



Tipo de artículo: Artículos de revisión
Temática: Inteligencia Artificial
Recibido: 28/12/2024 | Aceptado: 20/01/2025 | Publicado: 30/03/2025

Identificadores persistentes:
DOI: [10.48168/innosoft.s23.a208](https://doi.org/10.48168/innosoft.s23.a208)
ARK: [ark:/42411/s23.a208](https://nbn-resolving.org/ark:/42411/s23.a208)
PURL: [42411/s23.a208](https://nbn-resolving.org/urn:nbn:pe:ulasalle:innosoft-23-a208)

El Rol de la Inteligencia Artificial en la Ciencia de Datos: Perspectivas Teóricas y Desafíos Emergentes

The Role of Artificial Intelligence in Data Science: Theoretical Perspectives and Emerging Challenges

Anthony JeanPaul Reyes Risco*¹, Juan Aurelio De La Cruz Gamarra², Marcelino Torres Villanueva³[\[0000-0002-9797-1510\]](https://orcid.org/0000-0002-9797-1510)

¹Universidad Nacional de Trujillo. Trujillo, Perú. areyesri@unitru.edu.pe

²Universidad Nacional de Trujillo. Trujillo, Perú. t1053300821@unitru.edu.pe

³Universidad Nacional de Trujillo. Trujillo, Perú. mtorres@unitru.edu.pe

*Autor para correspondencia: areyesri@unitru.edu.pe

Resumen

El artículo exploró cómo la Inteligencia Artificial (IA) y la Ciencia de Datos han revolucionado la creación y análisis de información al integrar metodologías avanzadas que superan las barreras tradicionales en la interpretación de datos complejos. Se abordaron conceptos fundamentales y desafíos técnicos y éticos actuales, destacando la automatización del ciclo de vida analítico mediante AutoML, la implementación de modelos explicables y la gestión de sesgos algorítmicos. La investigación examinó también las limitaciones de la IA en el procesamiento de datos no estructurados y su interacción con tecnologías emergentes como blockchain y la computación cuántica. Los resultados subrayaron la importancia de establecer normativas que garantizan el equilibrio entre innovación tecnológica y la protección de derechos humanos en un contexto de big data y decisiones automatizadas. Concluye enfatizando que el impacto de la IA trasciende lo técnico, consolidándola como motor de avance interdisciplinario, impulsando tanto el progreso del conocimiento humano como aplicaciones prácticas sostenibles, siempre bajo enfoques éticos y regulados.

Palabras claves: AutoML, Big Data, Ciencia de Datos, Ética algorítmica, Inteligencia Artificial, Modelos Predictivos.

Abstract

The article explored how Artificial Intelligence (AI) and Data Science have revolutionized the creation and analysis of information by integrating advanced methodologies that overcome traditional barriers in the interpretation of complex data. Fundamental concepts and current technical and ethical challenges were addressed, highlighting the automation of the analytical lifecycle through AutoML, the implementation of explainable models and the management of algorithmic biases. The research also examined the limitations of AI in processing unstructured data and its interaction with emerging technologies such as blockchain and quantum computing. The results highlighted the importance of establishing regulations that guarantee the balance between technological innovation and the protection of human rights in a context of big data and automated decisions. It concludes by emphasizing that the impact of AI transcends the technical, consolidating it as an engine of interdisciplinary progress, promoting both the progress of human knowledge and sustainable practical applications, always under

ethical and regulated approaches.

Keywords: *AutoML, Big Data, Data Science, Algorithmic Ethics, Artificial Intelligence, Predictive Models.*

Introducción

En la actualidad, La Ciencia de Datos y la Inteligencia Artificial (IA) han surgido como disciplinas muy importantes para el ámbito de la ciencia y la tecnología. Según [1] se puede definir a la Ciencia de Datos como el campo interdisciplinario dedicado a extraer información y conocimiento de diferentes tipos de datos; mediante métodos analíticos, estadísticos y computacionales. Por otro lado, [2] define a la inteligencia artificial como el desarrollo de sistemas y algoritmos capaces de realizar diversas tareas, las cuales típicamente requerirían de la inteligencia de un humano; como, por ejemplo: el reconocimiento de patrones, la toma de decisiones y el aprendizaje automático.

Estas disciplinas comparten múltiples puntos de intersección teórica, particularmente en el uso de algoritmos avanzados para procesar grandes volúmenes de datos. Por ejemplo, la inteligencia artificial proporciona técnicas como el aprendizaje automático (Machine Learning), que se ha convertido en una herramienta esencial para la Ciencia de Datos, según lo indica [3]. A su vez, la Ciencia de Datos aporta la metodología necesaria para manejar datos a gran escala; desde la recolección de datos, hasta la interpretación de estos mismos. El estudio teórico de la interacción entre IA y Ciencia de Datos resulta de suma importancia para comprender cómo estas disciplinas no solo se complementan, sino también cómo están transformando el mundo científico y empresarial actual. En nuestro mundo, el cual cada vez está más impulsado por los datos; estas tecnologías están redefiniendo procesos clave, desde la personalización en servicios digitales hasta la automatización de decisiones en tiempo real según [4].

Además; hoy en día, debido al auge del big data y las tecnologías de procesamiento masivo, han ampliado las oportunidades y desafíos en este campo. Por ejemplo, la integración de datos no estructurados provenientes de redes sociales, sensores y plataformas digitales requiere enfoques teóricos robustos para maximizar el valor de la información [5]. En este artículo, se explorarán las interacciones teóricas entre la IA y la Ciencia de Datos, destacando sus fundamentos conceptuales, desafíos emergentes y perspectivas futuras, con el objetivo de ofrecer una visión integral sobre el papel de estas disciplinas en la evolución del conocimiento humano.

Materiales y Metodología computacional

Se ha analizado la literatura sobre la interacción entre la inteligencia artificial y la ciencia de datos, explorando marcos teóricos clave para comprender los desafíos éticos, técnicos y metodológicos asociados con el desarrollo y uso de modelos avanzados.

Fundamentos Conceptuales

La Ciencia de Datos es, como su nombre lo indica, una ciencia que combina conocimientos de habilidades computacionales, estadísticos y también experiencia en dominios específicos para extraer información valiosa a partir de datos estructurados y no estructurados permitiendo crear modelos predictivos usando algoritmos de machine Learning.

La ciencia de datos no solo se enfoca en analizar el contexto en el que se encuentran los datos utilizados. El proceso central de la Ciencia de Datos se puede resumir en cuatro etapas:

- **Recolección:** implica la captura de datos desde diferentes fuentes, como bases de datos, sensores, plataformas web y redes sociales.
- **Limpieza:** identificar y corregir errores, eliminar datos redundantes y preparar los conjuntos de datos para su posterior análisis. La calidad de los datos es crítica, ya que incluso los modelos más complejos no pueden compensar datos deficientes.
- **Análisis:** aplicación de métodos estadísticos y algoritmos para identificar patrones, realizar predicciones y generar insights accionables [6].
- **Interpretación:** etapa final, donde se contextualizan los resultados obtenidos para informar decisiones y generar valor en aplicaciones prácticas.

La Ciencia de Datos ha evolucionado desde enfoques tradicionales basados en estadística hasta metodologías más complejas habilitadas por la IA y el aprendizaje automático, lo que amplía su capacidad de problemas multidimensionales y heterogéneos [4].

Interacciones Teóricas entre la IA y La Ciencia de Datos

La relación entre la inteligencia Artificial (IA) y la Ciencia de Datos se basa en una interdependencia conceptual y metodológica que ha redefinido los enfoques en la creación, implementación y evaluación de modelos analíticos. Mientras que la IA proporciona herramientas algorítmicas para procesar y analizar datos complejos, la Ciencia

de Datos establece el marco metodológico que guía la integración de estas herramientas en contextos reales [1]. A continuación, se analizan tres áreas clave donde estas disciplinas interactúan teóricamente.

■ **Teoría de la modelización basada en IA**

La modelización basada en IA representa un cambio de paradigma en la construcción de modelos predictivos y descriptivos. Tradicionalmente, la Ciencia de Datos se apoyaba en enfoques estadísticos como regresión lineal o análisis de varianza. Sin embargo, los avances en la IA han introducido modelos capaces de adaptarse dinámicamente a datos complejos de alta dimensionalidad, como las redes neuronales profundas y los modelos de ensamblaje (ensemble learning) [3].

Desde una perspectiva teórica, los algoritmos de IA trascienden la parametrización predefinida, aprendiendo directamente de los datos y ajustándose iterativamente para optimizar su desempeño. Esto ha generado una nueva categoría de modelización basada en patrones emergentes, lo que desafía los paradigmas tradicionales centrados en hipótesis preestablecidas [7].

Además, la capacidad de la IA para procesar datos no estructurados ha revolucionado campos como el procesamiento del lenguaje natural y la visión por computadora, ampliando significativamente el alcance de la Ciencia de Datos. Sin embargo, esta dependencia de datos masivos plantea preguntas fundamentales sobre la generalización de los modelos en contextos con datos escasos o sesgados [4].

■ **Impacto de los algoritmos en la calidad del análisis de datos**

La calidad del análisis de datos depende en gran medida de las capacidades teóricas y prácticas de los algoritmos utilizados. Aunque los algoritmos de IA han demostrado ser herramientas poderosas para identificar patrones complejos y realizar predicciones, no están exentos de limitaciones. Por ejemplo: los modelos de aprendizaje profundo son notoriamente opacos, lo que dificulta su interpretabilidad y evaluación desde una perspectiva teórica [8].

Los problemas de explicabilidad también afectan la confianza en los resultados generados por los algoritmos. La falta de transparencia en la toma de decisiones de modelos como los de aprendizaje automático ha llevado al desarrollo de enfoques como es la IA Explicable (XAI), la cual busca proporcionar interpretaciones coherentes y comprensibles de las predicciones realizadas [8].

Otra limitación muy importante es la sensibilidad de los algoritmos al sesgo en los datos. Si los datos de entrenamiento contienen sesgos inherentes, estos se reflejan en los resultados, perpetuando inequidades y afectando la validez teórica de los modelos. Esto resalta la importancia de integrar conceptos éticos y epistemológicos en el diseño y la implementación de sistemas de inteligencia artificial.

■ **Relación con otras disciplinas teóricas**

El desarrollo de y aplicación de la IA en la ciencia de datos no puede comprenderse de manera aislada;

está profundamente interconectado con otras disciplinas teóricas como la estadística, las matemáticas y la epistemología:

- **Estadística:** Los fundamentos estadísticos siguen siendo esenciales en la Ciencia de Datos, proporcionando las bases para la validación de modelos y el análisis de incertidumbres. La IA, en particular, ha adoptado conceptos estadísticos como los métodos bayesianos para la predicción y el modelo probabilístico [9].
- **Matemáticas:** La teoría de optimización, el álgebra lineal y el cálculo diferencial son pilares para el diseño de algoritmos de aprendizaje automático. Estas herramientas permiten resolver problemas de ajuste de parámetros y minimización de errores en grandes espacios de búsqueda.
- **Epistemología:** La Ciencia de Datos y la IA plantean preguntas fundamentales sobre la naturaleza del conocimiento derivado de datos. Por ejemplo, ¿Hasta qué punto los modelos automatizados generan conocimiento real? ¿O simplemente reproducen patrones de datos? Este cuestionamiento resalta la necesidad de una integración teórica más profunda entre IA y epistemología [10].

Avances y Desafíos en el Uso de IA en Ciencias de Datos

La aplicación de la Inteligencia Artificial (IA) en la Ciencia de Datos ha generado avances significativos en múltiples áreas. Sin embargo, estos avances también han puesto en tela de juicio una serie de desafíos teóricos y prácticos. Este apartado analiza tres áreas clave: la automatización del ciclo de vida de la ciencia de datos, la ética y la explicabilidad e interpretabilidad de modelos.

Automatización del ciclo de vida de la ciencia de datos

Uno de los avances más destacados en el área de la ciencia de datos, es la automatización del ciclo de vida analítico mediante técnicas de AutoML (Auto Machine Learning). El objetivo principal de AutoML es reducir la intervención humana en tareas críticas como la selección de modelos [11].

Desde el punto de vista teórico, AutoML representa un cambio en la idea de la práctica de la Ciencia de Datos. La automatización no solo democratiza el acceso a herramientas avanzadas, sino que también cuestiona las barreras tradicionales en términos técnicos. Sin embargo, el uso extensivo de AutoML plantea interrogantes sobre la calidad y confiabilidad de los modelos generados automáticamente, especialmente en escenarios donde la interpretación de resultados es fundamental [12].

La automatización puede dar a conocer las decisiones tomadas por el sistema, dificultando la replicación de estudios y comprometiendo la transparencia de los resultados [13].

Ética y Sesgos Algorítmico

El uso de la IA en la ciencia de datos ha puesto en evidencia el debate ético sobre los sesgos inherentes en los algoritmos. Los modelos de IA tienden a amplificar las desigualdades presentes en los datos de entrenamiento, lo que puede llevar a decisiones discriminatorias en diferentes áreas como empleo, salud y sistema judicial [14].

Desde un punto de vista teórico, los sesgos algoritmos son un reflejo de las limitaciones de conocimientos en la ciencia de datos. Los datos no son neutrales, están influenciados por decisiones humanas sobre que recolectar, como recolectar y cómo procesarlo [15].

Explicabilidad e interpretabilidad de modelos

Uno de los desafíos teóricos en la aplicación de IA a la Ciencia de Datos es la explicabilidad de los modelos complejos, como las redes neuronales profundas. Estos modelos, aunque altamente precisos, operan como cajas negras lo que dificulta comprender cómo generan sus predicciones [16].

La explicabilidad se ha convertido en una prioridad debido a la importancia en la toma de decisiones basada en IA. Herramientas como LIME y SHAP han sido desarrolladas para proporcionar interpretaciones locales y globales de los modelos, pero estas técnicas a menudo se enfrentan a limitaciones teóricas y prácticas [8].

También, la falta de explicabilidad tiene implicaciones éticas y legales. En sectores regulados, como la medicina y finanzas, la incapacidad para justificar decisiones algorítmicas puede conducir a problemas de confianza y responsabilidad [17]. El futuro de la investigación en este ámbito apunta a modelos híbridos que combinan la precisión de las redes profundas con la transparencia de enfoques más tradicionales.

Desafíos Teóricos Emergentes

El impacto de la Inteligencia Artificial en la Ciencia de Datos se extiende más allá de sus aplicaciones actuales, generando una serie de desafíos teóricos que redefinen los límites del conocimiento y la capacidad analítica. Este apartado aborda tres áreas críticas:

1. Limitaciones de la IA para Analizar Datos no Estructurados

Los datos no estructurados, como imágenes, texto y señales de audio, representan una gran proporción de los datos generados globalmente. Aunque la IA ha avanzado significativamente en el procesamiento de este tipo de información mediante técnicas como el procesamiento de Lenguaje Natural y la visión por

Computadora, persisten desafíos fundamentales.

Desde una perspectiva teórica, el procesamiento de lenguaje natural enfrenta limitaciones inherentes en la comprensión semántica profunda y el contexto pragmático de los textos. Aunque modelos avanzados como GPT y BERT han demostrado capacidades notables en la generación y clasificación de textos, siguen siendo incapaces de interpretar matices culturales, ironías o referencias implícitas de manera consistente [18]. Por ejemplo: la falta de datos etiquetados para ciertos lenguajes o dominios técnicos puede reducir la efectividad de estos modelos [19].

En Visión por Computadora, los modelos de redes neuronales convolucionales (CNN) han transformado la capacidad de la inteligencia artificial para reconocer patrones visuales, Sin embargo, la generalización de estos sistemas sigue siendo limitada en escenarios con condiciones de iluminación adversas, ángulos no convencionales o datos incompletos. Además, los algoritmos actuales suelen ser vulnerables a perturbaciones adversarias, lo que plantea dudas sobre su robustez en entornos críticos como la medicina o la seguridad [20].

2. Desafíos de Escalabilidad en el Análisis de Big Data

El análisis de big data plantea cuestiones teóricas significativas relacionadas con la escalabilidad de los modelos de IA. A medida que los volúmenes de datos crecen exponencialmente, los enfoques tradicionales de procesamiento y almacenamiento son insuficientes. La Inteligencia Artificial debe adaptarse para manejar la velocidad, variedad y volumen característicos de big data [21].

Un desafío importante radica en la eficiencia computacional. Algoritmos como las redes neuronales profundas requieren recursos computacionales masivos, tanto en términos de procesamiento como de energía. Esto a llevado a investigaciones en optimización algorítmica, como la implementación de técnicas de compresión de modelos y el uso de hardware especializado, como unidades de procesamiento tensorial (TPU) [22].

Otro aspecto crítico es la integración de datos heterogéneos provenientes de múltiples fuentes. Teóricamente, esto plantea preguntas sobre la interoperabilidad y la consistencia de los datos, ya que los métodos actuales a menudo no son lo suficientemente robustos para combinar información estructurada y no estructurada de manera coherente [23].

3. Implicaciones de la IA para la Evolución del Conocimiento Humano

La Inteligencia Artificial ha pasado de ser una herramienta analítica a un actor en la generación de conocimiento científico. Este fenómeno plantea preguntas fundamentales sobre el papel de los humanos en la creación de conocimiento. Por ejemplo, los sistemas de inteligencia artificial se están utilizando para proponer ideas en áreas como la biología y la física, ayudando a que los descubrimientos científicos se realicen de manera más rápidamente.

Sin embargo, esta capacidad también introduce dilemas teóricos sobre la autonomía y la creatividad de la Inteligencia Artificial. Si los sistemas de IA pueden generar conocimiento que trasciende la comprensión humana, ¿cómo se valida y utiliza dicho conocimiento? Además, se plantea la cuestión de si la Inteligencia artificial puede ser considerada un “agente epistémico”, capaz de contribuir al corpus del conocimiento humano de manera autónoma [10].

Por último, la creciente dependencia de la IA en la investigación científica podría conducir a una pérdida de habilidades cognitivas humanas en ciertas áreas. Esto resalta la necesidad de equilibrar la automatización con el desarrollo de habilidades críticas, asegurando que los humanos sigan desempeñando un papel activo en la interpretación y aplicación del conocimiento generado [24].

Resultados y discusión

Perspectivas Futuras

El impacto de la IA en la ciencia de datos sigue siendo un área de transformación continua que influye tanto en los paradigmas teóricos como en las aplicaciones prácticas. Este apartado detalla las principales líneas de evolución que podrían consolidarse en el futuro: la redefinición de la ciencia de datos como disciplina, la integración de la IA con tecnologías emergentes, y la evolución ética y normativa en el desarrollo de la IA.

Impacto de la IA en la evolución de la ciencia de datos como disciplina

La interacción entre la IA y la ciencia de datos no solo ha ampliado las capacidades técnicas, sino que también está redefiniendo el marco conceptual de esta disciplina. En particular, se observa un desplazamiento hacia paradigmas más integrados y automatizados. Por ejemplo, el uso de arquitecturas avanzadas como redes neuronales convolucionales y transformadores permite abordar problemas de alta dimensionalidad y relaciones no lineales con mayor precisión [7], [25]

En el ámbito teórico, la ciencia de datos está pasando de ser un campo orientado al análisis descriptivo a convertirse en una disciplina predictiva y prescriptiva, donde la IA actúa no solo como una herramienta, sino como un motor de descubrimiento. Esto genera preguntas sobre la redefinición del papel científico de datos, que debe adaptarse para operar en entornos altamente automatizados. Además, esta evolución plantea desafíos relacionados con la transparencia y reproducibilidad en entornos donde los modelos son cada vez más complejos [21].

Integración de la IA con Tecnologías Emergentes

El futuro de la IA en la Ciencia de Datos estará marcado por su convergencia con tecnologías emergentes como blockchain y computación cuántica. Estas integraciones tienen el potencial de abordar limitaciones actuales y

abrir nuevas posibilidades teóricas y prácticas.

Blockchain:

El uso de blockchain en Ciencia de Datos puede transformar cómo se almacena, procesa y comparte la información, especialmente en términos de seguridad y transparencia. En un marco teórico, blockchain puede ser visto como un facilitador de la descentralización de datos, permitiendo que los modelos de IA operen en entornos más seguros y auditables [26]. Esto es particularmente relevante en sectores regulados, como la salud y las finanzas, donde la trazabilidad y la integridad de los datos son esenciales.

Computación Cuántica:

La computación cuántica promete resolver problemas de optimización y modelado que son intratables para la computación clásica. En el contexto de la Ciencia de Datos, esto podría permitir análisis más eficientes y precisos en tiempo real, especialmente en aplicaciones como predicción del mercado, análisis de redes neuronales y simulaciones complejas [27]. Sin embargo, el desafío teórico reside en desarrollar algoritmos híbridos que combinen los enfoques cuánticos con los tradicionales, asegurando que los avances tecnológicos sean accesibles y aplicables en entornos prácticos [28].

Evolución Ética y Normativa

La rápida adopción de la IA plantea desafíos éticos y normativos que deben ser abordados de manera sistemática. En el futuro, la gobernanza de la IA será esencial para garantizar que estas tecnologías sean desarrolladas y utilizadas de manera justa, transparente y responsable.

Marco Ético:

Las cuestiones éticas, como el sesgo algorítmico, la discriminación y la privacidad, han sido ampliamente debatidas en la literatura [14], [15]. Sin embargo, la IA del futuro deberá enfrentarse a dilemas aún más complejos, como la responsabilidad de las decisiones autónomas y la equidad en sistemas de IA altamente automatizados. Por ejemplo, el desarrollo de principios éticos universales, como los propuestos por la Unión Europea en su reglamento sobre IA, será crucial para mitigar riesgos [29].

Normativa y Gobernanza:

El desarrollo de marcos regulatorios específicos para la IA será necesario para equilibrar la innovación y la protección de los derechos humanos. Esto incluye establecer estándares para la explicabilidad de los modelos, la evaluación de riesgos y la supervisión continua de los sistemas implementados [10]. Las políticas internacionales, como la Declaración de Ética en la Inteligencia Artificial de UNESCO (2021), ofrecen una base inicial, pero requieren una adaptación constante a medida que las tecnologías evolucionan.

En un nivel teórico, la investigación futura deberá explorar cómo estas regulaciones pueden implementarse

de manera efectiva sin sofocar la innovación tecnológica. Esto podría incluir el desarrollo de herramientas de auditoría algorítmica y enfoques para evaluar el impacto social de los modelos de IA [24].

Conclusiones

En conclusión, La convergencia entre la Inteligencia Artificial y la Ciencia de Datos ha establecido un marco transformador en el análisis de datos y la generación de conocimiento, superando las barreras tradicionales en la interpretación de datos complejos. Este artículo ha identificado múltiples interacciones clave, como la automatización del ciclo de vida analítico, la implementación de enfoques éticos y el avance de modelos explicables.

No obstante, persisten desafíos significativos. La necesidad de superar limitaciones en el procesamiento de datos no estructurados y garantizar la escalabilidad de los modelos en un contexto de big data representa áreas críticas para la investigación futura. Asimismo, la creciente dependencia de la IA en la toma de decisiones automatizadas subraya la urgencia de desarrollar regulaciones que equilibren la innovación tecnológica y la protección de los derechos humanos.

Finalmente, el análisis resalta que el impacto de la IA no solo recae en su capacidad técnica, sino también en su influencia sobre la evolución del conocimiento humano, posicionándose como un motor central para descubrimientos interdisciplinarios y aplicaciones prácticas. En este sentido, se requiere un esfuerzo conjunto para integrar enfoques teóricos robustos con desarrollos tecnológicos sostenibles y éticos, consolidando la IA como un catalizador de progreso en la Ciencia de Datos.

Contribución de Autoría

Anthony JeanPaul Reyes Risco: Conceptualización, Investigación, Metodología, Validación, Análisis formal, Redacción - borrador original, Escritura, revisión y edición. **Juan Aurelio De La Cruz Gamarra:** Conceptualización, Investigación, Metodología, Análisis formal, Escritura, revisión y edición. **Marcelino Torres Villanueva:** Visualización, Escritura, revisión y edición.

Referencias

- [1] F. Provost and T. Fawcett, *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media, 2013. [Online]. Available: https://www.academia.edu/38731456/Data_Science_for_Business
- [2] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pear-

- son, 2021. [Online]. Available: https://api.pageplace.de/preview/DT0400.9781292401171_A41586057/preview-9781292401171_A41586057.pdf
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, 2016. [Online]. Available: [http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yoshua%20Bengio,%20Aaron%20Courville%20\(z-lib.org\).pdf](http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yoshua%20Bengio,%20Aaron%20Courville%20(z-lib.org).pdf)
- [4] P. Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, 2015.
- [5] B. Marr, *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Wiley, 2016. [Online]. Available: https://www.academia.edu/40809251/Bernard_marrbig_data_in_practice_how
- [6] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2011. [Online]. Available: http://sves.org.in/ecap/Resources/_53.pdf
- [7] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, pp. 436–444, 2015. [Online]. Available: https://www.researchgate.net/profile/Y-Bengio/publication/277411157_Deep_Learning/links/55e0cdf908ae2fac471ccf0f/Deep-Learning.pdf
- [8] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should i trust you? explaining the predictions of any classifier,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144. [Online]. Available: <https://dl.acm.org/doi/epdf/10.1145/2939672.2939778>
- [9] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012. [Online]. Available: <https://github.com/kerasking/book-1/blob/master/ML%20Machine%20Learning-A%20Probabilistic%20Perspective.pdf>
- [10] L. Floridi, “Big data and their epistemological challenge,” *Philosophy & Technology*, pp. 435–437, 2012. [Online]. Available: <https://link.springer.com/article/10.1007/s13347-012-0093-4>
- [11] X. He, K. Zhao, and X. Chu, “Automl: A survey of the state-of-the-art,” *Knowledge-Based Systems*, vol. 212, p. 106622, 2021. [Online]. Available: https://www.researchgate.net/publication/334963534_AutoML_A_Survey_of_the_State-of-the-Art
- [12] M.-A. Zöllner and M. F. Huber, “Benchmark and survey of automated machine learning frameworks,” *Journal of Artificial Intelligence Research*, vol. 70, pp. 409–472, 2021. [Online]. Available: <https://www.jair.org/index.php/jair/article/download/11854/26651/25924>

- [13] F. Hutter, L. Kotthoff, and J. Vanschoren, *Automated Machine Learning: Methods, Systems, Challenges*. Springer, 2019. [Online]. Available: https://www.automl.org/wp-content/uploads/2019/05/AutoML_Book.pdf
- [14] C. O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown, 2016. [Online]. Available: https://edisciplinas.usp.br/pluginfile.php/4605464/mod_resource/content/1/%28FFLCH%29%20LIVRO%20Weapons%20of%20Math%20Destruction%20-%20Cathy%20O%27Neal.pdf
- [15] S. Barocas, M. Hardt, and A. Narayanan, *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press, 2023. [Online]. Available: <https://fairmlbook.org/pdf/fairmlbook.pdf>
- [16] Z. C. Lipton, “The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery,” *Communications of the ACM*, vol. 61, no. 10, pp. 36–43, 2018. [Online]. Available: <https://dl.acm.org/doi/epdf/10.1145/3281635>
- [17] F. Doshi-Velez and B. Kim, “Towards a rigorous science of interpretable machine learning,” *arXiv preprint arXiv:1702.08608*, 2017. [Online]. Available: <https://arxiv.org/pdf/1702.08608>
- [18] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, and D. Amodei, “Language models are few-shot learners,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020. [Online]. Available: <https://arxiv.org/pdf/2005.14165>
- [19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “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*, 2019, pp. 4171–4186. [Online]. Available: https://eva.fing.edu.uy/pluginfile.php/524749/mod_folder/content/0/BERT%20Pre-training%20of%20Deep%20Bidirectional%20Transformers%20for%20Language%20Understanding.pdf
- [20] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, “Intriguing properties of neural networks,” in *Proceedings of the International Conference on Learning Representations*, 2014. [Online]. Available: <https://arxiv.org/pdf/1312.6199>
- [21] L. Zhou, S. Pan, J. Wang, and A. V. Vasilakos, “Machine learning on big data: Opportunities and challenges,” *Elsevier*, vol. 237, pp. 350–361, 2017. [Online]. Available: <https://pdf.sciencedirectassets.com/271597/1-s2.0-S0925231217X00106/1-s2.0-S0925231217300577/am.pdf>

- [22] J. Dean, D. Patterson, and Y. Cliff, “A new golden age in computer architecture: Empowering the machine-learning revolution,” *IEEE Micro*, vol. 38, no. 2, pp. 21–29, 2018. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/8259424/>
- [23] M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014. [Online]. Available: <https://www.cs.unibo.it/~montesi/CBD/Articoli/SurveyBigData.pdf>
- [24] I. Rahwan, M. Cebrian, N. Obradovich, J. Bongard, J.-F. Bonnefon, C. Breazeal, J. W. Crandall, N. A. Christakis, I. D. Couzin, M. O. Jackson, N. R. Jennings, E. Kamar, and M. Wellman, “Machine behaviour,” *Nature*, vol. 568, no. 7753, pp. 477–486, 2019. [Online]. Available: <https://www.nature.com/articles/s41586-019-1138-y>
- [25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
- [26] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008. [Online]. Available: <https://bitcoin.org/bitcoin.pdf>
- [27] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, and J. M. Martinis, “Quantum supremacy using a programmable superconducting processor,” *Nature*, vol. 574, no. 7779, pp. 505–510, 2019. [Online]. Available: <https://www.nature.com/articles/s41586-019-1666-5>
- [28] J. Preskill, “Quantum computing in the nisq era and beyond,” *Quantum*, vol. 2, p. 79, 2018. [Online]. Available: <https://quantum-journal.org/papers/q-2018-08-06-79/pdf/>
- [29] European Commission, “Proposal for a regulation laying down harmonized rules on artificial intelligence,” Tech. Rep. COM/2021/206 final, 2021. [Online]. Available: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF