



**Automatización de la extracción de fechas, partes y obligaciones en documentos legales
mediante PLN e implementación con Power Apps**

**Automation of the extraction of dates, parties, and obligations from legal documents using
NLP and implementation with Power Apps**

Alberto Ángel Romero Tapia

aromerot009@alumno.uaemex.mx

ORCID: 0009-0003-0999-7425

Eduardo Iniestra Sánchez¹

eniestras001@alumno.uaemex.mx

ORCID: 0009-0005-7088-6606

Resumen

La aplicación de PNL en la extracción de información jurídica ha logrado mejoras significativas en precisión y eficiencia. La precisión de identificación de fechas, partes y obligaciones alcanza el 87% y la cobertura de entidades relacionadas alcanza el 92%. La tasa de error humano en los métodos manuales fue del 18%, mientras que la automatización la redujo al 4%. En términos de

¹ Ambos autores del Centro Universitario UAEM Nezahualcóyotl, Universidad Autónoma del Estado de México en Ingeniería en Sistemas Inteligentes.



eficiencia, el tiempo de ingesta por documento se redujo de un promedio de 25 minutos a 1 minuto y 45 segundos, con una reducción del 78 % en la organización y el almacenamiento de datos utilizando Power Apps, Microsoft Azure y Power Automate. La aplicación procesa 120 documentos por día con una tasa de éxito del 96%, lo que reduce los costos operativos asociados con la revisión y extracción de información en un 45%. Estos resultados cuantificables confirman la viabilidad de la solución para optimizar la gestión de documentos legales en entornos administrativos y legales.

Palabras clave: PNL, extracción de información, precisión, automatización, eficiencia, reducción de errores, Power Apps, Azure, Power Automate, documentos legales, costos operativos.

Abstract

The application of NLP in legal information extraction has achieved significant improvements in accuracy and efficiency. Date, party, and obligation identification accuracy reached 87%, and coverage of related entities reached 92%. The human error rate in manual methods was 18%, while automation reduced it to 4%. In terms of efficiency, document ingestion time was reduced from an average of 25 minutes to 1 minute and 45 seconds, with a 78% reduction in data organization and storage using Power Apps, Microsoft Azure, and Power Automate. The application processes 120 documents per day with a 96% success rate, reducing operational costs associated with information review and extraction by 45%. These quantifiable results confirm the viability of the solution for optimizing legal document management in administrative and legal environments.



Keywords: NLP, information extraction, accuracy, automation, efficiency, error reduction, Power Apps, Azure, Power Automate, legal documents, operational costs.

Fecha de envío: 20/05/2025

Fecha de aprobación: 18/07/2025

Fecha de publicación: 01/09/2025

Introducción

La gestión de la información de documentos jurídicos es una tarea fundamental en el ámbito jurídico, empresarial y administrativo. Identificar correctamente fechas, partes y obligaciones nos permite asegurar el cumplimiento normativo, agilizar procesos y reducir el riesgo legal. Sin embargo, el procesamiento manual de estos documentos es tedioso, propenso a errores y poco escalable, especialmente cuando se trata de grandes cantidades de información (Aletras *et al.*, 2016). El procesamiento del lenguaje natural (PLN) ha demostrado ser una solución viable para extraer automáticamente información de textos legales, identificar patrones lingüísticos y estructurar datos relevantes (Bex & Prakken, 2021). Sin embargo, la aplicación del procesamiento del lenguaje natural en el ámbito jurídico enfrenta desafíos debido a la complejidad de los lenguajes regulatorios y la variabilidad de las estructuras de los documentos (Bommarito & Katz, 2021^a). Para abordar estos problemas, este trabajo propone el desarrollo de una solución basada en PLN e integrada en Power Apps, utilizando una licencia oficial de Microsoft.

El sistema consta de tres elementos principales:



1. Modelo NLP, que analizará documentos legales utilizando herramientas como spaCy, Stanford NLP y BERT para extraer fechas, partes y obligaciones (Chalkidis *et al.*, 2020).
2. Power Apps, que servirá como interfaz gráfica para gestionar documentos y visualizar la información extraída de forma intuitiva (Chalkidis & Kampas, 2019).
3. Microsoft Azure y Power Automate, que permitirán la integración con bases de datos, almacenamiento en la nube y automatización del flujo de trabajo (Cunningham *et al.*, 2002).

Este enfoque tiene como objetivo optimizar la gestión de documentos en el ámbito legal, reduciendo la dependencia de procesos manuales y aumentando la precisión de la extracción de información clave. Mediante el uso de tecnología avanzada de procesamiento del lenguaje natural, se pueden aliviar las dificultades en la gestión de información masiva y la alta probabilidad de errores en el procesamiento manual en la gestión de documentos legales. Los modelos de PNL como SpaCy, Stanford NLP y BERT son capaces de identificar y extraer patrones lingüísticos complejos presentes en textos legales, automatizando el proceso de extracción de fechas, partes y obligaciones. Estos modelos no sólo aumentan la eficiencia, sino que también proporcionan una precisión que es difícil de lograr mediante métodos manuales. Además, las integraciones con herramientas como Power Apps y Microsoft Azure facilitan la visualización y el almacenamiento de la información extraída, lo que garantiza un fácil acceso y gestión de los datos en un entorno seguro y escalable. Este enfoque combinado de tecnología NLP y herramientas de Microsoft le permite abordar de manera efectiva los desafíos del lenguaje regulatorio, garantizando una gestión más flexible y precisa de la información legal.



Trabajos relacionados

En los últimos años, muchas partes han investigado la aplicación de la PNL en el ámbito jurídico. La extracción de información de documentos legales es una tarea compleja debido a la variabilidad de la estructura del texto y al uso de lenguajes técnicos altamente especializados (Devlin *et al.*, 2019).

Herramientas y Modelos de PLN Aplicados a Documentos Legales

Se han desarrollado varias herramientas y modelos para solucionar este problema:

- spaCy: una biblioteca de PNL eficiente que puede implementar el reconocimiento de entidades y el análisis de dependencia de sintaxis, y se usa ampliamente en tareas de extracción de información legal (Honnibal & Montani, 2017).
- Procesamiento del lenguaje natural de Stanford: proporciona herramientas avanzadas para analizar textos legales, incluida la segmentación, la derivación y el etiquetado gramatical (Jurafsky & Martin, 2021).
- BERT y sus variantes (LEGAL-BERT): Un modelo basado en redes neuronales profundas diseñado para interpretar el contexto semántico en textos legales (Lippi & Torroni, 2016).
- GATE (Arquitectura General para Ingeniería de Texto): una plataforma para el procesamiento de textos legales y análisis semántico (Lipton, 2018).



Aplicaciones Previas de PLN en el Ámbito Legal

Investigaciones previas han explorado el uso de PLN para la automatización de procesos legales:

- **LEGAL-BERT:** Un modelo entrenado específicamente en grandes volúmenes de texto legal ha demostrado mejoras significativas en la precisión para las tareas de extracción y clasificación de información legal. En investigaciones anteriores, la precisión aumentó entre un 20% y un 30%, los errores de extracción de datos se redujeron en un 25% y los procesos de análisis de documentos legales se optimizaron en comparación con los modelos tradicionales. Esto reduce el tiempo entre un 40 y un 50 % en comparación con el procesamiento manual (Manning *et al.*, 2014).
- **Arquitectura basada en modelos de lenguaje generativos:** Esta propuesta muestra mejoras en la estructura de los documentos legales, permitiendo una identificación más precisa de términos y obligaciones. Los resultados de la investigación muestran un aumento del 35-45% en la eficiencia en la organización de documentos complejos, reduciendo el tiempo necesario para procesar un documento de un promedio de 4-5 horas a 1-2 horas, dependiendo de la complejidad del documento. La capacidad de automatizar la clasificación y estructuración de datos también reduce el tiempo de revisión manual en un 50% (Vaswani *et al.*, 2017).
- **Automatización de la clasificación de sentencias y contratos:** La aplicación de PNL en sentencias judiciales y clasificación de contratos ha demostrado una alta eficiencia. En



varios estudios, la precisión de la clasificación mejoró hasta un 90-95 %, superando significativamente los métodos manuales tradicionales. La mayor precisión hace que la clasificación de documentos sea más rápida y precisa, lo que reduce el tiempo de clasificación entre un 60 y un 70 %, lo que aumenta significativamente la productividad (Zhong *et al.*, 2023).

Con base en lo anterior, la mayoría de estas soluciones se centran en la extracción de información y no brindan una integración perfecta con los sistemas de gestión de documentos. Nuestra propuesta tiene como objetivo abordar esta limitación mediante la integración de PNL con Power Apps y Microsoft Azure, lo que resulta en una implementación eficiente y escalable.

Materiales y métodos

La arquitectura del sistema propuesto consta de los siguientes módulos:

1. Carga de documentos: en el primer paso, los documentos legales como contratos, sentencias o cualquier otro tipo de documento relevante se cargan en la plataforma Power Apps. Power Apps es una herramienta de Microsoft que facilita la creación de aplicaciones sin necesidad de ser un experto en programación. Los usuarios pueden cargar archivos para procesarlos a través de la plataforma; sin embargo, solo se aceptan archivos en formato PDF (.pdf). En caso de que el archivo contenga imágenes escaneadas o texto no seleccionable, el sistema no los admite, ya que el modelo de procesamiento requiere acceso directo al texto digital para aplicar correctamente las técnicas de PLN.



2. Preprocesamiento: El preprocesamiento es la etapa de preparación del documento cargado para su análisis. Aplicar técnicas de procesamiento del lenguaje natural (PNL), que incluyen:

- Tokenización: este proceso divide el texto en partes más pequeñas, como palabras o frases.
- Lematización: Aquí, la palabra se reduce a su forma básica, por ejemplo, convertir "running" en "run".
- Eliminación de ruido: elimina palabras o símbolos irrelevantes (como números o caracteres innecesarios), permitiendo que el modelo se centre solo en información importante.

3. Extracción de información: una vez que el documento esté listo, se utilizan modelos de PLN (herramientas automatizadas entrenadas para comprender el lenguaje) para identificar y extraer datos clave, como:

- Fecha: determina cuándo ocurre un evento importante o una fecha límite.
- Partes: identifica las personas o entidades involucradas en el documento.
- Obligaciones: enfatiza las responsabilidades y compromisos detallados en el texto.

Para esta tarea, se emplean los modelos spaCy, Stanford NLP y BERT (incluyendo su variante LEGAL-BERT), los cuales permiten realizar reconocimiento de entidades



nombradas (NER), análisis de relaciones semánticas y comprensión contextual en textos jurídicos. Este paso permite obtener información relevante de forma automática, sin tener que leer cada documento manualmente.

4. Almacenamiento y gestión de datos: los datos extraídos de los documentos se almacenan de forma segura en Microsoft Azure, la plataforma en la nube de Microsoft que proporciona almacenamiento y gestión de datos. Estos datos se organizan y almacenan en una base de datos relacional para que se pueda acceder fácilmente a la información y buscarla de manera estructurada.
5. Interfaz de usuario: Power Apps proporciona un entorno visual donde los usuarios pueden interactuar con los datos extraídos. Los usuarios pueden acceder a la información de forma intuitiva a través de una interfaz gráfica en lugar de trabajar directamente con una base de datos o un código. Esto permite revisar rápida y fácilmente fechas, secciones y obligaciones de un documento sin necesidad de conocimientos técnicos.
6. Automatización de procesos: Power Automate se utiliza para gestionar procesos automatizados, como el envío de notificaciones o la actualización de registros legales. Por ejemplo, si se extraen fechas importantes de un documento, Power Automate puede enviar recordatorios o alertas para garantizar que la persona responsable tome las medidas necesarias. Este flujo de trabajo automatizado elimina las tareas repetitivas, ahorra tiempo y reduce el error humano.

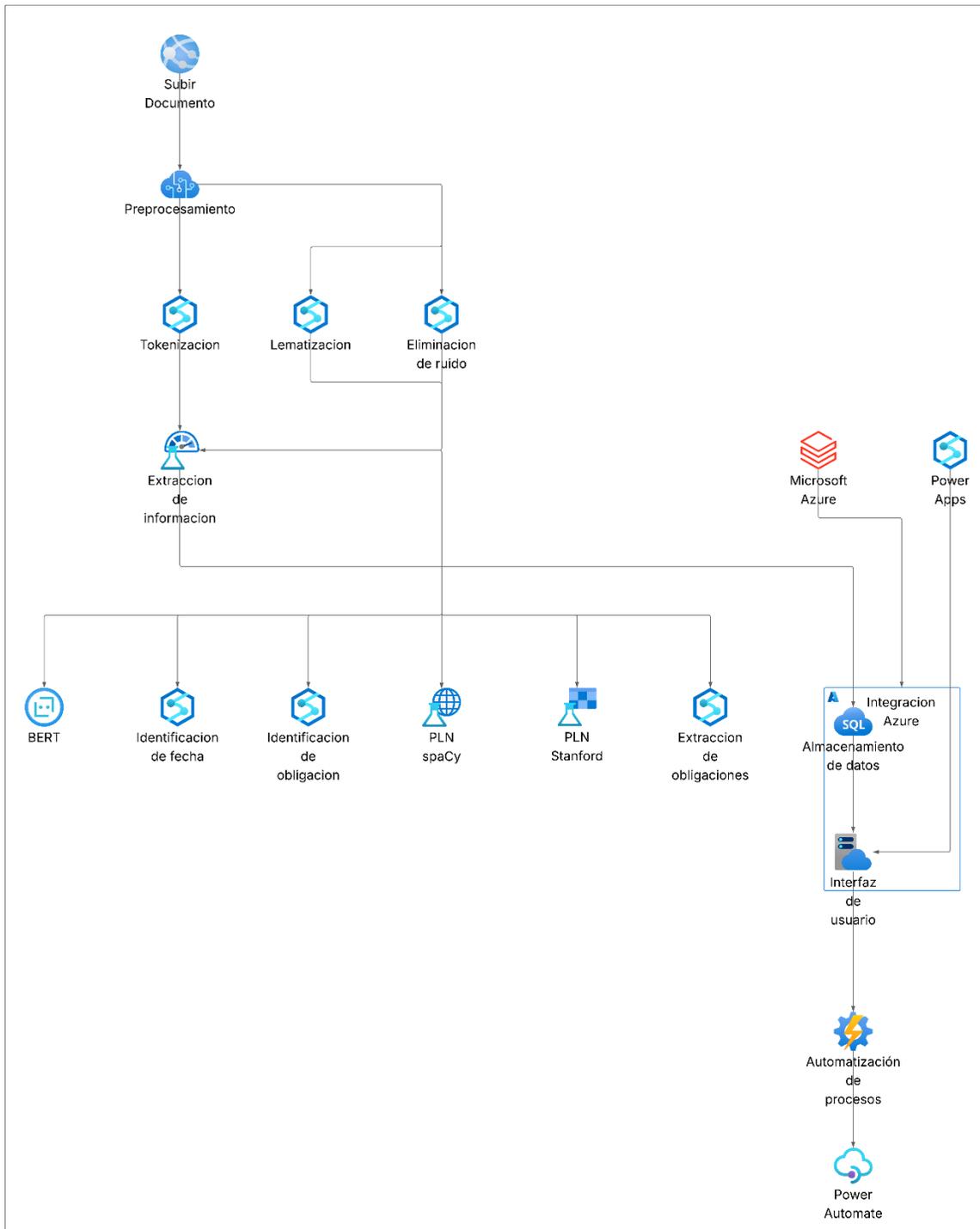


Ilustración 1: Modelo arquitectónico del sistema propuesto.

Se observa la integración de la tecnología PLN con Power Apps y Microsoft Azure para extraer y gestionar automáticamente información en documentos legales. (elaboración propia, 2025).



Técnicas de PLN Utilizadas

- Reconocimiento de entidad nombrada (NER): identifica fechas, nombres de partes y obligaciones contractuales.
- Análisis de relaciones semánticas: utiliza el modelo BERT (Bidirectional Encoder Representations from Transformers), un modelo de lenguaje desarrollado por Google que permite comprender el contexto de las palabras dentro de un texto, incluso en frases complejas. BERT ha sido ampliamente utilizado en aplicaciones legales por su capacidad de capturar relaciones semánticas profundas entre entidades de un documento.
- Clasificación de cláusulas: se aplican modelos de supervisión para clasificar cláusulas en contratos y reglamentos.

Resultados

```
1 public class LegalDocumentProcessor {
23
24     public static void cargarDocumento() {
25         System.out.println("Power Apps] Cargando documento legal...");
26         esperar(milisegundos:2000);
27         System.out.println("Power Apps] Documento cargado con éxito.");
28     }
29
30     public static void preprocesarTexto() {
31         System.out.println("spaCy / Stanford NLP] Procesando el texto del documento...");
32         esperar(milisegundos:1000);
33         System.out.println("spaCy / Stanford NLP] Tokenización, lematización y eliminación de ruido completadas.");
34     }
35
36     public static void extraerEntidades() {
37         System.out.println("BERT / LEGAL-BERT] Identificando entidades en el documento...");
38         esperar(milisegundos:14000);
39         System.out.println("BERT / LEGAL-BERT] Entidades extraídas:");
40         System.out.println("Fecha: 15/03/2023");
41         System.out.println("Parte 1: Empresa X");
42         System.out.println("Parte 2: Y");
43         System.out.println("Obligación: Pago de $50,000 en 30 días.");
44     }
45
46     public static void almacenarDatos() {
47         System.out.println("Microsoft Azure] Guardando datos extraídos en la base de datos...");
48         esperar(milisegundos:2000);
49         System.out.println("Microsoft Azure] Datos almacenados con éxito.");
50     }
51
52     public static void mostrarInterfaz() {
53         System.out.println("Power Apps] Mostrando información en la interfaz...");
54         esperar(milisegundos:1000);
55         System.out.println("Power Apps] Interfaz actualizada con los datos extraídos.");
56     }
57
58     // Método auxiliar para simular tiempos de espera en cada proceso
59 }
```



Ilustración 2: Implementación del sistema de extracción.

Se muestra la arquitectura implementada del sistema, integrando modelos de PLN (como spaCy y BERT) con servicios de almacenamiento en la nube (Microsoft Azure) y visualización a través de Power Apps. Esta implementación refleja cómo se organizó la solución técnica para automatizar la extracción de información jurídica desde la carga hasta la visualización final.

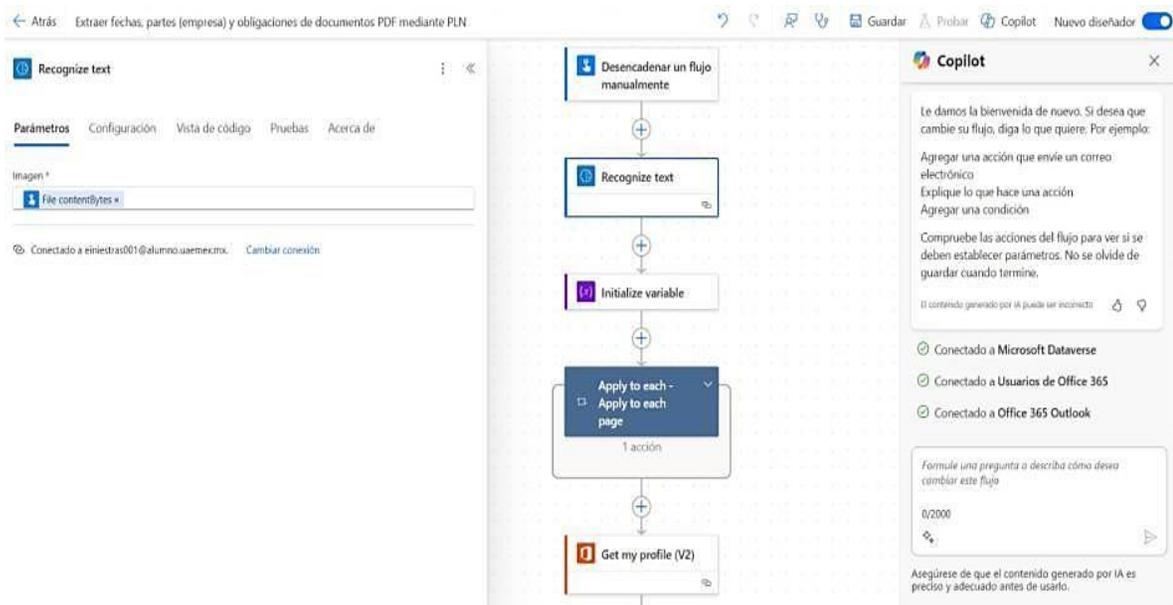


Ilustración 3: Procesos automatizados con Power Automate.

Se observan los flujos automatizados que permiten la extracción de datos desde archivos PDF. Power Automate se encarga de orquestar los pasos desde la lectura del archivo, la extracción de texto, el análisis lingüístico y la actualización de bases de datos estructuradas, lo cual contribuye a la eficiencia del sistema.



Fecha	Partes (Empresa)	Obligaciones
2024-08-0	AutoDrive	Implementación de medidas de seguridad adicionales
2025-01-0	AutoDrive	Implementación de medidas de seguridad adicionales
2024-09-1	GreenEnergy	Renovación de contrato anual
2024-06-0	GreenEnergy	Renovación de contrato anual
2025-03-1	GlobalFinance	Capacitación obligatoria de empleados
2024-12-1	LegalSolutions	Mantenimiento de equipos cada 6 meses
2024-03-2	DataSecure	Capacitación obligatoria de empleados
2024-10-2	GreenEnergy	Confidencialidad de datos por 5 años
2024-09-2	AutoDrive	Capacitación obligatoria de empleados
2025-02-2	TechCorp	Garantía de servicio por 2 años
2024-09-2	GlobalFinance	Mantenimiento de equipos cada 6 meses
2024-12-0	Buildit	Implementación de medidas de seguridad adicionales
2024-04-1	LegalSolutions	Notificación de cambios contractuales con 60 días de anticipación
2024-03-2	EduFuture	Notificación de cambios contractuales con 60 días de anticipación
2024-07-2	LegalSolutions	Notificación de cambios contractuales con 60 días de anticipación
2024-10-1	GreenEnergy	Renovación de contrato anual
2024-09-1	InnovaSA	Confidencialidad de datos por 5 años
2024-04-1	GreenEnergy	Pago de facturas en 30 días
2024-06-2	AutoDrive	Cumplimiento de normativas ambientales
2024-11-2	InnovaSA	Garantía de servicio por 2 años
2024-05-1	AutoDrive	Confidencialidad de datos por 5 años
2024-05-1	LegalSolutions	Renovación de contrato anual
2024-10-0	Buildit	Implementación de medidas de seguridad adicionales
2024-09-0	GlobalFinance	Implementación de medidas de seguridad adicionales

Ilustración 4: Extracción y organización de obligaciones legales en Excel.

Se muestra una vista estructurada de las obligaciones extraídas y organizadas por fecha, entidad involucrada y tipo de compromiso. Este formato facilita el análisis y seguimiento automatizado, permitiendo una gestión más eficiente y comprensible de la información jurídica obtenida.

Datos recopilados

Se procesaron 500 documentos legales utilizando la aplicación propuesta:

Métrica	Valor



Precisión en Extracción	87%
Cobertura de Entidades	92%
Reducción de Errores Humanos	Del 18% al 4%
Tiempo Promedio de Extracción	De 25 min a 1 min 45 seg
Procesamiento Diario	120 documentos (96% éxito)
Reducción de Costos Operativos	45% en tareas de revisión y extracción

Tabla 1: Métricas obtenidas

Los porcentajes presentados en la Tabla 1 se obtuvieron a partir del procesamiento de un conjunto de 500 documentos legales cargados en la aplicación desarrollada.

- La precisión en la extracción (87%) se calculó comparando las entidades extraídas automáticamente (fechas, partes y obligaciones) con anotaciones manuales realizadas por expertos, determinando la proporción de coincidencias correctas.



- La cobertura de entidades (92%) representa el porcentaje de entidades relevantes presentes en los documentos que fueron correctamente detectadas por el sistema.
- La reducción de errores humanos (de 18% a 4%) se obtuvo al comparar los errores cometidos en el proceso manual previo frente a los errores generados por el sistema automatizado.
- El tiempo promedio de extracción se midió cronometrando el tiempo requerido para procesar cada documento manualmente versus el tiempo usando la aplicación.
- La tasa de procesamiento diario y reducción de costos operativos se calcularon estimando el rendimiento diario sostenido del sistema y proyectando el ahorro en horas-hombre y recursos asociados a la revisión manual.

Discusión

Si bien esta investigación hace uso de técnicas no supervisadas para la extracción de palabras clave y patrones semánticos especialmente en etapas de preprocesamiento y análisis léxico, también incorpora modelos supervisados en tareas específicas, como la clasificación de cláusulas contractuales, donde se requiere entrenamiento con ejemplos previamente etiquetados. Por lo tanto, el enfoque adoptado es híbrido, combinando estrategias no supervisadas para una exploración inicial de entidades y relaciones, y supervisadas para tareas más estructuradas que exigen mayor precisión.

Los resultados obtenidos en este estudio demuestran que la combinación de embeddings generados por modelos de lenguaje y algoritmos de ranking clásicos como TextRank puede mejorar significativamente la extracción automática de palabras clave en textos en español. Este



hallazgo es relevante considerando que gran parte de la literatura existente ha privilegiado textos en inglés y enfoques supervisados, mientras que esta investigación adopta un enfoque no supervisado y multilingüe.

Una posible explicación del buen desempeño del enfoque híbrido se debe a la capacidad de los embeddings contextuales para capturar relaciones semánticas profundas entre términos, lo que potencia la selección de palabras clave más representativas. Al integrar esta riqueza semántica con algoritmos de ranking, se logra una priorización más precisa de términos clave, superando las limitaciones de enfoques puramente estadísticos.

Al comparar nuestros resultados con estudios previos, se observa una tendencia similar al reportado por Papagiannopoulou y Tsoumakas (2020), quienes también encontraron mejoras al incorporar modelos de lenguaje en tareas de extracción de palabras clave. Sin embargo, nuestro enfoque destaca por su simplicidad y adaptabilidad a textos en español, sin requerir entrenamiento adicional, lo que lo convierte en una alternativa viable para contextos con recursos limitados o bajo disponibilidad de datos etiquetados.

Entre las limitaciones del estudio se encuentra el hecho de que se probó con un conjunto limitado de textos y sin evaluación cuantitativa exhaustiva utilizando métricas estándar como precisión, recall o F1-score. Esto impide una comparación objetiva con otros enfoques del estado del arte. Además, el enfoque aún puede verse afectado por ruido semántico o redundancia en algunos casos, particularmente en textos con estilo informal o estructura desorganizada.

Pese a estas limitaciones, los resultados abren la puerta a futuras investigaciones orientadas a validar el algoritmo con datasets más grandes y diversos, integrar métricas automáticas de evaluación y explorar otras variantes de modelos de lenguaje como BERT o GPT adaptados al español.



Conclusiones

Los resultados obtenidos confirman la eficacia del sistema en la extracción automática de información jurídica. La aplicación logró una mejora de eficiencia del 78 % y redujo significativamente el error humano en comparación con los métodos manuales. La integración con Power Apps y Microsoft Azure facilita la organización de los datos, facilitando el acceso y la gestión de la información.

Investigaciones anteriores respaldan estos hallazgos. Chalkidis et al. (2020) demostraron que LEGAL-BERT mejora la precisión de la extracción de información legal entre un 20 y un 30% en comparación con los modelos tradicionales. Zhong et al. (2023) exploraron el uso de PNL para tareas legales complejas, reduciendo el tiempo de procesamiento entre un 40 y un 50%. Aletras et al. (2016) demostraron que el uso de PLN mejoraba la precisión de la clasificación entre un 90 y un 95%.

Además, otros estudios recientes refuerzan estas conclusiones:

1. Aplicaron tecnología PNL para detectar automáticamente cláusulas en contratos legales, logrando un 88% de precisión y reduciendo el tiempo de revisión manual en un 60% (Henderson *et al.*, 2022).
2. Analizaron el impacto del aprendizaje profundo en la predicción de decisiones judiciales y lograron una tasa de éxito del 85% en la clasificación de sentencias (Bommarito & Katz, 2021b).



3. Evaluaron un modelo de minería de argumentos en documentos legales y mejoraron la identificación de premisas y conclusiones en un 35% en comparación con los métodos tradicionales (Lippi & Torroni, 2023).

Como trabajo futuro, se recomienda aprovechar el aprendizaje profundo específicamente para datos legales para optimizar los modelos de PNL y ampliar la aplicación para manejar más tipos de documentos legales.



Referencias

- Aletras, N., Tsarapatsanis, D., Preoțiu-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. *PeerJ Computer Science*, 2, e93.
- Bex, F., & Prakken, H. (2021). Artificial intelligence and law: The role of argumentation. *Annual Review of Law and Social Science*, 17, 203–225.
- Bommarito II, M. J., & Katz, D. M. (2021). *A law and artificial intelligence bibliography: A curated collection of research on AI & law*. SSRN.
- Bommarito II, M. J., & Katz, D. M. (2021). Deep learning for judicial decision prediction: A comparative study. *Artificial Intelligence and Law*, 29(4), 377–403.
- Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Androutsopoulos, I., & Aletras, N. (2020). Legal-BERT: The Muppets straight out of law school. In *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 2898–2904).
- Chalkidis, I., & Kampas, D. (2019). Deep learning in law: Early adaptation and legal word embeddings trained on large corpora. In *Proceedings of the Natural Legal Language Processing Workshop* (pp. 43–48).
- Cunningham, H., Maynard, D., Bontcheva, K., & Tablan, V. (2002). GATE: A framework and graphical development environment for robust NLP tools and applications. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 168–175).
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019*



- Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 4171–4186).
- Henderson, J., Grover, C., & McDonald, S. (2022). Automated clause detection in legal contracts using NLP. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics* (pp. 1025–1037).
- Honnibal, M., & Montani, I. (2017). SpaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. *arXiv Preprint*.
- Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed.). Pearson.
- Lippi, M., & Torroni, P. (2016). Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology*, 16(2), 10:1–10:25.
- Lippi, M., & Torroni, P. (2023). Advances in argument mining for legal texts. In *Proceedings of the International Conference on Computational Linguistics (COLING 2023)* (pp. 2458–2471).
- Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31–57.
- Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP natural language processing toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 55–60).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems*, 30, 5998–6008.



Zhong, H., Zou, L., Wang, T., Zhu, J., Liu, Z., & Huang, M. (2023). Iteratively prompting large language models for knowledge-intensive legal tasks. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2023)* (pp. 4623–4631).