

Titulo

Primeros pasos en machine learning: Dominando los árboles de decisión

Descripción

Taller hand-on completo, orientado a quienes nunca han trabajado en proyectos de machine learning.

El objetivo es que cada participante pueda terminar con un modelo de machine learning completamente funcional en su notebook.

Se verán tanto temas teóricos como prácticos, directamente relacionados a la creación de modelos de machine learning como al análisis y procesamiento de los datos que alimentarán esos modelos.

Conoceremos los diferentes tipos de algoritmos disponibles y su utilización.

Ahondaremos en uno de los modelos predictivos más conocidos, el árbol de decisión.

Abordaremos el tratamiento de los datos con los que nuestro modelo “aprenderá”; qué sucede con los datos faltantes y los outliers, cómo afectan a nuestro modelo y las diferentes estrategias para tratarlos.

Utilizaremos diferentes técnicas que nos permitirán ir mejorando los valores arrojados por nuestro modelo.

Prerequisitos

Saber python (no es necesario ser un experto)

Haber trabajado alguna vez en un Jupyter notebook

Temas a tratar

- Análisis exploratorio de datos
- Detección de outliers y datos faltantes, estrategias para su tratamiento
- Diferencia entre aprendizaje supervisado y no supervisado, algoritmos disponibles
- Árboles de decisión
- Estrategias de separación de datos en training/testing/validación
- Matriz de confusión, métricas para evaluación de modelos
- Optimización de hiper-parámetros, diferentes estrategias
- Comprensión de la importancia de las variables en los modelos, uso de los valores de Shapley
- Feature engineering