

- **Glosario de Términos**

Este glosario tiene como objetivo proporcionar definiciones claras y concisas de los términos clave utilizados a lo largo del libro.

- **Algoritmo:**

- Un conjunto finito y ordenado de instrucciones o reglas bien definidas y no ambiguas que se siguen paso a paso para resolver un problema específico o realizar una tarea. Los algoritmos son la base de la computación.
 - - Ejemplo:* Una receta de cocina es un algoritmo para preparar un plato.

- **Inteligencia Artificial (IA):**

- Un campo de la informática que se enfoca en la creación de sistemas y programas capaces de simular procesos de inteligencia humana. Esto incluye la capacidad de razonar, aprender, resolver problemas, percibir el entorno, comprender el lenguaje y tomar decisiones.
 - - Ejemplo:* Los asistentes virtuales como Siri o Alexa utilizan IA.

- **Aprendizaje Automático (Machine Learning, ML):**

- Una rama de la IA que se centra en el desarrollo de algoritmos que permiten a las computadoras aprender de los datos sin ser explícitamente programadas. Los algoritmos de ML mejoran su rendimiento a través de la experiencia.
 - - Ejemplo:* Los filtros de spam de correo electrónico utilizan ML para identificar correos no deseados.

- **Aprendizaje Supervisado:**

- Un tipo de aprendizaje automático donde los algoritmos aprenden a partir de datos etiquetados. Los datos etiquetados incluyen tanto la entrada como la salida deseada, permitiendo al algoritmo aprender a mapear entradas a salidas.
 - - Ejemplo:* Entrenar un algoritmo para clasificar imágenes de gatos y perros, donde cada imagen está etiquetada como "gato" o "perro".

- **Aprendizaje No Supervisado:**

- Un tipo de aprendizaje automático donde los algoritmos aprenden a partir de datos no etiquetados. El algoritmo busca patrones, estructuras o agrupaciones ocultas en los datos sin la guía de una salida predefinida.
 - - Ejemplo:* Agrupar clientes en diferentes segmentos de mercado basándose en su comportamiento de compra.

- **Aprendizaje por Refuerzo:**

- Un tipo de aprendizaje automático donde un agente aprende a tomar decisiones interactuando con un entorno. El agente recibe recompensas o penalizaciones por sus acciones, aprendiendo así a maximizar las recompensas a lo largo del tiempo.
 -

- Ejemplo:* Entrenar un robot para navegar por un laberinto.

- **Fraude:**

- Un acto intencional de engaño o tergiversación diseñado para obtener una ganancia financiera o de otro tipo a expensas de otra persona o entidad. Implica la violación de la confianza.
 - - Ejemplo:* Fraude con tarjetas de crédito, fraude fiscal.

- **Corrupción:**

- El abuso de poder público o privado para obtener un beneficio personal o de grupo. Implica el uso indebido de la autoridad para fines deshonestos.
 - - Ejemplo:* Soborno de funcionarios públicos, malversación de fondos.

- **Sesgo Algorítmico:**

- Una tendencia sistemática en un algoritmo que causa que produzca resultados injustos, discriminatorios o desproporcionados hacia ciertos grupos o individuos. El sesgo algorítmico puede surgir de los datos de entrenamiento sesgados o del diseño del algoritmo.
 - - Ejemplo:* Un algoritmo de contratación que favorece a los hombres sobre las mujeres debido a datos históricos de contratación sesgados.

- **Datos Abiertos (Open Data):**

- Datos, especialmente información gubernamental, que están disponibles para que cualquiera los acceda, utilice y comparta, sin restricciones de derechos de autor, patentes u otros mecanismos de control.
 - - Ejemplo:* Datos sobre presupuestos gubernamentales, estadísticas de criminalidad.

- **Procesamiento del Lenguaje Natural (NLP):**

- Una rama de la IA que se ocupa de la interacción entre las computadoras y el lenguaje humano. El NLP permite a las computadoras entender, interpretar y generar lenguaje humano en forma de texto o voz.
 - - Ejemplo:* Traducción automática, análisis de sentimientos en redes sociales.

- **Análisis de Redes:**

- Un conjunto de técnicas y métodos para estudiar las relaciones y conexiones entre entidades (personas, organizaciones, conceptos, etc.). El análisis de redes ayuda a comprender la estructura y dinámica de sistemas complejos.
 - - Ejemplo:* Analizar redes de colusión en la contratación pública.

- **Detección de Anomalías:**

- El proceso de identificar puntos de datos, eventos o observaciones que se desvían significativamente del comportamiento normal o esperado. Las anomalías pueden indicar eventos inusuales, errores o fraude.

- - Ejemplo:* Detectar transacciones bancarias fraudulentas que se desvían del patrón de gasto habitual de un cliente.

- **Explicabilidad de la IA (XAI):**

- Un campo de la IA que se centra en el desarrollo de técnicas y métodos para hacer que los modelos y decisiones de IA sean más transparentes, comprensibles e interpretables para los humanos. La XAI busca aumentar la confianza y la responsabilidad en los sistemas de IA.
 - - Ejemplo:* Proporcionar razones claras y comprensibles para la decisión de un modelo de IA de denegar un préstamo.

- **Bibliotecas usadas en análisis de datos y recursos adicionales**

- Bibliotecas:

- ✓ Faker: <https://pypi.org/project/Faker/>
- ✓ Pandas: <https://pandas.pydata.org/>
- ✓ Random : <https://docs.python.org/3/library/random.html>
- ✓ Numpy : <https://numpy.org/>
- ✓ Datetime: <https://docs.python.org/es/3.13/library/datetime.html>
- ✓ Matplotlib : <https://matplotlib.org/>
- ✓ Seaborn: <https://seaborn.pydata.org/>
- ✓ Collections: <https://docs.python.org/3/library/collections.html>
- ✓ Scikit-learn: <https://scikit-learn.org/>
- ✓ Warnings: <https://docs.python.org/es/3/library/warnings.html>
- ✓ Nltk: <https://www.nltk.org/>
- ✓ Re: <https://docs.python.org/es/3.13/library/re.html>
- ✓ Imblearn: <https://imbalanced-learn.org/stable/>

- **Recursos Adicionales**

- **Plataformas de Aprendizaje en Línea:**

- **Coursera** : Ofrece cursos y especializaciones sobre inteligencia artificial, aprendizaje automático, ciencia de datos y ética de la IA, impartidos por universidades y organizaciones de todo el mundo. Se recomienda buscar cursos específicos como "Machine Learning" de Andrew Ng o "AI for Everyone". <https://www.coursera.org/>
- **edX** : Similar a Coursera, ofrece una amplia gama de cursos sobre IA, ciencia de datos y temas relacionados. Se recomienda buscar cursos de universidades como el MIT o Harvard. : <https://www.edx.org/>
- **Udacity** : Ofrece "Nanodegrees" y cursos prácticos sobre ciencia de datos, aprendizaje automático y desarrollo de IA. www.udacity.com

- **Organizaciones y Agencias:**

- Transparency International (www.transparency.org)
- Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC) (www.unodc.org)

- Financial Crimes Enforcement Network (FinCEN) (www.fincen.gov)
- Open Government Partnership (OGP) (www.opengovpartnership.org)
- Association of Certified Fraud Examiners (ACFE) (www.acfe.com)

● **Publicaciones y Sitios Web Especializados:**

- Revistas académicas sobre IA, ciencia de datos, seguridad informática, gobernanza y derecho.
- Blogs y sitios web de expertos en IA, ética de la tecnología y lucha contra la corrupción.
- Repositorios de investigación y bases de datos académicas (por ejemplo, arXiv: <https://arxiv.org/>, Google Scholar: <https://scholar.google.com/>)

● **Herramientas y Software:**

- Listado de software y plataformas de análisis de datos e IA relevantes para la detección de fraude y corrupción (por ejemplo, Python, R, herramientas de visualización de datos).

● **Lenguajes de Programación:**

- **Python:** Consulta la documentación oficial de Python (python.org) para aprender sobre el lenguaje y sus bibliotecas relevantes para el análisis de datos (NumPy, Pandas) y aprendizaje automático (scikit-learn, TensorFlow, Keras). : <https://www.python.org/>
- **R:** Revisa la documentación oficial de R (r-project.org) para aprender sobre el lenguaje y sus paquetes para el análisis estadístico y el aprendizaje automático. <https://cran.rstudio.com/>

● **Plataformas de Análisis de Datos:**

- **Jupyter Notebook:** Consulta la documentación oficial para aprender a utilizar esta herramienta interactiva para el análisis de datos y la programación en Python y R. : <https://www.anaconda.com/download>
- **Google Colab:** Revisa la documentación para aprender a utilizar esta plataforma gratuita basada en la nube para ejecutar código Python. <https://colab.google/>

● **Bibliotecas de Aprendizaje Automático:**

- **scikit-learn:** Consulta la documentación oficial (scikit-learn.org) para aprender sobre las diversas técnicas de aprendizaje automático implementadas en esta librería de Python.: <https://scikit-learn.org/stable/>
- **TensorFlow:** Revisa la documentación oficial (tensorflow.org) para aprender sobre esta biblioteca de aprendizaje automático de código abierto : <https://www.tensorflow.org/?hl=es-419>
- **Keras:** Consulta la documentación oficial (keras.io) para aprender sobre esta API de alto nivel para redes neuronales.: <https://keras.io/>

Acceso del lector al material del libro:¹

- Para un mejor manejo del Drive, Colaboratory y archivos se sugiere al lector poseer una cuenta de gmail gratuita.
- Para el lector que no tenga familiaridad con los repositorios de código Github, el lector accediendo al drive del autor, apoyando el mouse en el archivo que quiere abrir, despliega con boton derecho su apertura.

- Repositorio del autor:

https://github.com/Viny2030/Libro_Algoritmos_contra_fraude_corrupcion

- Acceso al drive del autor:

https://drive.google.com/drive/folders/1Ks8Xtsly0okZjhOTfB9m622EJKGkoLix?usp=drive_link

- Acceso al colab de los conjuntos de datos:

https://github.com/Viny2030/Libro_Algoritmos_contra_fraude_corrupcion/blob/main/dataset.xlsx

<https://drive.google.com/file/d/19bHYYJqBv4Gvxx3DgklYbcRtP3imOes0/view?usp=sharing>

- Acceso al colab del código de los algoritmos:

<https://drive.google.com/file/d/1T4m2u7-JWM7CsaGBc8gywSTwhYfrcNcd/view?usp=sharing>

¹ El lector puede abrir los archivos de los conjuntos de datos y de los algoritmos directamente del drive.