

**Revisión de herramientas de Inteligencia Artificial para extracción de texto en enunciados
de razonamiento lógico**

Jorge Andrés Ordoñez Moncada

Asesor

Julio Eduardo Mejía Manzano

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería (ECBTI)

Programa

2025

Resumen

El crecimiento exponencial de los datos textuales, proyectado para ser 175 veces mayor en 2025 que en 2011, plantea desafíos y oportunidades clave en la extracción de información mediante herramientas de inteligencia artificial. Este trabajo se centra en una revisión bibliográfica de herramientas para el procesamiento y extracción de lenguaje natural en enunciados de razonamiento lógico. La investigación se llevará a cabo en la biblioteca virtual de la Universidad Nacional a Distancia, con el fin de consolidar la información en tablas para clasificar y resaltar las fortalezas y debilidades de los hallazgos más relevantes. Los resultados del trabajo evidencian que, aunque los modelos basados en Deep Learning han avanzado significativamente, aún presentan dificultades para capturar detalles críticos. En este contexto, los enfoques híbridos, que combinan redes neuronales con razonamiento lógico estructurado, emergen como una alternativa prometedora. Estas herramientas han demostrado un alto potencial para superar las limitaciones de los métodos tradicionales, permitiendo abordar problemas de inferencia lógica con mayor precisión y adaptabilidad.

Palabras clave: Deep learning, procesamiento de lenguaje natural, razonamiento lógico

Abstract

The exponential growth of textual data, projected to be 175 times larger in 2025 than in 2011, poses key challenges and opportunities in information extraction using artificial intelligence tools. This paper focuses on a literature review of tools for natural language processing and extraction in logical reasoning statements. The research, carried out in the virtual library of the Universidad Nacional a Distancia, allowed to consolidate the information in tables to classify and highlight the strengths and weaknesses of the most relevant findings. The results show that, although Deep Learning-based models have advanced significantly, they still present difficulties in capturing critical details. In this context, hybrid approaches, which combine neural networks with structured logical reasoning, emerge as a promising alternative. These tools have demonstrated a high potential to overcome the limitations of traditional methods, allowing to address logical inference problems with greater accuracy and adaptability.

Key words: Deep learning, natural language processing, logical reasoning.

Tabla de contenido

Introducción	7
Planteamiento del problema.....	8
Justificación	10
Objetivos	12
Objetivo General	12
Objetivos Específicos.....	12
Marco Conceptual Y Teórico.....	13
Inteligencia Artificial	13
Lógica.....	15
Procesamiento de Lenguaje Natural (PNL)	16
Metodología	17
Método	19
Resultados	20
Herramientas de procesamiento y extracción de lenguaje natural	21
Clasificación de la información	24
Identificación de debilidades y fortalezas de cada una de las herramientas existentes.	25
Conclusiones	29
Recomendaciones	30
Apéndice	32

Lista de Tablas

Tabla 1: Actividades y Justificación por Objetivo.....	18
Tabla 2: Descripción de las herramientas encontradas sobre el procesamiento de lenguaje natural	21
Tabla 3: Clasificación de las herramientas encontradas sobre el procesamiento de lenguaje natural	24
Tabla 4: Fortalezas y debilidades de las herramientas encontradas sobre el procesamiento de lenguaje natural.....	25

Lista de Apéndices

Apéndice A: Formato de extracción de datos usado en la presente investigación.	29
--	-----------

Introducción

En la actualidad, el manejo de grandes volúmenes de datos es una realidad en sectores como la educación, la tecnología y la investigación. Este crecimiento exponencial de datos plantea desafíos significativos en la forma de procesar e interpretar información de manera automática y eficiente. Uno de los campos que requiere especial atención es la extracción de lenguaje natural aplicado a problemas que involucran razonamiento lógico, debido a la difícil comprensión del lenguaje natural.

El razonamiento lógico, formalizado a través de herramientas como el álgebra booleana, ha sido fundamental en las matemáticas y la lógica desde el siglo XIX. No obstante, trasladar estas representaciones simbólicas al ámbito de la inteligencia artificial ha evidenciado limitaciones. Modelos actuales de procesamiento de lenguaje natural (PLN) enfrentan dificultades al interpretar enunciados complejos, especialmente cuando estos contienen elementos implícitos o estructuras lingüísticas ambiguas. Estas limitaciones pueden generar interpretaciones erróneas y resultados inexactos, lo que reafirma la necesidad de desarrollar sistemas más precisos y funcionales.

A pesar de los avances recientes en inteligencia artificial, como los modelos de aprendizaje profundo, los desafíos relacionados con la ambigüedad del lenguaje, el costo computacional y la aplicabilidad de los resultados siguen siendo una barrera para su implementación efectiva.

Planteamiento del problema

Durante los desarrollos de la matemática pura en el siglo XIX se determinó la importancia de liberar a los razonamientos deductivos de todo tipo de intuición para darle validez formal, en ese sentido fue necesario desarrollar toda una serie de símbolos que representaran las preposiciones y otros que representaran sus relaciones. Este lenguaje propio convirtió los problemas de lógica en algo muy parecido a los problemas de álgebra, por lo que se denominó a esta forma de lógica algebra Booleana (Skansi 2020).

Este importante paso en las matemáticas puras, esta representación por medio de los símbolos lógicos de los enunciados de los problemas ha generado grandes inconvenientes en tanto en los estudiantes universitarios como en los lectores de lenguaje inteligente (NPL) e incluso fueron de las primeras evidencias de errores en las primeras versiones de ChatGTP. El planteamiento automático de los problemas matemáticos son un problema de investigación que combina los dominios de tecnología de procesamiento, el lenguaje natural, el aprendizaje automático y la educación (Mandal & Naskar, 2021).

La problemática es la debilidad de las técnicas de extracción de texto en la interpretación de enunciados lógicos que se deben a que no puede resolverse mediante una identificación de patrones, sino que requiere de una profunda comprensión semántica y capacidad de inferencia (Wang, Zhang, Gao, Song, Guo & Shen, 2018).

En particular la deficiente interpretación de enunciados y la no identificación del sujeto cuando no se escribe explícitamente, cuando se confunden los conectores lógicos con conjunciones presentes en el enunciado. Todo esto genera que los planteamientos sean erróneos, que información presente en el texto se pierda o se mal interprete generando soluciones erróneas y que se requiera una revisión manual que se transforma en un incremento de costos.

Se ha desarrollado toda una rama de la Inteligencia Artificial como lo es el procesamiento de lenguaje natural (PLN) y se ha aplicado a problemas matemáticos alcanzando niveles de acierto superiores al 80% (Ontanon, Ainslie, Cvicek & Fisher,2022). ¿Cuáles son las herramientas y avances existentes de inteligencia artificial sobre la extracción de texto en problemas relacionados con la inferencia lógica?

Justificación

Según la Fundación de Innovación Barkinter (2018), el volumen de datos existentes en 2025 será 175 veces más que en 2011. Con este crecimiento exponencial de datos textuales disponibles, desde redes sociales hasta documentos académicos, blogs y noticias, hay una necesidad creciente de extraer información valiosa de estos datos de manera eficiente y precisa. Una adecuada extracción de texto permite automatizar tareas que anteriormente requerían una gran cantidad de trabajo manual, como el análisis de sentimientos, la clasificación de documentos, la extracción de información de formularios, la resolución de problemas matemáticos, entre otros.

Estos sistemas de extracción han sido entrenados para ser cada vez más precisos en su ejecución. Sin embargo, aún no están en la capacidad de captar detalles esenciales, como lo afirman los autores Dasgupta y Gershman (2020) las redes modernas se vuelven cada vez más complejas, y se acercan a las capacidades humanas solo en algunos dominios, la pregunta radica en las representaciones y decisiones que toman comparadas con las de los humanos. Aun cuando el sistema de extracción logre captar el texto adecuadamente, no significa que el texto extraído tenga sentido, según Wang y Pan (2021) la mayoría de los modelos de Deep Learning funcionan como cajas negras, carecen de capacidades de raciocinio, que normalmente son esenciales para el desarrollo de problemas complejos.

El lenguaje natural es inherentemente ambiguo, lo que puede dificultar aún más la extracción precisa de información. En el campo de la lógica matemática, a menudo se enfrentan problemas complejos que requieren una comprensión profunda del lenguaje natural para ser resueltos (Xuguang, Xin, Peilin, Guangyao & Wenwu, 2023). Detectar las palabras o frases que se refieren a entidades previamente mencionadas en el texto puede resultar complicado. Las expresiones lógicas desempeñan un papel importante en muchas aplicaciones informáticas, como

el razonamiento automatizado y la programación lógica, y la verificación forma de software y hardware. (Xin, Yi, Zetao, Ho-fung, Tao, 2023).

Además de ser un campo por explorar más a profundidad, según Lauriola, Lavelli y Aioli (2022) hay otros problemas que necesitan ser contemplados, como el costo computacional y la reproducibilidad de los resultados y la falta de interpretación de estos. Con la revisión bibliográfica, se pretende recomendar diferentes herramientas de inteligencia artificial para la extracción de lenguaje natural en enunciados de razonamiento lógico.

Objetivos

Objetivo General

Identificar las herramientas de inteligencia artificial existentes para el procesamiento y extracción de lenguaje natural en enunciados de inferencia lógica.

Objetivos Específicos

Buscar en la literatura herramientas existentes sobre el procesamiento y extracción de lenguaje natural en enunciados de inferencia lógica.

Clasificar los hallazgos, estudios y trabajos relacionados con la aplicación del procesamiento y extracción de lenguaje natural.

Identificar debilidades y fortalezas de cada una de las herramientas existentes.

Resumir los hallazgos de forma descriptiva para presentar los resultados y que sirvan como referente para investigaciones futuras dando recomendaciones sobre las herramientas más usadas en la investigación.

Marco Conceptual Y Teórico

Inteligencia Artificial

La Inteligencia Artificial (IA) se encuentra en el centro de numerosos debates actuales pues existe un escaso entendimiento común sobre las distinciones y semejanzas entre la inteligencia humana y la artificial. En estos debates, temas cruciales como la confiabilidad y la ética suelen estar impregnados de argumentos en los que la IA debiera ser exactamente como un ser humano. La búsqueda de una inteligencia artificial que se asemeje a la humana se percibe a menudo como el estándar máximo de la IA.

Butazzo (2023) indica que la inteligencia artificial está logrando avances extraordinarios a un ritmo sin precedentes, alcanzando y superando las capacidades humanas en muchas tareas que antes se consideraban inalcanzables para las máquinas, como la traducción de idiomas, la composición musical, la detección de objetos, los diagnósticos médicos, la programación de software y muchas otras. Así, dependiendo la perspectiva desde que se observe, la inteligencia artificial puede ser vista como una ayuda a la resolución de problemas que toman bastante tiempo para el ser humano descifrar y muchas veces hasta formular.

El aprendizaje juega un papel fundamental en cómo los seres humanos piensan y comprenden el mundo que les rodea. Utilizamos nuestro conocimiento abstracto para procesar una gran cantidad de información y mejorar nuestra comprensión de lo que percibimos. Como parte de la Inteligencia Artificial, los modelos de Machine Learning tienen la capacidad de imitar estas capacidades cognitivas humanas debido a su naturaleza adaptable. El Machine Learning se refiere a un conjunto de métodos utilizados para resolver una variedad de problemas del mundo real utilizando sistemas informáticos, los cuales pueden aprender a resolver un problema en lugar de ser programados explícitamente para hacerlo (Koza, 1996).

Con la aplicación de sistemas de IA con una autonomía cada vez mayor, se considera cada vez más la necesidad de abordar las cuestiones reales y complejas de la inteligencia a nivel humano, más ampliamente, de la inteligencia artificial general. Una de estas problemáticas de la realidad es la solución de problemas de inferencia lógica, en la que la extracción e interpretación del texto juega un papel fundamental (Liu, Zhang, Lu, Ma, Ho, 2024).

Un objetivo a largo plazo de la inteligencia artificial es lograr un razonamiento preciso. Este tipo de razonamiento es crucial para tareas de procesamiento del lenguaje natural (PLN), como la traducción automática, la respuesta a preguntas específicas y los juegos de texto. Combinar el aprendizaje de redes neuronales y el razonamiento simbólico es crucial para una IA sólida. Pocos estudios pioneros han logrado avances en tareas de razonamiento lógico que requieren entradas particionadas de instancias (por ejemplo, datos secuenciales). (Han, Cai y Dai, 2023).

Los modelos actuales pueden aprender muchos aspectos del lenguaje a partir de grandes conjuntos de datos, a menudo fallan en manejar casos más sutiles como el conteo, la copia, el razonamiento posicional y la comprensión de antónimos. Los problemas relacionados con el procesamiento del lenguaje natural, como el reconocimiento de voz, la minería de datos basada en texto y la generación de texto o voz, están adquiriendo una importancia creciente. Para abordar eficazmente muchos de estos desafíos, es esencial primero analizar y procesar las estructuras sintácticas de las oraciones (Jaf, Calder , 2019). Resolver estos problemas requiere un conocimiento externo de la lógica, especialmente en dominios específicos y tareas detalladas. Incluso con grandes volúmenes de datos, es difícil que los modelos aprendan perfectamente la lógica humana, ya que tienden a adoptar sesgos del conjunto de datos en lugar de comprender el proceso de razonamiento subyacente (Yao, Xu, Shi, & Xu 2018).

Lógica

La lógica se dedica a analizar la estructura del razonamiento. Esta disciplina utiliza reglas y técnicas para determinar la validez de un argumento. Tiene un amplio espectro de aplicaciones, incluyendo la filosofía, las matemáticas, la computación y la física. En filosofía, la lógica se emplea para evaluar la validez de los razonamientos y discernir el significado correcto de frases con posibles interpretaciones variadas. En matemáticas, sirve para demostrar teoremas y deducir resultados útiles en investigaciones. En computación, se aplica para la revisión y verificación de programas (Jimenez & Hernandez, 2018).

Debido al continuo avance y mejora de la sociedad humana, así como al crecimiento en cantidad y calidad del conocimiento resultante, han surgido y continúan surgiendo nuevas disciplinas relacionadas con los procesos del pensamiento. A continuación, se abordarán ciertos conceptos y componentes de la lógica matemática.

Las proposiciones son estructuras lingüísticas con sentido, y de las cuales se puede decir que son verdaderas o falsas. Sobre el contenido de verdad de una proposición, a través de los axiomas o principios de la lógica, se han hecho las siguientes consideraciones: Toda proposición o es falsa o es verdadera, según el principio de la bivalencia. De este principio se pueden deducir los dos teoremas siguientes. El teorema de la tercera posibilidad excluida, que expresa: Toda proposición es falsa o verdadera. El teorema de la contradicción excluida, que expresa: ninguna proposición es falsa y verdadera al mismo tiempo.

A partir de dos o más proposiciones podemos formar los enlaces de proposiciones utilizando los conectores lógicos (y, o, si ... entonces, si y solo si, entre otros). En el lenguaje

común se utiliza un procedimiento similar; a partir de las oraciones más sencillas formamos oraciones compuestas utilizando conjunciones. (Carmenates, Amat, Gamboa, González, 2005, p.4)

La expresión lógica, un tipo de lenguaje formal, desempeña un papel esencial en diversas aplicaciones informáticas, como el razonamiento automatizado, la programación lógica y la verificación de software y hardware. En la argumentación computacional, los argumentos estructurados se han estudiado y formalizado durante décadas utilizando modelos que pueden expresarse en un marco lógico (Bench-Capon & Dunne, 2007).

Sin embargo, aunque las técnicas recientes están empezando a alcanzar un rendimiento excelente en diversas tareas, todavía quedan varios problemas por resolver como lo indican Lauriola, Lavelli y Aiolli (2022), el coste computacional y la reproducibilidad de los resultados carecen de interpretabilidad y precisión lógica.

Procesamiento de Lenguaje Natural (PNL)

El procesamiento del lenguaje natural (PNL) es una rama de la inteligencia artificial llena de tareas intrincadas, sofisticadas y desafiantes relacionadas con el lenguaje, como la traducción automática, la respuesta a preguntas, el resumen, etc. El PNL implica el diseño e implementación de modelos, sistemas y algoritmos para resolver problemas prácticos en la comprensión de los lenguajes humanos.

El PNL se puede categorizar en dos áreas principales: investigación fundamental y aplicada. La primera se enfoca en problemas generales que son fundamentales para desarrollar sistemas complejos basados en el lenguaje humano. Estos incluyen el modelado del lenguaje, el análisis morfológico, el análisis sintáctico y el análisis semántico. Por otro lado, el PNL aplicada

aborda cuestiones prácticas como la extracción automática de información relevante de textos, la traducción entre idiomas, la síntesis de documentos, la respuesta automática a preguntas, y la clasificación y agrupación de documentos. (Lauriola, Lavelli, Aioli, 2022)

Como lo afirman Salvatore, Finger, Hirata y Patriot (2023), el Deep learning es una de las áreas, para PNL, de la ciencia de datos de más rápido crecimiento. Se refiere a una clase de algoritmos que se basan en redes neuronales artificiales optimizadas para trabajar con datos no estructurados como imágenes, voz, videos y texto. Si bien las técnicas de Deep Learning se remontan a mediados de los años 80, su verdadero potencial sólo se ha materializado en los últimos años (Kotu & Deshpande,2019). Sin embargo, cómo lo afirman los autores Galassi, Kersting, Lippi, Shao y Torroni (2020) el Deep Learning está aportando notables contribuciones al campo de la minería de la argumentación, pero los enfoques existentes todavía necesitan llenar el vacío que existe hacia la realización de tareas de razonamiento avanzadas.

Metodología

Tabla 1: Actividades y Justificación por Objetivo

Objetivo 1: Buscar en la literatura herramientas existentes sobre el procesamiento y extracción de lenguaje natural en enunciados de inferencia lógica	
Actividad	Justificación
Revisión de literatura y herramientas actuales mediante análisis documental y revisión bibliográfica.	Encontrar información pertinente, que permita hacer una comparación de avances existentes sobre el procesamiento de lenguaje natural en enunciados de inferencia lógica.
Objetivo 2: Clasificar los hallazgos, estudios y trabajos relacionados con la aplicación del procesamiento y extracción de lenguaje natural.	
Actividad	Justificación
Diseñar una tabla que permita organizar los hallazgos y situarlos dentro del mundo de la inteligencia artificial.	Se realiza la creación de una tabla como una herramienta visual y práctica para organizar los hallazgos debido a la amplitud y complejidad de la inteligencia artificial, es necesario establecer una forma de sintetizar y estructurar la información de manera efectiva.
Objetivo 3: Identificar debilidades y fortalezas de cada una de las herramientas existentes.	
Actividad	Justificación
Proporcionar una visión integral que facilite la selección de la herramienta más adecuada, contribuyendo al desarrollo de recomendaciones informadas para futuros investigadores interesados en esta área.	La identificación de las fortalezas y debilidades de cada herramienta resulta fundamental, ya que permite realizar una comparación detallada que facilite la selección de la opción más prometedora. Este análisis comparativo no solo contribuye a una mejor

	comprensión de las capacidades y limitaciones de las herramientas disponibles, sino que también orienta hacia la recomendación de la solución más adecuada para abordar los desafíos específicos planteados en el procesamiento y extracción de lenguaje natural.
Objetivo 4: Resumir los hallazgos de forma descriptiva para presentar los resultados y que sirvan como referente para investigaciones futuras dando recomendaciones sobre las herramientas más usadas en la investigación.	
Actividad	Justificación
Descripción detallada de la herramienta de inteligencia artificial que se recomendará para el procesamiento de lenguaje natural en problemas de razonamiento lógico basado en el análisis de fortalezas y debilidades.	Sugerir una herramienta para quienes deseen continuar explorando herramientas de inteligencia artificial aplicadas al procesamiento de lenguaje natural en enunciados de razonamiento lógico.

Método

Se emplean dos métodos principales:

Método bibliográfico/documental:

Revisión exhaustiva de hallazgos, estudios y trabajos en los que se aplica inteligencia artificial para el procesamiento y extracción de lenguaje natural en enunciados de inferencia lógica.

Método analítico-descriptivo:

Análisis descriptivo de herramientas de inteligencia artificial existentes para el procesamiento y extracción de lenguaje natural en enunciados de inferencia lógica.

Resultados

Herramientas de procesamiento y extracción de lenguaje natural

Los avances recientes en el campo de la inteligencia artificial han llevado a mejoras sin precedentes en las aplicaciones de Procesamiento del Lenguaje Natural (PNL), generando un creciente interés entre la comunidad de aprendizaje automático. En la revisión bibliográfica, se encontraron las siguientes herramientas de inteligencia artificial orientadas a la extracción de texto en problemas de razonamiento lógico:

Tabla 2: Descripción de las herramientas encontradas sobre el procesamiento de lenguaje natural

Herramienta	Definición/explicación	Referencia
Deep Reinforcement Learning (DRL)	<p>Modelo que utiliza reglas lógicas explícitamente durante el razonamiento. Esta herramienta fue aplicada en tres fases:</p> <p>Estado: Una lista de tuplas relacionales en la memoria de trabajo (como entorno), y las frases de la historia (como entrada).</p> <p>Transición de estado: La inserción, eliminación o modificación de la lista de tuplas, basada en acciones y reglas lógicas.</p> <p>Acción: Generación de nuevas tuplas (o ninguna) utilizadas para modificar o recuperar de la lista de tuplas. Las acciones típicas pueden descomponerse.</p>	<p>https://www-sciencedirect-com.bibliotecavirtual.unad.edu.co/science/article/abs/pii/S0893608018301941?via%3Dihub</p>

<p>Neural-Symbolic Systems (NeSy) - Statistical Relational Learning (SRL) - DeepProbLog</p>	<p>Este modelo es propuesto para explotar el potencial de los enfoques simbólico y sub-simbólico para el procesamiento de lenguaje natural, combinando ambos resultados en sistemas capaces de modelar el conocimiento y las restricciones con un formalismo lógico, al tiempo que se mantiene la potencia computacional de las redes profundas (Neural Networks).</p>	<p>https://www.frontiersin.org/journals/big-data/articles/10.3389/fdata.2019.00052/full</p>
<p>Modelos de Transformadores (BERT, RoBERTa, ALBERT, XLNet), LSTMs (Long Short-Term Memory)</p>	<p>Uso de modelos modernos de Deep Learning (DL), como el transformador preentrenado BERT, así como modelos adicionales que se basan en redes LSTM, para la inferencia de lenguaje natural. Se comparan cinco modelos DL (y variaciones de los mismos) en ocho conjuntos de datos de inferencia de lenguaje natural ampliamente utilizados. Se entrena y ajusta los hiperparámetros de cada modelo para obtener el mejor rendimiento en cada conjunto de datos.</p>	<p>https://www.mdpi.com/2076-3417/13/4/2577</p>
<p>Deep Neural Networks (DNN) & Logic Networks</p>	<p>Deep Neural Networks (DNNs) es una herramienta que transforma secuencias de texto de entrada en características abstractivas y produce salidas</p>	<p>https://direct.mit.edu/coli/article/47/4/775/106773/Variational-Deep-Logic-</p>

	<p>probabilísticas.</p> <p>Logic Networks herramienta utilizada que parametriza el proceso de inferencia lógica mediante una jerarquía de capas formada por una capa de átomos y una capa de reglas.</p> <p>Se presenta un modelo que hereda la capacidad de representación del Deep Learning y, al mismo tiempo, simula el proceso de aprendizaje de Logic Rules mediante una novedosa red lógica compuesta por una jerarquía de una capa de átomos, una capa de reglas y una capa de salida.</p>	<p>Network-for-Joint-Inference</p>
<p>LOGICINFERENCE, SNLI , CLUTRR</p>	<p>LOGICINFERENCE es un conjunto de datos diseñado que se centra en la inferencia mediante lógica proposicional y un pequeño subconjunto de lógica de primer orden, representados tanto en notación lógica semiformal como en lenguaje natural.</p>	<p>https://arxiv.org/pdf/2203.15099</p>

Clasificación de la información

Cada una de las herramientas mencionadas pertenece al campo de la Inteligencia Artificial (IA). A continuación, se muestra una tabla que organiza estas herramientas según su naturaleza y campo dentro de la IA:

Tabla 3: Clasificación de las herramientas encontradas sobre el procesamiento de lenguaje natural

Macro	Clasificación	Sub-Categoría	Técnica o Herramienta
Artificial Intelligence (AI)	Machine Learning	Deep Learning	Deep Reinforcement Learning (DRL)
Natural Language Processing (NLP)	Neural network Techniques	Text processing	Neural-Symbolic Systems (NeSy) - Statistical Relational Learning (SRL) - DeepProbLog
Artificial Intelligence (AI)	Neural Networks / Machine Learning	Modern Deep Learning (DL) models	Modelos de Transformadores (BERT, RoBERTa, ALBERT, XLNet), LSTMs (Long Short-Term Memory)
Artificial Intelligence (AI)	Neural Networks / Machine Learning - Logic	Variational Deep Logic Network	Deep Neural Networks (DNN) & Logic Networks

Artificial Intelligence (AI)	Model Evaluation	Logical Inference	LOGICINFERENCE, SNLI, CLUTRR

Identificación de debilidades y fortalezas de cada una de las herramientas existentes.

Para cada una de las herramientas se identificaron fortalezas y debilidades en cada uno de los proyectos en los que se utilizaron dichas herramientas. A continuación, se presenta la información:

Tabla 4: Fortalezas y debilidades de las herramientas encontradas sobre el procesamiento de lenguaje natural

Técnica o Herramienta	Fortalezas	Debilidades
Deep Reinforcement Learning (DRL)	Esta herramienta presenta eficiencia en el uso de datos ya que es capaz de entrenar con cerca de 1.000 ejemplos, lo que lo hace altamente eficiente en comparación con otros modelos que requieren conjuntos de datos mucho más grandes.	Presenta dependencia de la activación de reglas lógicas. Si las reglas no se activan correctamente o si la vinculación de variables no es precisa, el modelo puede fallar en tareas de razonamiento más complejas.
	Humanización del razonamiento. Utiliza un enfoque más cercano al razonamiento humano al activar reglas lógicas a través de un proceso de decisión secuencial.	Limitaciones en la interpretación de relaciones abstractas. Aunque las relaciones abstractas son útiles, pueden no capturar completamente todos los niveles del lenguaje y las

		interacciones en contextos complejos.
	Desempeño en benchmarks. Logra un bajo porcentaje de error en tareas complejas como las de bAbI-20, destacándose frente a otros modelos que dependen de grandes volúmenes de datos.	
	Integración de memoria externa. La memoria externa ayuda a gestionar las relaciones complejas y mejorar la precisión en la toma de decisiones.	
Neural-Symbolic Systems (NeSy) - Statistical Relational Learning (SRL) - DeepProbLog	La integración de enfoques simbólicos y sub-simbólicos puede proporcionar ventajas en el desarrollo, mientras se aprovecha el poder computacional y la capacidad de Deep Learning.	El costo computacional puede ser elevado.
	Permite manejar la incertidumbre, y las restricciones lógicas pueden hacer que los sistemas sean más precisos y explicativos.	
Modelos de		Falta de generalización efectiva entre

Transformadores (BERT, RoBERTa, ALBERT, XLNet), LSTMs (Long Short-Term Memory)	Evaluación exhaustiva de modelos en múltiples benchmarks ampliamente utilizados.	conjuntos de datos disímiles, un desafío clave en NLI.
	Uso de modelos modernos (transformadores) que muestran avances significativos en inferencia.	Dependencia de la calidad y método de recolección de datos, lo que limita su desempeño.
	Evaluación de la capacidad de generalización entre datasets, identificando limitaciones cruciales.	Resultados dependientes de la similitud entre datasets (por ejemplo, MNLI y SNLI).
	Inclusión de nuevos enfoques como WANLI y ChaosNLI para abordar sesgos y errores en anotaciones.	Críticas a los métodos de anotación que no reflejan la diversidad de interpretaciones humanas.
Deep Neural Networks (DNN) & Logic Networks	Enfoque novedoso en la integración de redes neuronales profundas y razonamiento lógico. La combinación de la capacidad de aprendizaje automático con la lógica formal mejora la capacidad del modelo para capturar tanto las características implícitas como las explícitas en las relaciones entre entidades.	El uso de un enfoque híbrido que combina redes neuronales profundas y lógica puede ser costoso en términos de recursos computacionales, especialmente cuando se aplican en tareas grandes o de datos complejos.

	<p>Flexibilidad en la incorporación de reglas predefinidas y aprendizaje de reglas. La capacidad de integrar tanto reglas fijas como reglas aprendidas mejora la capacidad del modelo para adaptarse a tareas más complejas.</p>	<p>Dependencia de datos de entrenamiento. Aunque el modelo puede aprender nuevas reglas, sigue siendo muy dependiente de la calidad y cantidad de los datos de entrenamiento, lo que puede afectar su rendimiento en dominios con datos limitados o no etiquetados.</p>
	<p>Mejor rendimiento en tareas de extracción de información: Los resultados experimentales muestran que el modelo supera a otros enfoques en una variedad de tareas de extracción de información con un rendimiento mejorado sin necesidad de conocimiento previo.</p>	<p>Desafíos en la optimización de los dos módulos: La alternancia entre los pasos E y M del aprendizaje variacional puede presentar dificultades en la optimización simultánea de los dos módulos (aprendizaje profundo y razonamiento lógico).</p>
<p>LOGICINFERENCE, SNLI, CLUTRR</p>	<p>Cobertura amplia de razonamiento lógico: LOGICINFERENCE incluye tareas variadas que evalúan desde traducción a lógica hasta inferencias multi-paso en lenguaje formal y natural.</p>	<p>Limitaciones en la variedad lingüística. El uso de lenguaje generado automáticamente puede no capturar matices o complejidades reales del lenguaje humano, limitando la generalización.</p>
	<p>Evaluación robusta de casos críticos (contradicciones, preguntas obvias) aumentan la capacidad del dataset para evaluar la precisión y robustez de los modelos.</p>	<p>Foco limitado en lógica: Solo cubre una parte de la lógica de primer orden y carece de razonamiento numérico, restringiendo su aplicabilidad a otras áreas.</p>

	<p>Generación de cadenas de inferencia. Obliga a los modelos a producir pasos intermedios mejora la interpretabilidad y ayuda a reducir razonamientos "alucinados".</p>	<p>Dependencia de modelos complejos. Aunque se evalúan diferentes modelos, el diseño del dataset puede requerir arquitecturas complejas que no son accesibles para todos los entornos.</p>
--	---	--

Tras identificar las fortalezas y debilidades de cada una de las herramientas, se concluye que el modelo o herramienta que mejores resultados arroja es Neural-Symbolic Systems (NeSy) - Statistical Relational Learning (SRL) – DeepProbLog, ya que es un modelo híbrido que aborda a profundidad tanto el PNL como la inferencia lógica. Además, dentro de sus debilidades, solo se identifica su costo.

Conclusiones

Se evidenció que la combinación de redes neuronales y enfoques simbólicos representa una contribución significativa al desarrollo de soluciones de inteligencia artificial para el procesamiento y extracción de lenguaje natural en enunciados de inferencia lógica. Este enfoque híbrido ofrece capacidades más robustas y adaptativas frente a los desafíos inherentes a este tipo de tareas.

Asimismo, los avances en modelos de Deep Learning y técnicas híbridas, como los sistemas neuronales-simbólicos (NeSy) y DeepProbLog, han demostrado un gran potencial para superar limitaciones, pues estas herramientas combinan el poder del Deep Learning con estructuras lógicas predefinidas, lo que las posiciona como alternativas prometedoras para abordar la complejidad de los problemas relacionados con la inferencia lógica.

Una limitación significativa de este trabajo fue el tiempo reducido para realizar una investigación más exhaustiva y la disponibilidad de textos en idiomas diferentes a inglés. A pesar de estos retos, los hallazgos refuerzan la relevancia de la revisión de modelos o herramientas que combinen lógica, procesamiento del lenguaje natural y aprendizaje automático.

Recomendaciones

Se recomienda que futuras investigaciones profundicen en el estudio y aplicación de enfoques híbridos, como los sistemas neuronales-simbólicos (NeSy) y DeepProbLog, debido a su capacidad para combinar el aprendizaje profundo con estructuras lógicas predefinidas. Además, se sugiere que futuras investigaciones se planteen como proyectos colaborativos o de mayor duración, que permitan explorar con mayor profundidad las herramientas y evaluar sus resultados

en escenarios aplicados. Esto facilitará una evaluación más precisa del impacto y la efectividad de estas herramientas en el procesamiento de lenguaje natural aplicado a enunciados de razonamiento lógico.

Dado que la mayoría de las fuentes relevantes y los desarrollos recientes están disponibles en inglés, se recomienda que los futuros investigadores desarrollen habilidades sólidas en el idioma para facilitar el acceso a la literatura científica y técnica. Además, podría ser útil crear una base de datos compartida o un repositorio con resúmenes y análisis de artículos clave, lo que podría ahorrar tiempo y esfuerzos a nuevos investigadores en este campo.

Apéndice

Apéndice A: Formato de extracción de datos usado en la presente investigación.

<https://docs.google.com/spreadsheets/d/1VNp1nUJjoUeb6Gg4pXQxNPGGPyGwdKIR/edit?pli=1&gid=906827578#gid=906827578>

Referencias Bibliográficas

- Bench-Capon, T., Dunne, P. E. (2007). Argumentation in artificial intelligence. *Artif. Intell.* 171, 619–641. doi: 10.1016/j.artint.2007.05.001.
<https://www.sciencedirect.com/science/article/pii/S0004370207000793>
- Buttazzo G (2023) Rise of artificial general intelligence: risks and opportunities. *Front. Artif. Intell.* 6:1226990. doi: 10.3389/frai.2023.1226990
- Carmenates, O., Amat, M., Gamboa, M., González, O. (2005). las inferencias lógicas: una vía para desarrollar el aprendizaje del escolar de secundariabásica. Universidad Pedagógica “PepitoTey”.
http://sedici.unlp.edu.ar/bitstream/handle/10915/24629/Documento_completo.pdf3Fsequence=1
- Dasgupta, I., Guo, D., Gershman, S. J., Goodman, N.D. (2020). Analyzing Machine-Learned Representations: A Natural Language Case Study. *Wiley. Cognitive Science: A multidisciplinary journal.*
- Fundación Innovación Bankinter. (2018). En 2025 el volumen de datos será 175 veces más que en 2011. <https://www.fundacionbankinter.org/noticia/big-data-2025-volumen-datos>.
- Galassi A., Kersting K., Lippi M., Shao X., Torroni P. (2020). Neural-Symbolic. Argumentation Mining: An Argument in Favor of Deep Learning and Reasoning. *Frontiers.* <https://www.frontiersin.org/articles/10.3389/fdata.2019.00052/full>
- Han, Z., Cai, LW., Dai, WZ. (2023). Abductive subconcept learning. *Sci. China Inf. Sci.* 66,122103. <https://doi-org.bibliotecavirtual.unad.edu.co/10.1007/s11432-020-3569-0>
- Jaf S., Calder C., (2019). Deep Learning for Natural Language Parsing, *IEEE Access*, vol.7 <https://ieeexploreieeeorg.bibliotecavirtual.unad.edu.co/document/8827454/footnotes#footnotes>

- Jimenez, J., Hernandez A. (2018). Lógica Matemática. Centro Interdisciplinario de Investigación y Docencia en Educación Técnica. https://unicauca.edu.co/matematicas/eventos/log&co/MATERIAL/Elementos_Logica/Textos/Biblioteca/Libros/Libro_008/Logica_Matematica
- Koza, J. R., Bennett, F. H., Andre, D., & Keane, M. A. (1996). Automated design of both the topology and sizing of analog electrical circuits using genetic programming. In J. S. Gero, & F. Sudweeks (Eds.), *Artificial Intelligence in Design '96*. Springer. https://doi-org.bibliotecavirtual.unad.edu.co/10.1007/978-94-009-0279-4_9
- Lauriola, A. Lavelli, F. Aiolli. (2022). An introduction to deep learning in natural language processing: models, techniques, and tools. <https://www-sciencedirect-com.bibliotecavirtual.unad.edu.co/science/article/pii/S0925231221010997?via=ihub#section-cited-by>
- Liu, Z., Zhang, J., Lu, J., Ma, W., & Ho, J. C. (2024). LogicPrpBank: A Corpus for Logical Implication and Equivalence. <https://arxiv.org/pdf/2402.09609>
- Mandal, S., & Naskar, S. K. (2021). Classifying and solving arithmetic math word problems—an intelligent math solver. *IEEE Transactions on Learning Technologies*, 14(1), 28-41. <https://doi.org/10.1109/TLT.2021.3057805>
- Ontanon, S., Ainslie, J., Cvicek, V., & Fisher, Z. (2022). LogicInference: A new Dataset for Teaching Logical Inference to seq2seq Models. In *ICLR2022 Workshop on the Elements of Reasoning: Objects, Structure and Causality*. <https://arxiv.org/abs/2203.15099>
- Salvatore S., Finger M., Hirata R., Patriota A. (2023). A resampling-based method to evaluate NLI models. *Natural Language Engineering*. Published online

- Skansi, S. (2020). Guide to Deep Learning Basics. Springer International. Publishing. https://doi-org.bibliotecavirtual.unad.edu.co/10.1007/978-3-030-37591-1_9
- Wang, L., Zhang, D., Gao, L., Song, J., Guo, L., & Shen, H. T. (2018). MathDQN:Solving Arithmetic Word Problems via Deep Reinforcement Learning. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). <https://doi.org/10.1609/aaai.v32i1.11981>
- Wang, WY., Pan, SJ. (2021). Variational Deep Logic Network for Joint Inference of Entities and Relations. MIT Press Direct. Computational Linguistics. <https://direct.mit.edu/coli/article/47/4/775/106773/Variational-Deep-Logic-Network-for-Joint-Inference-of-Neural-Perception-and-Logical-Reasoning>. IEEE Transactions on Pattern Analysis and Machine Intelligence <https://ieeexplore.ieee.org.bibliotecavirtual.unad.edu.co/document/9831049>
- Xin W., Yi C., Zetao L., Ho-fung L., Tao W. (2023). Generating Natural Language From Logic Expressions With Structural Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. <https://ieeexplore-ieee-org.bibliotecavirtual.unad.edu.co/document/10091889>
- Yao Y., Xu J., Shi J., Xu B.(2018). Learning to activate logic rules for textual reasoning. Science Direct: Journal and Books. Neural networks. <https://doi.org/10.1016/j.neunet.2018.06.012>