los pares no relacionados estén más alejados. Esto probablemente sea útil en casos de desequilibrio de clases o señales débiles (comunes en la categoría ADE-Fármaco).

La Figura 4 muestra un mapa de calor de las puntuaciones F1 para cada tipo de relación generado por nuestro modelo. La intensidad (o valor) del color indica el rendimiento: las celdas más brillantes indican un F1 más alto. Este resumen visual muestra que nuestro modelo alcanza un F1 alto en la mayoría de las relaciones y destaca, en particular, la mejora sustancial de la relación ADE-Fármaco con respecto a los valores de referencia. A pesar de esta mejora, el rendimiento de ADE-Fármaco sigue siendo inferior al de relaciones como Fuerza-Fármaco. La Figura 5 presenta un gráfico de barras agrupado que compara las puntuaciones F1 micro promedio, moderadas, de nuestro modelo y valores de referencia previos. El gráfico muestra un claro aumento en el rendimiento de ADE-Fármaco: el listón de nuestro modelo se sitúa muy por encima de todos los sistemas anteriores. Al mismo tiempo, las relaciones de alta frecuencia, como Fuerza-Fármaco, se mantienen cerca del límite en todos los modelos, lo que confirma que la atención dual combinada con el aprendizaje contrastivo potencia la relación más difícil sin reducir el rendimiento en las más fáciles. En comparación con trabajos previos, nuestro método proporciona una solución más refinada y específica para cada tarea. La mejora significativa en la categoría de la relación ADE-Fármaco subraya la solidez de nuestro enfoque para extraer relaciones clínicamente relevantes, impulsando así el trabajo hacia la extracción de relaciones de extremo a extremo en el dominio ADE.

CONCLUSIONES Y TRABAJO FUTURO

Propusimos un marco integral de extracción de relaciones para tareas clínicas de PLN, específicamente enfocado en la extracción de relaciones ADE. La integración del mecanismo de atención dual y el aprendizaje contrastivo mejora significativamente la capacidad de capturar interacciones entre entidades y sus relaciones contextuales más amplias, abordando matices relacionales y semánticos. Nuestra evaluación en el conjunto de datos n2c2 ADE 2018 muestra un excelente rendimiento, alcanzando una media F1 del 93,31%. Cabe destacar que nuestro método mejoró significativamente el rendimiento en la categoría crítica de relación ADE-fármaco, alcanzando una puntuación F1 del 83,31%, lo que representa un aumento sustancial con respecto a los parámetros de referencia anteriores. Estos resultados destacan la capacidad de nuestro modelo para superar las limitaciones existentes, como los desequilibrios de clase y los contextos ambiguos, a la vez que ofrece predicciones robustas y clínicamente significativas. Más allá de las mejoras técnicas, esta extracción de relaciones mejorada tiene importantes implicaciones clínicas. Una identificación más precisa de los pares ADE-fármaco en las HCE puede permitir la detección temprana de efectos secundarios perjudiciales, respaldar las iniciativas de farmacovigilancia y, en última instancia, mejorar la seguridad del paciente. Por ejemplo, al identificar automáticamente asociaciones entre fármacos y ADE poco reconocidas previamente, el modelo podría alertar a los médicos sobre posibles riesgos de la medicación, reduciendo los reingresos hospitalarios e informando sobre una prescripción más segura. Estas aplicaciones ilustran el valor de nuestro

enfoque no solo para los parámetros de referencia del PLN, sino también para los resultados de la atención médica en el mundo real.

A pesar de estos avances, aún existen oportunidades de mejora. Nuestro trabajo futuro se centrará en mejorar la detección de la amplitud de entidades mediante técnicas más avanzadas, como modelos basados en la amplitud o la vinculación de entidades, para abordar ambigüedades en textos clínicos. Además, la incorporación de minería negativa dura en el marco de aprendizaje contrastivo puede refinar aún más el espacio de incrustación y mejorar la discriminación del modelo de relaciones sutiles. Para gestionar las dependencias temporales y causales inherentes a las relaciones de ADE, nuestro objetivo es integrar el razonamiento temporal y los mecanismos de inferencia causal. También planeamos extender nuestro modelo a corpus clínicos más amplios, como MIMIC-III o MIMIC-IV, para evaluar su generalización en diversos conjuntos de datos. También se priorizará la mejora de la explicabilidad del modelo mediante mapas de calor de atención interpretables para aumentar la confianza y la usabilidad para los profesionales de la salud.

FINANCIAMIENTO

Los autores no recibieron patrocinio para realizar este estudio-artículo.

CONFLICTO DE INTERESES

No existe ningún conflicto de intereses relacionado con el tema del trabajo.

CONTRIBUCIÓN DE AUTORES

Conceptualización, curación de datos, análisis formal, investigación, visualización, redacción (borrador original), redacción (corrección y edición): Kashtriya, P. y Singh, P.

REFERENCIAS

Aguilar-Canto, F., Cardoso-Moreno, M., Jiménez, D., & Calvo, H. (2023). GPT-2 versus GPT-3 and Bloom: LLMs for LLMs Generative Text Detection. *CEUR Workshop Proceedings*, *3496*.

Alsentzer, E., Murphy, J. R., Boag, W., Weng, W.-H., Jin, D., Naumann, T., & McDermott, M. B. A. (2019). *Publicly Available Clinical BERT Embeddings*. http://arxiv.org/abs/1904.03323

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners*. http://arxiv.org/abs/2005.14165

Christopoulou, F., Tran, T. T., Sahu, S. K., Miwa, M., & Ananiadou, S. (2020). Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods. *Journal of the American Medical Informatics Association*, *27*(1), 39-46. https://doi.org/10.1093/jamia/ocz101

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. http://arxiv.org/abs/1810.04805

El-allaly, E., Sarrouti, M., En-Nahnahi, N., & Ouatik El Alaoui, S. (2021). MTTLADE: A multi-task transfer learning-based method for adverse drug events extraction. *Information Processing & Management*, *58*(3), 102473. https://doi.org/10.1016/j.ipm.2020.102473

Gao, T., Yao, X., & Chen, D. (2021). *SimCSE: Simple Contrastive Learning of Sentence Embeddings*. http://arxiv.org/abs/2104.08821

Gurulingappa, H., Mateen-Rajpu, A., & Toldo, L. (2012). Extraction of potential adverse drug events from medical case reports. *Journal of Biomedical Semantics*, *3*(1), 15. https://doi.org/10.1186/2041-1480-3-15

Henry, S., Buchan, K., Filannino, M., Stubbs, A., & Uzuner, Ö. (2020). 2018 n2c2 shared task on adverse drug events and medication extraction in electronic health records. *Journal of the American Medical Informatics Association*, 27(1), 3–12. https://doi.org/10.1093/jamia/ocz166

IOM. (2007). Preventing Medication Errors: Quality Chasm Series. En *The National Academies Press, Washington, DC*. Institute of Medicine.

Ji, Z., Wei, Q., & Xu, H. (2019). *BERT-based Ranking for Biomedical Entity Normalization*. http://arxiv.org/abs/1908.03548

Johnson, A. E. W., Bulgarelli, L., & Pollard, T. J. (2020). Deidentification of free-text medical records using pre-trained bidirectional transformers. *Proceedings of the ACM Conference on Health, Inference, and Learning*, 214-221. https://doi.org/10.1145/3368555.3384455

Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L. H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Anthony Celi, L., & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data*, *3*(1), 160035. https://doi.org/10.1038/sdata.2016.35

Kim, Y. H., Kim, C., & Kim, Y. S. (2024). Language Model-Based Text Augmentation System for Cerebrovascular Disease-Related Medical Reports. *Applied Sciences*, *14*(19).

Kim, Y., & Meystre, S. M. (2020). Ensemble method–based extraction of medication and related information from clinical texts. *Journal of the American Medical Informatics Association*, *27*(1), 31-38. https://doi.org/10.1093/jamia/ocz100

Kumar, A., Chaudhary, K., Singh, P., & Thapa, J. (2020). Contextual Data Augmentation for Improving Natural Language Understanding in Biomedical Domain. *IEEE BIBM*.

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2019). *BioBERT: a pre-trained biomedical language representation model for biomedical text mining*. https://doi.org/10.1093/bioinformatics/btz682

Li, Y., Tao, W., Li, Z., Sun, Z., Li, F., Fenton, S., Xu, H., & Tao, C. (2024). Artificial intelligence-powered pharmacovigilance: A review of machine and deep learning in clinical text-based adverse drug event detection for benchmark datasets. *Journal of Biomedical Informatics*, *152*, 104621. https://doi.org/10.1016/j.jbi.2024.104621

Liu, S., Tang, B., Chen, Q., & Wang, X. (2016). Drug-Drug Interaction Extraction via Convolutional Neural Networks. *Computational and Mathematical Methods in Medicine*, *2016*, 1-8. https://doi.org/10.1155/2016/6918381

Modi, S., Kasmiran, K. A., Mohd Sharef, N., & Sharum, M. Y. (2024). Extracting adverse drug events from clinical Notes: A systematic review of approaches used. *Journal of Biomedical Informatics*, *151*, 104603. https://doi.org/10.1016/j.jbi.2024.104603

Murphy, S. N., Weber, G., Mendis, M., Gainer, V., Chueh, H. C., Churchill, S., & Kohane, I. (2010). Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *Journal of the American Medical Informatics Association*, *17*(2), 124-130. https://doi.org/10.1136/jamia.2009.000893

Peng, Y., Yan, S., & Lu, Z. (2018). Transfer Learning in Biomedical Natural Language Processing: An Evaluation of BERT and ELMo on Ten Benchmarking Datasets. *Proceedings of the 18th BioNLP Workshop and Shared Task*, 58-65. https://doi.org/10.18653/v1/W19-5006

Peng, Y., Yan, S., & Lu, Z. (2020). Biomedical Sentence Similarity Estimation with Transfer Learning and Entity-aware Contrastive Learning. *BioNLP Workshop - ACL*.

Roberts, K., Demner-Fushman, D., & Tonellato, P. J. (2009). Identifying and classifying adverse drug events in discharge summaries. *AMIA Annual Symposium Proceedings*, 601–605.

Sun, M., Zhang, Z., Han, X., Liu, Z., Jiang, X., & Liu, Q. (2019). *ERNIE: Enhanced Language Representation with Informative Entities*. http://arxiv.org/abs/1905.07129

Thapa, S., Islamaj, R., Rogers, K., & Demner-Fushman, D. (2022). Transfer Learning and Domain Adaptation with Pre-trained Embeddings for Clinical Natural Language Processing: A Case Study on Adverse Drug Event Classification. *Journal of the American Medical Informatics Association*, *29*(6), 1019–1027. https://academic.oup.com/jamia/search-results?page=1&q=TransferLearningandDomainAdaptationwithPre-trainedEmbeddingsforClinicalNaturalLanguageProcessing%3AACaseStudyonAdverseDrugEventClassification&fl_SiteID=5396&SearchSourceType=1&allJ

Vig, J. (2019a). *A Multiscale Visualization of Attention in the Transformer Model*. http://arxiv.org/abs/1906.05714

Vig, J. (2019b). *Visualizing Attention in Transformer-Based Language Representation Models*. http://arxiv.org/abs/1904.02679

Wei, Q., Ji, Z., Li, Z., Du, J., Wang, J., Xu, J., Xiang, Y., Tiryaki, F., Wu, S., Zhang, Y., Tao, C., & Xu, H. (2020). A study of deep learning approaches for medication and adverse drug event extraction from clinical text. *Journal of the American Medical Informatics Association*, *27*(1), 13-21. https://doi.org/10.1093/jamia/ocz063

WHO. (2008). *The importance of pharmacovigilance - Safety Monitoring of medicinal products*. World Health Organization.

Yang, X., Bian, J., Fang, R., Bjarnadottir, R. I., Hogan, W. R., & Wu, Y. (2020). Identifying relations of medications with adverse drug events using recurrent convolutional neural networks and gradient boosting. *Journal of the American Medical Informatics Association*, *27*(1), 65-72. https://doi.org/10.1093/jamia/ocz144

Zhang, F., Zheng, W., Yu, J., & Qin, Y. (2020). *Dual Attention-Based BiLSTM-CNN for Named Entity Recognition*.

Zhang, Y., Chen, Q., Yang, Z., Lin, H., & Lu, Z. (2019). BioWordVec, improving biomedical word

embeddings with subword information and MeSH. *Scientific Data*, *6*(1), 52. https://doi.org/10.1038/s41597-019-0055-0

Zhang, Y., Sun, A., Jin, F., & Zhou, X. (2020). Contrastive learning for text classification. *AAAI*. http://arxiv.org/abs/2305.09269

Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., & Xu, B. (2016). Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 207-212. https://doi.org/10.18653/v1/P16-2034