

La inteligencia artificial en la investigación científica

Rogelio Jesús Corrales Díaz
Aned Esquerra Arguelles
Beatriz Alejandra García Ramos
Oscar Alejandro Hernández López
Alberto Martínez Noa
Gerardo Palafox Castillo
Gabriela Sánchez Yepez

30 de mayo 2023

1. Introducción

La inteligencia artificial (IA) ha emergido como una de las tecnologías más revolucionarias y prometedoras de nuestra era. Su capacidad para simular la inteligencia humana y realizar tareas complejas ha generado un gran impacto en diversos campos, y la investigación científica no es una excepción. En la búsqueda constante de avances y descubrimientos, los científicos han recurrido cada vez más a las tecnologías de inteligencia artificial para acelerar y optimizar su labor investigativa. En este ensayo, exploramos el creciente papel que desempeña la inteligencia artificial en la investigación científica. Analizaremos cómo las tecnologías de IA están transformando la forma en que los científicos abordan los desafíos, desde la recopilación y análisis de datos hasta la generación de hipótesis y la toma de decisiones. También examinaremos los beneficios y las posibles preocupaciones éticas asociadas con el uso de la inteligencia artificial en la investigación científica.

La integración de la inteligencia artificial en la investigación científica ha dado lugar a avances significativos en una amplia gama de disciplinas, incluyendo la medicina, la biología, la física, la química y muchas más. Por ejemplo, la Tabla 1 muestra herramientas de IA que se han utilizado para el descubrimiento de nuevos fármacos. Específicamente en el área de investigación de operaciones, la inteligencia artificial se ha utilizado para distintas tareas como: el análisis de grandes conjuntos de datos, pronóstico de demandas de distintos productos y/o servicios que, a su vez, sirve para identificar oportunidades de mejora en distintos procesos (inventario, transporte, planificación, etc). Diversos trabajos han estudiado aplicaciones de predicción de demanda [3, 13, 15]. Mehmood *et al.* [15] propusieron un esquema para predecir la demanda de transporte urbano utilizando usando *Big Data*. López Lázaro *et al.* [13] utilizan *Machine Learning* para predecir la demanda de efectivo en sucursales bancarias, esta información se emplea para optimizar el transporte de efectivo. Por su parte, Chung *et al.* [3] analizaron datos históricos de vuelos en distintos aeropuertos para predecir el retraso de los vuelos. Los resultados se usan como datos de entrada en un modelo de optimización para determinar el número necesario de tripulantes de reserva.

En cuanto a la creación de rutas mediante distintas herramientas de IA se encuentra el trabajo de Becker *et al.* [1], quienes analizan el rendimiento una red neuronal simplificada para tomar decisiones de ruteo sobre un conjunto de datos reales obtenidos por la terminal de automóviles del puerto de Hamburgo. Liu *et al.* [12] emplean distintas

herramientas de IA como *Big Data*, *Machine Learning*, *Internet of Things* para determinar rutas de transporte óptimas y dinámicas a través de las especificaciones del usuario para servicios de lavandería.

Sin embargo, a medida que la IA se vuelve más prominente en la investigación científica, también surgen preguntas y preocupaciones. ¿Cómo afecta la validez y la confiabilidad de los resultados científicos el uso de algoritmos de IA? ¿Qué implicaciones éticas y sociales deben ser consideradas al utilizar estas tecnologías en la investigación? Estas son algunas de las cuestiones que examinaremos en el desarrollo de este ensayo.

En resumen, la inteligencia artificial está desempeñando un papel cada vez más importante en la investigación científica, brindando nuevas oportunidades y desafíos a los investigadores. A medida que nos adentramos en una era impulsada por la tecnología, comprender el impacto y el potencial de la inteligencia artificial en la investigación científica se vuelve esencial para avanzar en el conocimiento y abordar los problemas más complejos que enfrenta nuestra sociedad.

2. Desarrollo

La última década ha visto un crecimiento considerable en el uso de la inteligencia artificial para la gestión de operaciones con el objetivo de encontrar soluciones a problemas que aumentan en complejidad y escala, en su participación en “AI and OR in management of operations: history and trends” [9], hace una revisión extensiva de 1200 artículos publicados en Scienedirect que abarca el período de 1995 a 2004, clasificando los documentos en cuatro áreas de gestión de operaciones: (a) diseño, (b) programación, (c) planificación y control de procesos y (d) calidad, mantenimiento y diagnóstico de fallas, sub clasificando las áreas en dependencia de los algoritmos de Inteligencia Artificial empleado para las soluciones y dejando un campo abierto al análisis de las tendencias de la década y las posibles soluciones futuras sugeridas.

Muchos de los campos relacionados con la Investigación de Operaciones que de cierta forma enriquecen esta área del conocimiento, cada vez desdibujan sus fronteras para irse integrando como técnicas que permiten lograr resultados interesantes y totalmente nuevos al combinarse [8, 18, 19, 20, 21], la Inteligencia Artificial ha tenido un progreso sin precedentes desde su surgimiento y adopción por la mayoría de las empresas y áreas de aplicación, las cadenas de suministro, los servicios de salud [6, 16], la agricultura, la industria financiera [5] e innumerables áreas del quehacer humano se ven atraídas por la capacidad de los métodos de Inteligencia Artificial de encontrar soluciones atractivas y nuevas a problemas complejos y de difícil solución [4, 11].

Una de las herramientas esenciales de la Inteligencia Artificial utilizada en Investigación de Operaciones es el aprendizaje automático, aprendizaje de máquina o también conocido como machine learning por su nombre en inglés. El aprendizaje de máquina es una rama de Inteligencia Artificial que implica algoritmos de capacitación para aprender de los datos y tomar predicciones o decisiones. En Investigación de Operaciones, el área de Machine Learning puede emplearse para el mantenimiento predictivo, el pronóstico de la demanda y la optimización de la cadena de suministro. Por ejemplo, los algoritmos Machine learning pueden analizar datos históricos para predecir fallas de equipos y mantenimiento del cronograma antes de que ocurran, minimizando el tiempo de inactividad y aumentando la eficiencia. Machine learning también puede analizar los datos del cliente para predecir los patrones de demanda y optimizar los niveles de inventario, reducir los desechos y mejorar la satisfacción del cliente.

Existen diversas herramientas de IA para apoyar la investigación. Entre ellas están:

- Semantic Scholar: Motor de búsqueda académica
- Bit.ai: Organización de investigación.
- Scholarcy: Resumen de la investigación.
- Scite: Evaluación de citas.

- Trinka: Redacción de documentos de investigación.
- chatGPT: modelo generativo del lenguaje

En particular, chatGPT puede apoyar de las siguientes formas:

- Proporcionar un acceso rápido y fácil a la información: los investigadores de operaciones pueden tener preguntas o necesitar información sobre un tema en particular. ChatGPT puede proporcionar rápidamente información relevante de su extensa base de datos de conocimiento, ayudando a los investigadores a ahorrar tiempo y esfuerzo [2, 7, 10, 14].
- Ayuda con el análisis de datos: la investigación de operaciones a menudo implica analizar grandes cantidades de datos. ChatGPT puede ayudar a los investigadores con tareas como la limpieza de datos, el preprocesamiento y el modelado, lo que hace que el proceso sea más eficiente [17].
- Generación de información: los investigadores de operaciones pueden necesitar generar información a partir de datos complejos. ChatGPT puede ayudar con el procesamiento y la comprensión del lenguaje natural para generar ideas a partir de datos no estructurados, como comentarios de los clientes, revisiones y datos de redes sociales.
- Proporcionar apoyo para tomar decisiones: los investigadores de operaciones pueden necesitar tomar decisiones estratégicas basadas en datos y modelos complejos. CHATGPT puede ayudar proporcionando información relevante y sugiriendo posibles cursos de acción basados en datos históricos y análisis predictivos.
- Explicando modelos y resultados: los investigadores de operaciones pueden necesitar explicar sus modelos y resultados a las partes interesadas que pueden no tener antecedentes técnicos. ChatGPT puede ayudar proporcionando explicaciones en un lenguaje simple que es fácil para el público no técnico de entender.

3. Limitaciones

En la etapa de revisión de literatura, se encontraron limitantes serios en algunas de las herramientas más populares de inteligencia artificial. Por ejemplo, se le solicitó a chatGPT hablar sobre procesos epidémicos en complejos simpliciales. La definición general que da del tema es bastante acertada, sin embargo, la literatura recomendada es problemática. Si bien los artículos 1 y 3 hablan sobre epidemias, no abordan el tema específico de epidemias en estructuras de orden mayor o epidemias en complejos simpliciales. Los artículos 4 y 5 no hablan del tema ni de algún tema relacionado. Por último, los artículos 2 y 6 no existen. Cabe destacar que R. Lambiotte, uno de los autores del “artículo 6”, sí es un investigador del área mencionada. Ante la inconsistencia, se le solicitó a chatGPT proporcionar el DOI de uno de estos artículos, y nos devuelve uno no existente. Otra de las herramientas que se exploró fue la de Paper Digest, misma a la que se le pidió hacer un resumen de literatura con los keywords Epidemics Simplicial Complex. El resultado fue mejor si se compara con chatGPT, empezando con el hecho de que no arrojó trabajos ficticios. El listado de literatura incluye trabajos relevantes sobre el área, sin embargo, también hay trabajos algo distanciados del tema. Además el resumen que proporciona y supone un atractivo de la herramienta, es muy básico. Parece el resultado de sacar una oración de cada artículo listado. Como investigador, Google Scholar ofrece mejores resultados.

Herramienta	Descripción	Sitio web
AlphaFold	Predicción de estructuras 3D de proteínas	https://deepmind.com/blog/alphafold
Chemputer	Formatos estandarizados para reportar procedimientos de síntesis química	https://zenodo.org/record/1481731
DeepChem	Una herramienta basada en Python para varias tareas predictivas relacionadas con el descubrimiento de fármacos	https://github.com/deepchem/deepchem
DeepNeuralNet- QSAR	Predicciones de actividad molecular	https://github.com/Merck/DeepNeuralNet-QSAR
DeepTox	Predicciones de toxicidad	www.bioinf.jku.at/research/DeepTox
Hit Dexter	Modelos de aprendizaje máquina para la predicción de moléculas que respondan a procedimientos bioquímicos	http://hitdexter2.zbh.uni-hamburg.de
Neural Graph Fingerprints	Predicción de propiedades de moléculas nuevas	https://github.com/HIPS/neural-fingerprint
ODDT	Herramientas para el uso en química-informática y modelado molecular	https://github.com/oddt/oddt
ORGANIC	Herramienta para generar moléculas con propiedades específicas	https://github.com/aspuru-guzik-group/ORGANIC
PPB2	Predicción polifarmacológica	http://ppb2.gdb.tools/
REINVENT	Desarrollo molecular <i>de novo</i> usando redes neuronales recurrentes y aprendizaje reforzado	https://github.com/MarcusOlivecrona/REINVENT
SCScore	Evaluador de la complejidad de sintetizar una molécula	https://github.com/connorcoley/scscore

Tabla 1: Herramientas de IA para el descubrimiento de fármacos

What about epidemic or contagion processes on simplicial complexes?

Simplicial complexes are higher-dimensional analogs of graphs and networks, and have been used to study a variety of complex systems, including epidemic and contagion processes. Here are some papers that deal with epidemic or contagion processes on simplicial complexes:



1. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, 2(3), 203-271.
2. Bauke, H. (2007). Epidemic threshold in clustered networks. *Physical Review E*, 75(2), 020901.
3. Bailey, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and its Applications*. John Wiley & Sons.
4. Kuznetsov, Y. A. (2013). *Elements of Applied Bifurcation Theory*. Springer.
5. Scardovelli, R., & Zaleski, S. (1999). Direct numerical simulation of free-surface and interfacial flow. *Annual Review of Fluid Mechanics*, 31(1), 267-292.
6. Delvenne, J. C., Evans, T., Omodei, E., & Lambiotte, R. (2015). Diffusion on networks. *Reviews of Modern Physics*, 87(3), 953-979.

These papers should provide you with a good understanding of the mathematical foundations of contagion processes on simplicial complexes, and give you some ideas for further study.

Can you give me the DOI of paper 6 from your response?

Certainly! Paper 6 from my previous response, "Diffusion on networks" by Delvenne, J. C., Evans, T., Omodei, E., & Lambiotte, R., was published in the *Reviews of Modern Physics* in 2015 and its DOI is: 10.1103/RevModPhys.87.953.



Figura 1: chatGPT sugiriendo artículos.



DOI NOT FOUND

10.1103/RevModPhys.87.953

This DOI cannot be found in the DOI System. Possible reasons are:

- The DOI is incorrect in your source. Search for the item by name, title, or other metadata using a search engine.
- The DOI was copied incorrectly. Check to see that the string includes all the characters before and after the slash and no sentence punctuation marks.
- The DOI has not been activated yet. Please try again later, and report the problem if the error continues.

Figura 2: DOI inexistente del artículo sugerido por chatGPT.

Related Work

- [1] Annalisa Buffa; Snorre H. Christiansen; "A Dual Finite Element Complex on The Barycentric Refinement", [MATH. COMPUT.](#), 2007. (IF: 6)
- [2] Li Ding; Ping Hu; "Contagion Processes on Time-Varying Networks with Homophily-Driven Group Interactions", [COMPLEX.](#), 2019.
- [3] Iacopo Iacopini; Giovanni Petri; Alain Barrat; Vito Latora; "Simplicial Models of Social Contagion", [NATURE COMMUNICATIONS](#), 2019. (IF: 6)
- [4] Joan T. Matamalas; Sergio Gómez; Alex Arenas; "Abrupt Phase Transition Of Epidemic Spreading In Simplicial Complexes", [ARXIV-PHYSICS.SOC-PH](#), 2019. (IF: 3)
- [5] Ginestra Bianconi; Sergey N. Dorogovtsev; "The Spectral Dimension Of Simplicial Complexes: A Renormalization Group Theory", [ARXIV-COND-MAT.DIS-NN](#), 2019. (IF: 3)
- [6] Wenyao Li; Xiaoyu Xue; Liming Pan; Tao Lin; Wei Wang; "Competing Spreading Dynamics in Simplicial Complex", [ARXIV-PHYSICS.SOC-PH](#), 2021. (IF: 3)
- [7] Dong Wang; Yi Zhao; Jianfeng Luo; Hui Leng; "Simplicial SIRS Epidemic Models with Nonlinear Incidence Rates", [CHAOS \(WOODBURY, N.Y.\)](#), 2021. (IF: 3)
- [8] Gerardo Palafox-Castillo; Arturo Berrones-Santos; "Stochastic Epidemic Model on A Simplicial Complex", [ARXIV-COND-MAT.STAT-MECH](#), 2022.
- [9] Huan Wang; Hai-Feng Zhang; Pei-Can Zhu; Chuang Ma; "Interplay of Simplicial Awareness Contagion and Epidemic Spreading on Time-varying Multiplex Networks", [CHAOS \(WOODBURY, N.Y.\)](#), 2022.
- [10] Wenjie Li; Yanyi Nie; Wenyao Li; Xiaolong Chen; Sheng Su; Wei Wang; "Two Competing Simplicial Irreversible Epidemics on Simplicial Complex", [CHAOS \(WOODBURY, N.Y.\)](#), 2022.

Summary of the Related Work

It can be seen as a realization of the simplicial cochain complex. (Buffa et. al., 2007) construct a new complex Y^* of finite element spaces on the barycentric refinement of the mesh which can be seen as a realization of the simplicial chain complex on the original (unrefined) mesh, such that the L^2 duality is non-degenerate on $Y^i \times X^{2-i}$ for each $i \in \{0,1,2\}$. (Ding et. al., 2019) investigate the impact of homophily-driven group interactions on the contagion process in temporal networks. (Iacopini et. al., 2019) introduce a higher-order model of social contagion in which a social system is represented by a simplicial complex and contagion can occur through interactions in groups of different sizes. (Matamalas et. al., 2019) cope with the problem of epidemic spreading, using the Susceptible-Infected-Susceptible (SIS) model, in simplicial complexes. Using the renormalization group here (Bianconi et. al., 2019) calculate the spectral dimension of the graph Laplacian of two classes of non-amenable d -dimensional simplicial complexes: the Apollonian networks and the pseudo-fractal networks. (Li et. al., 2021) analyze the model's critical phenomena. (Wang et. al., 2021) propose a simplicial susceptible-infected-recovered-susceptible (SIRS) model to investigate the epidemic spreading via combining the network higher-order structure with a nonlinear incidence rate. (Palafox-Castillo et. al., 2022) study stochastic epidemic model on a simplicial complex. A stochastic epidemic model on a simplicial complex is defined, generalizing the known Markovian SIR epidemic process on networks. The physical social interaction of individuals is not static but time-varying. Therefore (Wang et. al., 2022) propose a novel sUAU-tSIS model to characterize the interplay of simplicial awareness contagion and epidemic spreading on time-varying multiplex networks, in which one layer with 2-simplicial complexes is considered the virtual information layer to address the complex contagion mechanisms in awareness diffusion and the other layer with time-varying and memory effects is treated as the physical contact layer to mimic the temporal interaction pattern among population. A competing spread model for two simplicial irreversible epidemics (i.e., susceptible-infected-

Figura 3: Uso de herramienta Paper Digest.

4. Conclusiones

Hemos explorado el impacto de la inteligencia artificial en la investigación científica y hemos examinado tanto sus beneficios como sus limitaciones. Desde el análisis de grandes volúmenes de datos hasta la generación de hipótesis y la toma de decisiones, las tecnologías de inteligencia artificial han demostrado ser herramientas poderosas que han transformado la forma en que los científicos abordan los desafíos. Entre las herramientas populares de inteligencia artificial utilizadas en la investigación científica, el aprendizaje automático ha jugado un papel destacado. Algoritmos como las redes neuronales convolucionales han demostrado una capacidad excepcional para el reconocimiento de patrones en imágenes y han impulsado el avance en campos como la medicina y la astronomía.

Por otro lado, los sistemas de recomendación basados en IA, como el famoso algoritmo de filtrado colaborativo, han mejorado la eficiencia en la búsqueda y selección de información relevante en investigaciones científicas. Sin embargo, es importante destacar que, a pesar de los impresionantes avances, las tecnologías de inteligencia artificial también tienen limitaciones y desafíos a considerar. Una de las preocupaciones más destacadas es la falta de transparencia y explicabilidad en los resultados generados por los algoritmos de IA. En ocasiones, los modelos de aprendizaje automático pueden ser considerados como “cajas negras”, dificultando la comprensión de cómo se llega a una conclusión o recomendación en particular. Esto plantea preocupaciones en términos de confiabilidad y validez de los resultados científicos obtenidos a través de estos métodos. Además, debemos ser conscientes de las limitaciones inherentes a los conjuntos de datos utilizados en la investigación científica. La calidad y la representatividad de los datos pueden tener un impacto significativo en la precisión y la aplicabilidad de los resultados generados por los algoritmos de IA. El sesgo y la falta de diversidad en los datos pueden llevar a conclusiones inexactas o a resultados que no son generalizables a toda la población. A medida que avanzamos en la integración de la inteligencia artificial en la investigación científica, es fundamental abordar estas limitaciones y desafíos de manera ética y responsable. Se requiere una combinación de conocimiento científico, experiencia humana y la implementación de salvaguardias adecuadas para garantizar la precisión, la equidad y la transparencia en la aplicación de la inteligencia artificial en la investigación científica. A medida que los investigadores continúan explorando las posibilidades de la inteligencia artificial, es esencial seguir fomentando la colaboración

Referencias

- [1] Becker, T., Illigen, C., McKelvey, B., Hülsmann, M., y Windt, K. (2016). Using an agent-based neural-network computational model to improve product routing in a logistics facility. *International Journal of Production Economics*, 174:156–167.
- [2] Borji, A. (2023). A categorical archive of chatgpt failures. *arXiv preprint arXiv:2302.03494*.
- [3] Chung, S. H., Ma, H. L., y Chan, H. K. (2017). Cascading delay risk of airline workforce deployments with crew pairing and schedule optimization. *Risk Analysis*, 37(8):1443–1458.
- [4] Dogru, A. K. y Keskin, B. B. (2020). Ai in operations management: applications, challenges and opportunities. *Journal of Data, Information and Management*, 2:67–74.
- [5] Doumpos, M., Zopounidis, C., Gounopoulos, D., Platanakis, E., y Zhang, W. (2022). Operational research and artificial intelligence methods in banking. *European Journal of Operational Research*.
- [6] Elkefi, S. y Layeb, S. B. (2022). Artificial intelligence and operations research in a middle ground to support decision-making in healthcare systems in africa. *Africa Case Studies in Operations Research: A Closer Look into Applications and Algorithms*, pp. 51–69.
- [7] Frieder, S., Pinchetti, L., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A., y Berner, J. (2023). Mathematical capabilities of chatgpt. *arXiv preprint arXiv:2301.13867*.
- [8] Hebrard, E. y Musliu, N. (2020). *Integration of Constraint Programming, Artificial Intelligence, and Operations Research: 17th International Conference, CPAIOR 2020, Vienna, Austria, September 21–24, 2020, Proceedings*, volumen 12296. Springer Nature.
- [9] Kobbacy, K. A., Vadera, S., y Rasmy, M. H. (2007). Ai and or in management of operations: history and trends. *Journal of the Operational Research Society*, 58:10–28.

- [10] Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., Leon, L. D., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., y Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2(2):e0000198.
- [11] Kusiak, A. (1987). Artificial intelligence and operations research in flexible manufacturing systems. *INFOR: Information Systems and Operational Research*, 25(1):2–12.
- [12] Liu, C., Feng, Y., Lin, D., Wu, L., y Guo, M. (2020). Iot based laundry services: an application of big data analytics, intelligent logistics management, and machine learning techniques. *International Journal of Production Research*, 58(17):5113–5131.
- [13] López Lázaro, J., Álvaro Barbero Jiménez, y Takeda, A. (2018). Improving cash logistics in bank branches by coupling machine learning and robust optimization. *Expert Systems with Applications*, 92:236–255.
- [14] Megahed, F. M., Chen, Y.-J., Ferris, J. A., Knoth, S., y Jones-Farmer, L. A. (2023). How generative ai models such as chatgpt can be (mis) used in spc practice, education, and research? an exploratory study. *arXiv preprint arXiv:2302.10916*.
- [15] Mehmood, R., Meriton, R., Graham, G., Hennelly, P., y Kumar, M. (2016). Exploring the influence of big data on city transport operations: a markovian approach. *International Journal of Operations & Production Management*, 37.
- [16] Rao, A., Kim, J., Kamineni, M., Pang, M., Lie, W., y Succi, M. D. (2023). Evaluating chatgpt as an adjunct for radiologic decision-making. *medRxiv*, pp. 2023–02.
- [17] Sanmarchi, F., Golinelli, D., y Bucci, A. (2023). A step-by-step researcher’s guide to the use of an ai-based transformer in epidemiology: an exploratory analysis of chatgpt using the strobe checklist for observational studies. *medRxiv*, pp. 2023–02.
- [18] Schaus, P. (2022). *Integration of Constraint Programming, Artificial Intelligence, and Operations Research: 19th International Conference, CPAIOR 2022, Los Angeles, CA, USA, June 20-23, 2022, Proceedings*, volumen 13292. Springer Nature.
- [19] van Hoeve, W.-J. y Hooker, J. N. (2009). *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems: 6th International Conference, CPAIOR 2009 Pittsburgh, PA, USA, May 27-31, 2009 Proceedings*, volumen 5547. Springer.
- [20] Vuppapapati, C. (2022). *Artificial Intelligence and Heuristics for Enhanced Food Security*, volumen 331. Springer Nature.
- [21] Yi, L., Min, R., Kunjie, C., Dan, L., Ziqiang, Z., Fan, L., y Bo, Y. (2023). Identifying and managing risks of ai-driven operations: A case study of automatic speech recognition for improving air traffic safety. *Chinese Journal of Aeronautics*, 36(4):366–386.