
ACTIVIDAD PREVIA

- Solicitar llenar el foro «¿Qué aprendí esta semana?» del aula virtual.
- Cuestionar sobre la colinealidad al aplicar one-hot encoding.

RESULTADO DE APRENDIZAJE

RdA de la asignatura:

- **RdA 2:** Aplicar modelos de aprendizaje automático supervisado y no supervisado, así como su validación y optimización, en la resolución de problemas tanto reales como simulados.
- **RdA 3:** Resolver problemas prácticos mediante el uso de modelos de aprendizaje automático, ajustándolos para la mejora de su rendimiento y precisión.

RdA de la clase:

- Identificar las etapas esenciales del flujo de trabajo del entrenamiento de un modelo supervisado.
- Implementar la división de un conjunto de datos en entrenamiento y prueba mediante Python.
- Entrenar un modelo de regresión lineal y evaluar su desempeño.

INTRODUCCIÓN

Pregunta inicial: ¿Cómo se aplica el aprendizaje supervisado para predecir valores continuos en un conjunto de datos usando Python?

DESARROLLO

Actividad 1: El flujo de trabajo del aprendizaje supervisado

Esta actividad presenta los elementos fundamentales del flujo de trabajo: generación de datos, separación en conjuntos, entrenamiento y evaluación. Se desarrollará mediante clase magistral apoyada con ejemplos y discusión guiada.

¿Cómo lo haremos?

- **Presentación del flujo:** Se revisarán las etapas: definición del problema, construcción del conjunto de datos, división en entrenamiento y prueba, entrenamiento del modelo, predicción y métricas.
- **Discusión guiada:** Se plantearán preguntas para reflexionar sobre la importancia de cada etapa.

Actividad 2: Implementación en Python de un modelo de regresión lineal

La actividad se desarrollará mediante trabajo práctico guiado en Python, utilizando un conjunto de datos simulado.

¿Cómo lo haremos?

- **Implementación en Python:** Los estudiantes accederán a un cuaderno de Jupyter previamente preparado.

Enlace al cuaderno: [04-Mi-primer-modelo.ipynb](#).

- **Entrenamiento y evaluación:** Se ejecutará el modelo de regresión lineal, se obtendrán predicciones y se calcularán métricas como el error cuadrático medio.
- **Experimentación:**

- Cambiar el tamaño del conjunto de entrenamiento y observar el impacto en las métricas.
- Cambiar el modelo a uno de Máquina de Soporte Vectorial (SVM) y comparar resultados (se deberá comentar algunas partes del código).

Actividad 3: Taller de implementación

Los estudiantes trabajarán para implementar un modelo de regresión lineal desde cero, siguiendo los pasos aprendidos, sobre el conjunto de datos `kingCountyHouseData`.

¿Cómo lo haremos?

- **Desarrollo del taller:** Los estudiantes aplicarán los conceptos aprendidos para dividir los datos, entrenar el modelo y evaluar su desempeño. Se utilizará la siguiente plantilla:

Enlace al cuaderno: [01-Mi-segundo-modelo-Plantilla.ipynb](#).

- **Soporte y guía:** El instructor estará disponible para resolver dudas y proporcionar retroalimentación durante la actividad.

CIERRE

Verificación de aprendizaje:

- ¿Qué etapas conforman el flujo de trabajo del aprendizaje supervisado?
- ¿Cómo se divide un conjunto de datos en entrenamiento y prueba utilizando Python?
- ¿Qué hacen los métodos `fit()`, `transform()` y `predict()` en scikit-learn?

Tarea: Completar las tareas «Repositorio de GitHub para tareas» y «Mi segundo modelo», del aula virtual.