

Ejercicio 1: ¿Qué es “Machine Learning”?

Definiciones de Machine Learning:

Definición 1: Machine Learning es una rama de la inteligencia artificial que permite a las computadoras aprender de los datos y mejorar su rendimiento con el tiempo sin ser programadas explícitamente.

Definición 2: Machine Learning es el estudio de algoritmos y modelos estadísticos que las computadoras utilizan para realizar tareas específicas sin utilizar instrucciones explícitas, dependiendo en cambio de patrones e inferencias.

Definición 3: Machine Learning es un enfoque que permite a los sistemas informáticos aprender y hacer predicciones o decisiones basadas en datos.

Comparaciones:

Con Inteligencia Artificial (IA): Machine Learning es una subcategoría de la IA. Ambos buscan simular la inteligencia humana, pero Machine Learning se centra específicamente en el aprendizaje a partir de datos.

Con Análisis Estadístico: Ambos utilizan métodos para analizar datos, pero el análisis estadístico se basa en la inferencia a partir de muestras, mientras que Machine Learning se centra en la predicción a partir de datos completos.

Con Data Mining: El Data Mining se enfoca en la exploración y descubrimiento de patrones en grandes conjuntos de datos, mientras que Machine Learning utiliza esos patrones para hacer predicciones o tomar decisiones.

Aplicaciones: Machine Learning se aplica en diversas áreas, como reconocimiento de voz, visión por computadora, procesamiento de lenguaje natural y sistemas de recomendación.

## **Ejercicio 2:** Herramientas y plataformas para Machine Learning

Herramientas destacadas:

### **RapidMiner**

Características:

- Tratamiento de datos: Herramientas para limpieza y transformación de datos.
- Algoritmos: Incluye diversos algoritmos de aprendizaje automático.
- Prueba y validación: Métodos para evaluar modelos.
- Integración: Funcionalidades para integrarse con servicios en la nube.

### **Weka**

Características:

- Tratamiento de datos: Herramientas para preprocesamiento y visualización.
- Algoritmos: Amplia colección de algoritmos de machine learning.
- Prueba y validación: Métodos para validar modelos.
- Integración: Capacidad de importar y exportar datos en varios formatos.

## **Ejercicio 3:** Procesos de “Data Science”

Proceso CRISP-DM:

CRISP-DM (Cross-Industry Standard Process for Data Mining) es un modelo de proceso para el desarrollo de proyectos de Data Science. Consiste en seis fases:

- Comprensión del negocio: Entender el objetivo y el contexto del negocio.
- Comprensión de los datos: Recolectar y analizar los datos iniciales.
- Preparación de los datos: Limpiar y transformar los datos para el modelado.
- Modelado: Seleccionar y aplicar técnicas de modelado.
- Evaluación: Evaluar el modelo para asegurar que cumple los objetivos.
- Implementación: Implementar el modelo en el entorno de producción.

Otros procesos similares:

KDD (Knowledge Discovery in Databases): Se centra en el descubrimiento de conocimiento a partir de grandes bases de datos y es más amplio que CRISP-DM.

SEMMA (Sample, Explore, Modify, Model, Assess): Un enfoque que se centra en la exploración y modelado de datos.

Similitudes y diferencias: Todos estos procesos comparten el objetivo de transformar datos en conocimiento, pero pueden diferir en su enfoque y pasos específicos.

**Ejercicio 4:** Comparativa de herramientas

Característica	RapidMiner	Weka
Algoritmos soportados	Variados (clustering, clasificación)	Variados (clustering, clasificación)
Parámetros	Configuración de algoritmos	Configuración de algoritmos
Integración	Con servicios en la nube	Importación/exportación de formatos