From Data to Guidance: Artificial Intelligence for Judicial Process Assistance

A. Indurain-Ibero[1,2], M. Ferrero-Jaurrieta[2], and J. Fernández[2]

[1] Tracasa Instrumental, C. Cabárceno 6, 31621 Sarriguren, España
[2] Universidad Pública de Navarra, 31006 Pamplona, España
alfonso.indurain@unavarra.es

1. Datos

- **Datos de contacto:**
  - Nombre del doctorando: Alfonso Indurain Ibero
  - Email: alfonso.indurain@unavarra.es
- **Propuesta de título:** From Data to Guidance: Artificial Intelligence for Judicial Process Assistance
- **Directores:**
  - Francisco Javier Fernández Fernández, Universidad Pública de Navarra.
  - Mikel Ferrero Jaurrieta, Universidad Pública de Navarra.
- **Universidad:** Departamento de Estadística, Informática y Matemáticas de la Universidad Pública de Navarra (UPNA).
- **Fecha de primera matrícula:** 1 de septiembre de 2023.

2. Antecedentes

Los sistemas de recomendación son herramientas de gran valor que ayudan a sus usuarios a seleccionar las mejores opciones entre un gran abanico de posibilidades. Ya sean series, libros, o incluso documentos oficiales, estos sistemas aprenden de los usuarios para ayudarles en sus procesos de toma de decisión. Sin embargo, no todos los sistemas de recomendación son iguales ya que deben adaptarse a los datos o usuarios con los que deben tratar, y por tanto, deben desarrollarse con un objetivo en mente. Este fue el razonamiento de la empresa Tracasa Instrumental, encargada del diseño y creación de los sistemas de gestión procesal para el Gobierno de Navarra, cuando vio que su gestor procesal, AVANTIUS, disponía de una gran cantidad de plantillas judiciales a disposición de sus usuarios. Debido a la complejidad de los datos de carácter judicial, en esta tesis doctoral, se recoge todo el proceso de diseño, creación y análisis de un sistema de apoyo a la tramitación guiada basada en modelos de aprendizaje profundo, así como los retos y objetivos secundarios que han ido surgiendo durante su desarrollo y que han acabado derivando en diferentes áreas de estudio. Estas áreas incluyen la imputación de datos faltantes [1], la fusión de información mediante funciones de agregación [2, 3] y la cuantificación de incertidumbre [4].

3. Objetivos

El objetivo principal de esta tesis doctoral es la creación de un sistema de ayuda a la gestión procesal basado en modelos de aprendizaje profundo. Con este objetivo en mente, el trabajo se desglosa en los siguientes objetivos:

- Estudio de los sistemas de recuperación de información presentes en la literatura.
- Diseño y propuesta de una nueva familia de operadores de agregación diseñados para la imputación de datos faltantes para conjuntos de datos que no siguen una distribución normal.
- Desarrollo de una arquitectura, basada en modelos de transformer, para la predicción de procesamientos judiciales.
- Estudio y aplicación de métodos de fusión de información para diferentes fuentes de información.
- Estudio y aplicación de métodos de cuantificación de la incertidumbre en modelos de aprendizaje profundo para analizar el sistema desarrollado.

4. Metodología y Plan de Trabajo

Las tareas desarrolladas se pueden agrupar en tres marcos diferentes: el modelo de ayuda a la gestión procesal, imputación de datos faltantes y cuantificación de incertidumbre. Todas estas tareas y su planificación quedan recogidas en el cronograma incluido en la Tabla 1.

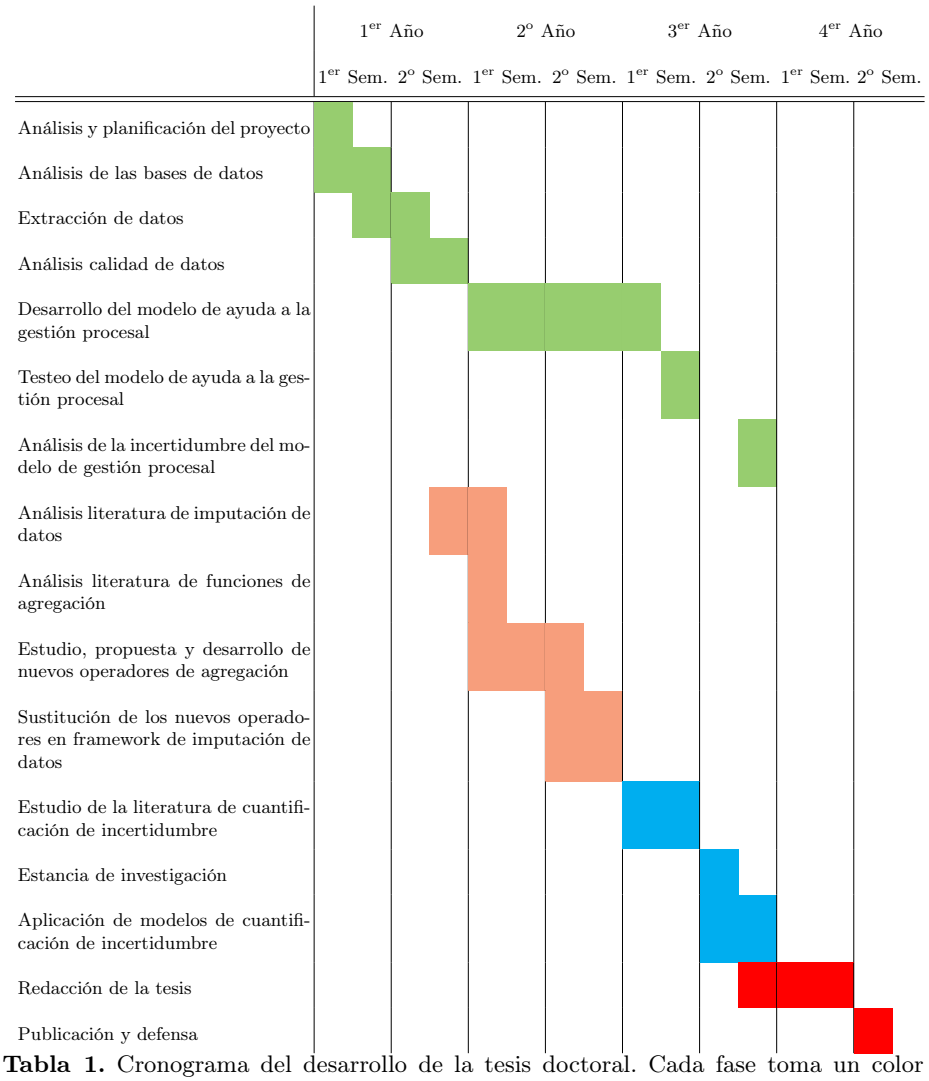


Tabla 1. Cronograma del desarrollo de la tesis doctoral. Cada fase toma un color distinto en el diagrama, representando los objetivos de la tesis. En verde las fases relacionadas con la creación del sistema de gestión procesal, en naranja las fases relacionadas con la imputación de datos faltantes, en azul las fases relacionadas con la cuantificación de incertidumbre y en rojo las fases relacionadas con la redacción y defensa de la tesis doctoral.

A continuación se desglosa en mayor detalle cada una de ellas.

4.1. Modelo de ayuda a la gestión procesal

El modelo de ayuda a la gestión procesal es el objetivo principal de la tesis. Para ofrecer una solución a este problema se decidió trabajar en la elaboración de un sistema de recomendación que proporcionase los mejores documentos a tramitar en cada paso del proceso. Para ello, se ha hecho uso de modelos *transformer* y mecanismos de fusión de información para combinar las diferentes fuentes de datos. El modelo resultante no solo aprende haciendo uso del histórico de tramitaciones previas, sino que es capaz de diferenciar entre diferentes usuarios y sus diferentes tendencias. Eso ha permitido generar un modelo que aprende cómo tramita el usuario y ofrece recomendaciones ajustadas a cada uno.

Las tareas que se han llevado a cabo para realizar esta parte del trabajo han sido:

- **Análisis y planificación del proyecto:** durante esta fase se realizó un estudio detallado del proceso de gestión procesal y de las necesidades de los usuarios finales del sistema.
- **Análisis de la base de datos:** se llevó a cabo un estudio exhaustivo de la estructura y contenido de las bases de datos utilizadas en el proyecto. Este proceso permitió identificar las entidades relevantes para la tramitación judicial.
- **Extracción de datos:** se desarrollaron los procedimientos necesarios para la recopilación y transformación de los datos procedentes de los sistemas de gestión procesal. Como resultado, se construyó un conjunto de datos estructurado que integra información histórica de expedientes e información identificativa de los usuarios.
- **Análisis de la calidad de los datos:** se evaluó la calidad de los datos extraídos mediante el estudio de aspectos como la completitud, consistencia y presencia de valores anómalos.
- **Desarrollo del modelo de ayuda a la tramitación:** se diseñó e implementó una arquitectura basada en modelos *transformer* y mecanismos de fusión de información capaces de integrar múltiples fuentes de datos. El modelo desarrollado aprende a partir del histórico de tramitaciones realizadas y de las características de los usuarios, permitiendo generar recomendaciones personalizadas sobre los documentos más adecuados para tramitar en cada etapa del procedimiento.
- **Test del modelo de ayuda a la tramitación:** se realizaron diferentes experimentos para evaluar el rendimiento del sistema propuesto en escenarios representativos de uso real. La evaluación incluyó el análisis de métricas de recomendación, permitiendo validar la capacidad del modelo para proporcionar recomendaciones relevantes.

Las tareas que quedan pendientes de realizar son:

- **Análisis de la incertidumbre del modelo de gestión procesal**: esta tarea se centrará en la aplicación y evaluación de técnicas de cuantificación de incertidumbre sobre el modelo desarrollado. El objetivo será estimar la confianza asociada a las recomendaciones generadas, identificar situaciones en las que el sistema presenta una mayor incertidumbre y estudiar mecanismos que permitan comunicar dicha información al usuario final. Este análisis contribuirá a incrementar la transparencia, robustez y fiabilidad del sistema de ayuda a la tramitación en un dominio crítico como el judicial.

4.2. Imputación de datos faltantes

La calidad de los datos constituye uno de los factores más determinantes en el desarrollo de sistemas de inteligencia artificial fiables. En particular, la presencia de valores faltantes representa un problema habitual en numerosos dominios. La utilización de conjuntos de datos incompletos puede introducir sesgos, reducir la capacidad predictiva de los modelos y comprometer la fiabilidad de los sistemas de apoyo a la toma de decisiones.

Con el objetivo de mitigar estos problemas, una de las líneas de investigación de esta tesis se centra en el estudio y desarrollo de técnicas de imputación de datos faltantes. Más concretamente, se investiga cómo el uso de operadores de agregación avanzados puede contribuir a mejorar la calidad de las imputaciones realizadas, especialmente en escenarios caracterizados por distribuciones de datos complejas. Las actividades desarrolladas en esta línea se describen a continuación:

- **Análisis de la literatura de imputación de datos faltantes:** se ha realizado una revisión sistemática de la literatura científica relacionada con la imputación de datos faltantes, prestando especial atención a los métodos basados en aprendizaje automático e inteligencia artificial. Este estudio ha permitido identificar las principales técnicas empleadas para el tratamiento de información incompleta, así como sus ventajas, limitaciones y ámbitos de aplicación.
- **Análisis de la literatura de funciones de agregación:** esta tarea ha abordado el estudio de las principales funciones y operadores de agregación utilizados en inteligencia artificial, poniendo especial interés en los Ordered Weighted Averaging (OWA) y sus modificaciones. Este estudio ha permitido encontrar un hueco de interés en la literatura de los operadores Induced Ordered Weighted Averaging (IOWA), directamente relacionado con las características de los datos del proyecto [5–12].
- **Propuesta y desarrollo de nuevos operadores de agregación para la imputación de datos faltantes:** tras el análisis realizado en la tarea anterior, se descubrió que no existía literatura previa de los operadores IOWA donde se tuviese en cuenta la distribución de los datos de entrada, característica importante para nuestro proyecto debido a la gran variedad en la distribución de nuestros datos. Esto ha resultado en la propuesta de una nueva familia de operadores, basados en los IOWA que han sido diseñados teniendo en mente la distribución de los datos de entrada.

- **Sustitución de los nuevos operadores en arquitecturas existentes de imputación de datos:** los operadores propuestos han sido incorporados en arquitecturas de imputación de datos con el fin de evaluar su impacto sobre el rendimiento de los modelos.

Este objetivo de la tesis ha resultado en la publicación del siguiente trabajo en la revista Applied Soft Computing [13]

4.3. Cuantificación de incertidumbre

La gestión de la incertidumbre constituye un aspecto fundamental en el desarrollo de sistemas de inteligencia artificial para la asistencia a la tramitación judicial. En este contexto, la calidad de los datos de entrada y las capacidades del modelo influyen directamente en la fiabilidad de las recomendaciones generadas. Sin embargo, incluso cuando se dispone de mecanismos de imputación para tratar datos incompletos y de modelos avanzados, persisten fuentes de incertidumbre derivadas tanto de la información disponible como de las limitaciones inherentes al propio modelo. La identificación, cuantificación y comunicación de esta incertidumbre resultan esenciales para garantizar un uso responsable de la tecnología, permitiendo detectar situaciones en las que las predicciones pueden ser menos fiables.

Con este objetivo en mente, se han realizado las siguientes tareas:

- **Estudio de la literatura: se ha estudiado la literatura actual de cuantificación de incertidumbre para modelos de IA:** Se ha realizado una revisión exhaustiva de la literatura científica sobre cuantificación de incertidumbre en modelos de inteligencia artificial, con especial atención a los modelos de lenguaje de gran escala (LLMs). Este estudio ha permitido identificar y analizar diferentes familias de enfoques para la estimación de la incertidumbre, incluyendo métodos basados en las probabilidades generadas por el modelo, técnicas de muestreo y agregación de respuestas, aproximaciones fundamentadas en la consistencia semántica de las predicciones, así como modificaciones en las arquitecturas orientadas a obtener estimaciones de confianza más fiables. Los conocimientos obtenidos a partir de esta revisión servirán como base para la selección, implementación y evaluación de las técnicas más adecuadas en el contexto de la asistencia a la tramitación judicial [4, 14–24].

Las tareas que están en desarrollo actualmente:

- **Aplicación de modelos de cuantificación de incertidumbre en datasets judiciales:** esta tarea comprende la implementación y adaptación de diferentes técnicas de cuantificación de incertidumbre sobre los conjuntos de datos judiciales seleccionados. El objetivo es analizar la capacidad de estos métodos para estimar la confianza de las predicciones generadas en escenarios reales de tramitación judicial. Posteriormente, se llevará a cabo

una evaluación comparativa exhaustiva de los modelos, considerando métricas relacionadas tanto con el rendimiento predictivo como con la calidad de las estimaciones de incertidumbre. Los resultados obtenidos permitirán identificar las técnicas más adecuadas para su integración en sistemas de asistencia judicial basados en inteligencia artificial, favoreciendo una toma de decisiones más transparente y fiable.

5. Resultados

Actualmente, esta tesis doctoral ha resultado en articulo, publicado en la revista Applied Soft Computing, titulado "RIOWA: Restricted Induced Ordered Weighted Averaging operators, applied to missing data imputation.". En este trabajo se ha introducido una nueva familia de operadores de agregación, basados en los operadores "Induced Ordered Weighted Averaging" (IOWA), especialmente diseñados para el tratamiento de datos con distribuciones no uniformes para la imputación de datos faltantes, siendo esta aplicación uno de los objetivos de la tesis doctoral.

6. Relevancia

La relevancia de esta tesis doctoral se puede observar tanto en el plano científico como en su aplicación.

Desde el punto de vista científico, la investigación realizada durante esta tesis doctoral ha resultado en el desarrollo de una nueva familia de operadores de agregación diseñados para la imputación de datos faltantes y con características previamente no presentes en operadores similares. Al mismo tiempo, el trabajo que actualmente se esta llevando acabo en el estudio para cuantificar la incertidumbre en LLMs aplicados en el ámbito judicial plantea un nuevo camino de estudio que no había sido tratado hasta ahora.

Sin embargo, la mayor aportación de este trabajo viene desde su aplicación en el sistema de gestión procesal AVANTIUS, donde se prevé que entre en funcionamiento durante los próximos meses. El modelo propuesto ya ha salido de la fase de desarrollo en laboratorio y se encuentra a la espera de sus primeras pruebas con usuarios reales.

Referencias

1. Z. Zhang, "Missing data imputation: Focusing on single imputation," *Annals of Translational Medicine*, vol. 4, no. 1, p. 9, 2016.
2. G. Beliakov, A. Pradera, T. Calvo, et al., *Aggregation Functions: A Guide for Practitioners*, vol. 221, Springer, 2007.
3. M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, *Aggregation Functions*, vol. 127, Cambridge University Press, 2009.

4. X. Liu, T. Chen, L. Da, C. Chen, Z. Lin, and H. Wei, "Uncertainty quantification and confidence calibration in large language models: A survey," in *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pp. 6107–6117, 2025.
5. R. R. Yager, On ordered weighted averaging aggregation operators in multicriteria decisionmaking, IEEE Transactions on systems, Man, and Cybernetics 18 (1) (1988) 183–190.
6. F. Herrera, E. Herrera-Viedma, J. L. Verdegay, Aggregating linguistic preferences: properties of lowa operator, in: Proc. 5th IFSA World Congress, Sao Paulo, Citeseer, 1995, pp. 153–156.
7. J. Fodor, J.-L. Marichal, M. Roubens, Characterization of the ordered weighted averaging operators, IEEE Transactions on Fuzzy Systems 3 (2) (2002) 236–240.
8. V. Torra, The weighted owa operator, International journal of intelligent systems 12 (2) (1997) 153–166.
9. R. R. Yager, Nonmonotonic owa operators, Soft computing 3 (3) (1999) 187–196.
10. R. R. Yager, D. P. Filev, Induced ordered weighted averaging operators, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 29 (2) (1999) 141–150.
11. R. R. Yager, Heavy owa operators, Fuzzy optimization and decision making 1 (2002) 379–397.
12. R. R. Yager, Generalized owa aggregation operators, Fuzzy optimization and decision making 3 (1) (2004) 93–107.
13. A. Indurain-Ibero, B. Bedregal, M. Ferrero-Jaurrieta, J. Fernandez, and H. Bustince, "RIOWA: Restricted induced ordered weighted averaging operators, applied to missing data imputation," *Applied Soft Computing*, p. 115188, 2026.
14. S. Kadavath, T. Conerly, A. Askell, T. Henighan, D. Drain, E. Perez, N. Schiefer, Z. Hatfield-Dodds, N. DasSarma, E. Tran-Johnson, et al., "Language models (mostly) know what they know," *arXiv preprint arXiv:2207.05221*, 2022.
15. Z. Jiang, F. F. Xu, J. Araki, and G. Neubig, "How can we know what language models know?" *Transactions of the Association for Computational Linguistics*, vol. 8, pp. 423–438, 2020.
16. Z. Lin, S. Trivedi, and J. Sun, "Generating with confidence: Uncertainty quantification for black-box large language models," *arXiv preprint arXiv:2305.19187*, 2023.
17. M. Xiong, Z. Hu, X. Lu, Y. Li, J. Fu, J. He, and B. Hooi, "Can LLMs express their uncertainty? An empirical evaluation of confidence elicitation in LLMs," *arXiv preprint arXiv:2306.13063*, 2023.
18. X. Wang, J. Wei, D. Schuurmans, Q. Le, E. Chi, S. Narang, A. Chowdhery, and D. Zhou, "Self-consistency improves chain of thought reasoning in language models," *arXiv preprint arXiv:2203.11171*, 2022.
19. P. Manakul, A. Liusie, and M. Gales, "SelfCheckGPT: Zero-resource black-box hallucination detection for generative large language models," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 9004–9017, 2023.
20. L. Kuhn, Y. Gal, and S. Farquhar, "Semantic uncertainty: Linguistic invariances for uncertainty estimation in natural language generation," *arXiv preprint arXiv:2302.09664*, 2023.
21. Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," *arXiv preprint arXiv:1506.02142*, 2015.

22. A. Azaria and T. Mitchell, "The internal state of an LLM knows when it's lying," in *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 967–976, 2023.
23. S. Zhang, S. Roller, N. Goyal, M. Artetxe, M. Chen, S. Chen, C. Dewan, M. Diab, X. Li, X. V. Lin, et al., "OPT: Open pre-trained transformer language models," *arXiv preprint arXiv:2205.01068*, 2022.
24. S. Farquhar, J. Kossen, L. Kuhn, and Y. Gal, "Detecting hallucinations in large language models using semantic entropy," *Nature*, vol. 630, no. 8017, pp. 625–630, 2024.