

Instancia evaluativa 1 - Machine Learning

Universidad de Medellín

2024-2

El trabajo que tienes ante ti representa el primer hito evaluativo dentro del marco del curso "Machine Learning", impartido durante el período académico 2024-2 en la Universidad de Medellín. En esta etapa inicial, se asigna un peso significativo del 12.5% sobre la evaluación total del curso.

Objetivo:

Desarrollar un modelo de machine learning para predecir el precio de casas utilizando regresión lineal. Como parte del proceso, realizarás un análisis exploratorio de datos (EDA) y llevarás a cabo el feature engineering para mejorar la calidad del modelo.

Instrucciones:

1. Obtención de Datos:

- Utiliza el conjunto de datos proporcionado en este reto de kaggle <https://www.kaggle.com/c/home-data-for-ml-course>

2. Análisis Exploratorio de Datos (EDA):

- Describe y visualiza las características del dataset.
- Identifica y trata valores nulos o inconsistentes.
- Explora la distribución de las variables y su relación con el precio de las casas.
- Identifica y analiza posibles outliers.
- Realiza análisis de correlación para identificar relaciones entre las características y el precio.

3. Feature Engineering:

- Crea nuevas características que puedan mejorar el rendimiento del modelo. (ex. one-hot encoding)
- Realiza transformaciones necesarias (por ejemplo, escalado, normalización).
- Considera la selección o eliminación de características redundantes o poco informativas.

4. Modelado:

- Aplica regresión lineal al dataset transformado. (Puede usar regularización)
- Ajusta el modelo y evalúa su desempeño utilizando métricas como RMSE (Root Mean Square Error)
- Hacer tuning del modelo.

5. Documentación y Presentación:

- Documenta cada uno de los pasos del proceso, incluyendo las decisiones tomadas y el razonamiento detrás de ellas.
- Presenta los resultados en un notebook, incluyendo gráficos y análisis.

6. Entrega:

- Sube el notebook finalizado con tus comentarios y análisis.

Si necesitas recursos adicionales o tienes preguntas, no dudes en consultarme.