

Tarea: Investigación sobre el uso de la IA en tu Carrera Profesional

Facundo Bautista Barbera A01066843

16 de septiembre de 2025

Introducción

Estoy estudiando Ingeniería en Ciencia de Datos, que básicamente se trata de crear sistemas que puedan analizar grandes cantidades de datos y sacar información útil de ellos. En mi carrera vamos a diseñar pipelines de datos, crear modelos que puedan predecir cosas y desarrollar sistemas de machine learning para ayudar a las empresas a tomar mejores decisiones. En esta tarea voy a investigar cómo la inteligencia artificial puede ayudarnos a ser más productivos en nuestro trabajo como futuros científicos de datos. Este documento es un boceto de una actividad mas completa, por lo que no debe ser tomada una versión final.

Casos de uso de la IA en mi carrera

La IA en ciencia de datos es algo que vemos de cerca en clases y proyectos. Para esta investigación escogí cuatro **casos de uso principales** que me parecieron interesantes para mostrar cómo la IA puede ayudarnos a ser más eficientes en nuestro trabajo.

1. **Optimización en Ciencia de Datos:** La IA contribuye en problemas de optimización matemática, heurísticas y optimización de hiperparámetros mediante sistemas AutoML.
2. **Análisis Multivariado:** Algoritmos de IA potencian la reducción de dimensionalidad, la clasificación y la regresión. Técnicas como autoencoders variacionales y t-SNE permiten generar representaciones más interpretables y robustas de datos complejos, facilitando el descubrimiento de patrones ocultos.

3. **Topología y Geometría de Datos:** La IA permite descubrir estructuras ocultas en conjuntos de datos complejos mediante el análisis topológico de datos (TDA). Aplicando conceptos de homología persistente y métricas topológicas, se pueden identificar características invariantes que son robustas al ruido y perturbaciones.
4. **Simulación y Modelado:** La IA apoya en la simulación de sistemas complejos y en la generación de escenarios hipotéticos. Los modelos generativos como GANs y modelos de difusión facilitan la creación de datos sintéticos para pruebas y validación, mejorando la toma de decisiones estratégicas.

Prompt General

Utilizando prompt engineering, se diseñó un prompt base para todos los casos de uso:

“Actúa como un experto en ciencia de datos. Con base en el siguiente tema: [TEMA ESPECÍFICO], investiga fuentes académicas recientes (2023-2025) y genera una presentación ejecutiva que explique: 1) Fundamentos teóricos del concepto, 2) Aplicaciones prácticas con ejemplos reales de empresas, 3) Métricas cuantitativas de mejora en productividad (porcentajes, tiempos reducidos, ROI), 4) Limitaciones actuales y consideraciones éticas, 5) Tendencias futuras. Incluye referencias verificables y código ejemplo cuando sea relevante.”

Metodología Detallada

Para cada caso de uso se siguió un proceso sistemático:

Para conseguir la información usé varias herramientas de IA como ChatGPT-5, Claude Opus 4.1, Gemini Pro 2.5, Microsoft Copilot y Meta AI (modelo local usando ML Studio, no logró hacer la presentación). Fui mejorando mis prompts poco a poco según las respuestas que me daban, aplicando técnicas de prompt engineering y también con ayuda de otras inteligencias artificiales. También tuve que verificar que las referencias que me sugerían fueran reales buscando directamente en bases de datos académicas, porque a veces las IA se inventan citas. (Nota: Logre hacer que prácticamente todas las referencias que me escribieron fueran reales.)

Después comparé las respuestas de las diferentes IA para ver qué tenían en común y en qué se diferenciaban. Tuve que editar y mejorar un poco el contenido que generaron, corrigiendo errores y haciendo que todo tuviera más sentido. Al final decidí entregar la presentación sin muchas modificaciones ya que no estaba seguro si era necesario modificar mucho para esta primera entrega.

Resultados y Hallazgos Principales

Impacto en la Productividad

En mi investigación encontré que la IA realmente puede ayudarnos mucho a ser más productivos en ciencia de datos. Creo que en particular (fuera de lo que encuentre) también puedo decir que el uso mas particular para la inteligencia artificial que tengo como científico de datos es que me ayuda con la programación.

Otra cosa interesante es que la optimización bayesiana con estrategias evolutivas funciona mejor que los métodos tradicionales para buscar hiperparámetros. Generalmente para este tipo de pruebas se hacen procesos como el grid search y también validaciones cruzadas. La IA puede ayudar a automatizar la decisión de cuáles hiperparámetros vamos a probar. Esto significa que no tenemos que ser expertos en todo desde el principio.

Conclusiones

Después de hacer toda esta investigación, creo que el uso de IA en ciencia de datos realmente ha cambiado mucho la forma en que trabajamos. Lo que más se puede destacar es cómo la IA puede ayudarnos a ser mucho más productivos automatizando tareas complejas y repetitivas, lo que nos deja tiempo para enfocarnos en la solución de los problemas.

Pero también me di cuenta de que es muy importante mantener un balance entre automatización y supervisión humana. La IA es una herramienta muy poderosa, pero siempre necesitamos validar constantemente los resultados y considerar aspectos éticos. Creo que el futuro va a estar marcado por la colaboración entre humanos y la IA, donde cada uno aporta lo que sabe hacer mejor.

En general, la inteligencia artificial está siendo una herramienta que tiene un crecimiento constante (en todos los aspectos) y no podemos detenerlo. Es mejor adaptarnos a utilizarla, sin quitar la parte del trabajo humano.

Referencias

1. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. Disponible en: <https://hastie.su.domains/ElemStatLearn/>
2. Bronstein, M., Bruna, J., Cohen, T., & Velicković, P. (2021). “Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges.” *arXiv preprint arXiv:2104.13478*. Disponible en: <https://arxiv.org/abs/2104.13478>
3. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Disponible en: <https://www.deeplearningbook.org/>
4. Vincent, A.M., & Jidesh, P. (2023). “An improved hyperparameter optimization framework for AutoML systems using evolutionary algorithms.” *Scientific Reports*, 13, 4737. <https://doi.org/10.1038/s41598-023-32027-3>
5. Bischl, B., Binder, M., Lang, M., et al. (2023). “Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges.” *WIREs Data Mining and Knowledge Discovery*, 13(2), e1484. <https://doi.org/10.1002/widm.1484>
6. Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., & Harrington, H. A. (2017). “A roadmap for the computation of persistent homology.” *EPJ Data Science*, 6, 17. <https://doi.org/10.1140/epjds/s13688-017-0109-5>
7. Chazal, F., & Michel, B. (2021). “An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists.” *Frontiers in Artificial Intelligence*, 4, 667963. <https://doi.org/10.3389/frai.2021.667963>
8. Davenport, T. H., & Bean, R. (2024). “Five Key Trends in AI and Data Science for 2024.” *MIT Sloan Management Review*. Disponible en: <https://sloanreview.mit.edu/article/five-key-trends-in-ai-and-data-science-for-2024/>
9. Dell’Acqua, F., et al. (2023). “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality.” *Harvard Business School*

Working Paper. Disponible en: <https://www.hbs.edu/faculty/Pages/item.aspx?num=64700>

10. Bick, A., Blandin, A., & Deming, D. (2025). "The Rapid Adoption of Generative AI." *Federal Reserve Bank of St. Louis Working Paper 2024-027C*. Disponible en: <https://doi.org/10.20955/wp.2024.027>
11. McKinsey Global Institute. (2023). "The economic potential of generative AI: The next productivity frontier." McKinsey & Company. Disponible en: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/>
12. Salehin, I., Islam, M. S., Saha, P., et al. (2024). "AutoML: A systematic review on automated machine learning with neural architecture search." *Journal of Information and Intelligence*, 2(1), 52-81. <https://doi.org/10.1016/j.jiixd.2023.10.002>
13. Su, Z., Liu, X., Hamdan, L. B., et al. (2025). "Topological Data Analysis and Topological Deep Learning Beyond Persistent Homology—A Review." *arXiv preprint arXiv:2507.19504*. Disponible en: <https://arxiv.org/abs/2507.19504>
14. Desai, M. (2024). "Introducing Code-First AutoML and Hyperparameter Tuning: Now in Public Preview for Fabric Data Science." *Microsoft Fabric Blog*. Disponible en: <https://blog.fabric.microsoft.com/en-us/blog/introducing-code-first-automl>
15. Carlsson, G. (2009). "Topology and data." *Bulletin of the American Mathematical Society*, 46(2), 255-308. <https://doi.org/10.1090/S0273-0979-09-01249-X>