

Proyecto Final - Introducción a la Inteligencia Artificial.

Descripción

El objetivo de este proyecto es que se apliquen los conocimientos adquiridos en el curso de Introducción a la Inteligencia Artificial, para resolver un problema real utilizando un dataset de tu elección.

Instrucciones.

Este proyecto se realizará de forma individual o en parejas y tendrá una duración aproximada de dos semanas.

Deberá seleccionar un tema que te resulte entretenido o significativo, buscar un dataset adecuado, definir una pregunta o problema a resolver, y diseñar un plan de acción para abordarlo utilizando alguno de los modelos revisados en clase.

Todo el proyecto debe ser documentado y presentado a través de un repositorio de GitHub que contenga el código, los resultados y un archivo `README.md` explicativo. Este último reemplazará a lo que sería un informe o reporte, por lo que sólo se evaluará lo respaldado en dicho repositorio. Se evaluará el aporte de cada miembro del equipo por separado, viendo sus commits en el repositorio. Si un miembro del equipo no realiza aportes, esta persona será evaluada con la nota mínima.

Plazos

26 de Junio: Inicio.

- Selección del dataset y problemática.
- Determinación del plan de trabajo

03 de Julio: Checkpoint

- Ejecución del plan de trabajo
- Documentación.

10 de Julio: Entrega.

Entregables

A más tardar el 4 de Julio, se debe adjuntar el link de un repositorio de Github, además de los nombres de los participantes del equipo. Los diferentes avances se irán evaluando según las fechas entregadas en la pauta (leer más abajo).

Evaluación

Fecha de evaluación	Criterio	Descripción	Puntaje máximo
04 de Julio	Definición del problema	Claridad y relevancia del problema planteado.	5 pts
	Plan de acción	Descripción del dataset, modelo(s) seleccionado(s), y estrategia de evaluación claramente explicados	5 pts
	Justificación del modelo.	Elección bien argumentada del modelo en relación al problema. Consideración de ventajas, limitaciones y pertinencia.	5 pts
11 de Julio	Implementación	Código limpio, reproducible y funcional. Uso correcto de librerías y buenas prácticas de programación. Debe considerar: EDA y feature engineering Feature selection. Reducción de dimensionalidad (según corresponda) Entrenamiento. Control de overfitting. Testeo. Visualización de resultados. Métricas.	20 pts
	Resultados	Presentación clara de resultados, interpretación adecuada de métricas, gráficas, discusión crítica de los hallazgos.	10 pts
	Documentación	Repositorio organizado, `README.md` completo y claro, estructura de carpetas adecuada. Dependencias correctamente reportadas en un archivo .yaml.	5 pts
	Conclusiones	Claridad en los hallazgos presentada en el README.md.	10 pts
Total			60 pts

Sobre la implementación

El proyecto debe ser implementado en un notebook .ipynb en el que se abarque desde la importación de los datos, hasta el reporte de métricas y gráficos. El notebook debe estar completamente ejecutable para obtener puntaje en el apartado de implementación.

Es requerimiento realizar un análisis de datos exploratorio y feature engineering.

Podrá utilizar al menos uno de los siguientes métodos vistos en clase:

- Regresión lineal
- Regresión logística
- Naive Bayes
- K-Nearest Neighbors (KNN)
- Decision Tree
- Random Forest
- Transfer Learning (Deep Learning)
 - Imágenes.
 - Texto.

El repositorio deberá contener:

- Un notebook .ipynb con código limpio, funcional y comentado.
- Un archivo `README.md` que explique:
 - Título del proyecto.
 - Descripción del problema.
 - Breve descripción del dataset y fuente.
 - Justificación del modelo (s) seleccionado.
 - Metodología aplicada (paso a paso).
 - Resultados obtenidos.
 - Conclusiones.
- Las dependencias de su proyecto en formato .yml.

Recomendaciones

- Explorar datasets en Kaggle. Busque un tema que lo motive o divierta: deportes, cine, salud, música, juegos, redes sociales, etc.
- Si tiene dudas sobre la pertinencia de su dataset o proyecto, consultar al ayudante o docente.
- En caso de requerir GPU y que ningún miembro del equipo posea, utilizar los notebooks de Kaggle o Collab.
- Documenta bien cada paso en GitHub para que el repositorio sea claro y fácil de entender.