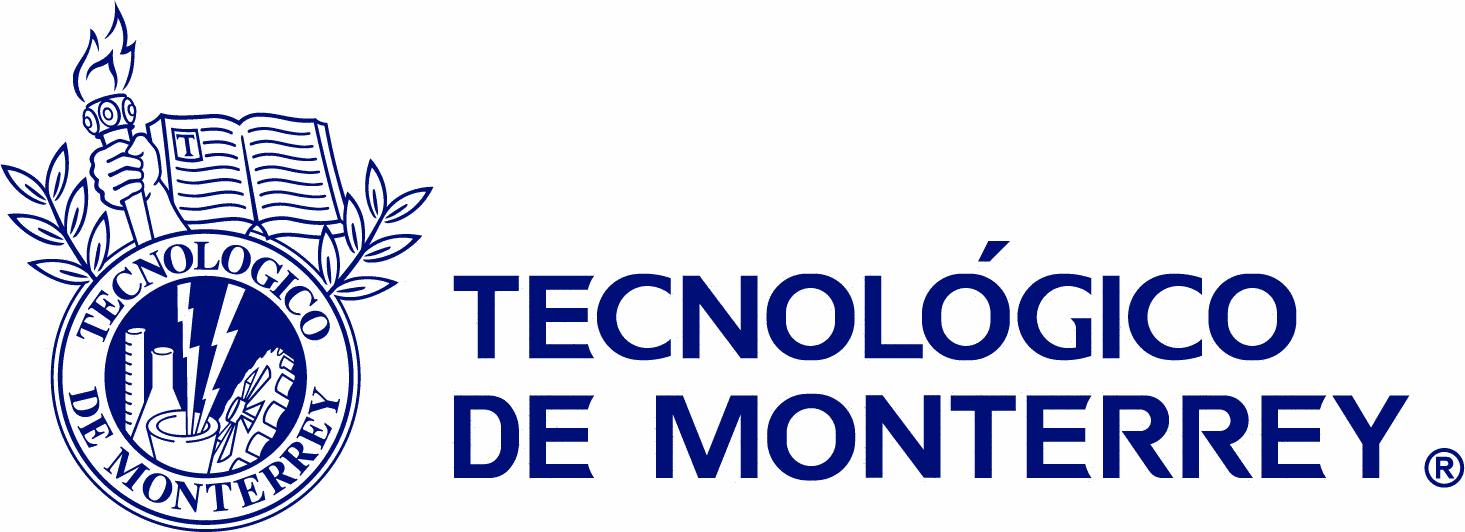
**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY**



**Maestría en Inteligencia Artificial con Especialidad en Ciencia de Datos**

**Materia:**

Modelos de Machine Learning (MLOps)

**Profesor Titular:**

Carlos L. Mejia



**PROYECTO INTEGRADOR**

**Alumno:**

A01688818 - Edgar Filiberto Nava Barrón

Agosto de 2023.

**INTRODUCCIÓN**

Los Modelos de Machine Learning son algoritmos que permiten al software aprender de los datos y mejorar su rendimiento en tareas específicas sin intervención humana directa. A través de la capacitación en patrones y relaciones ocultas en los datos, estos modelos tienen la capacidad de hacer predicciones, clasificar información, generar recomendaciones y mucho más. Su aplicación se extiende a diversas áreas, desde la detección de fraudes, hasta diagnósticos médicos.

Este proceso de aprendizaje se basa en el concepto de que el modelo pueda detectar patrones en los datos mucho más rápido y eficientemente que los métodos tradicionales de programación. En lugar de codificar reglas específicas para cada situación, los Modelos de Machine Learning se entrenan con ejemplos y, a medida que se exponen a más datos, pueden generalizar y adaptarse a situaciones nuevas. Es como enseñar a una máquina a reconocer rostros humanos al mostrarle una serie de fotografías etiquetadas en lugar de programar manualmente todas las características faciales.

Uno de los pilares fundamentales del Machine Learning es la diversidad de enfoques que ofrece para resolver problemas. Los Modelos de Machine Learning se pueden agrupar en diversas categorías, como el Aprendizaje Supervisado, donde los modelos se entrenan en conjuntos de datos etiquetados para predecir etiquetas futuras; el Aprendizaje No Supervisado, que explora patrones en datos no etiquetados para encontrar estructuras ocultas; y el Aprendizaje por Reforzamiento, donde los modelos aprenden a través de la interacción con un entorno y reciben recompensas por tomar decisiones óptimas.

En esta exploración de los Modelos de Machine Learning, nos sumergiremos en los conceptos esenciales que los respaldan, desde entender cómo se construyen y entrenan estos modelos hasta explorar las consideraciones éticas y los desafíos actuales del campo, nuestro objetivo es proporcionarte una base sólida para comprender y aprovechar el potencial de esta tecnología revolucionaria.

En este proyecto integrador enfocado en MLOps, se investigarán los conceptos clave de los marcos de ML y su aplicación en un enfoque integral. Se mostrarán los conceptos básicos y las herramientas fundamentales para desarrollar software en el campo de MLOps, cubriendo desde la configuración el entorno hasta las mejores prácticas para crear y desplegar modelos de ML.

**Objetivos (1)**

* Elegir el dataset de forma individual.
* Entender el problema del dataset
* Preparar el notebook con la línea base (modelo más simple que resuelva el problema)

**Datasets**

* Comenzar a explorar los datasets que han sido asignados.

**Entregables**

* Definición de la línea base y alcance del proyecto
* Notebook

**Recapitulación de proyecto - Parte 1**

Esta recapitulación consiste en analizar el contenido aprendido durante el módulo 2 y 3, que será usado para integrar y evolucionar el proyecto.

Durante el módulo 2, el alumno se enfocará en aprender los conceptos y comenzar a aplicarlos en los ejemplos que se le van a otorgar.

A lo largo del módulo 3, el alumno integrará los conocimientos adquiridos desde la primera sesión en su proyecto, se recomienda que la integración sea de forma inmediata, ya que las sesiones son consecutivas.

**Objetivos (2)**

* Incluir ambientes virtuales
* Uso continuo de GitHub
* Pruebas unitarias
* Pre-commits
* Refactorización
* Linting y formatting
* Estructura de directorios
* POO (Clases, métodos, transformadores, pipelines)
* GitHub Actions
* REST API - FastAPI
* Logging

**Entregables**

* Repositorio con el proyecto y con la integración de todos los objetivos.

Exploración de Data Set.

<https://github.com/EdgarF9799/Modulo3/tree/main/01_Introducction>

Trabajando con modelo de predicción, Pre-commits, Refactorización, Linting y formatting, Estructura de directorios, POO (Clases, métodos, transformadores, pipelines).

<https://github.com/EdgarF9799/Modulo3/tree/main/02_ProcesandoDataSet>