


Plataforma para Machine Learning “bigML”

Ejercicio práctico: Creación y entrenamiento de un árbol de clasificación haciendo uso de la plataforma bigML.

Indicaciones iniciales:

- Sigue las transparencias.
- Una vez cargado el dataset wine, toma nota del número de registros que incluye. Te será útil para compararlo después con el de entrenamiento y validación.
- Una vez seleccionado el dataset wine, editar la característica “class” y modificar el nombre por “Calidad” y establecerla como objetivo pinchando sobre el icono 
- Compara diversas características entre sí y observa la gráfica. Observa que si comparas una característica consigo misma el resultado es una línea.
- Genera los dataset de entrenamiento y validación. Asegúrate de que cargas el de entrenamiento. Observa que tiene menos registros que el inicial.
- Observa que información te ofrece cada nodo del árbol cuando pones el puntero sobre él.
- Prueba con otros dataset de los disponibles Free

Contesta a las siguientes cuestiones:

1. Pregunta: ¿Cuáles son las distintas formas de cargar o importar un dataset en la plataforma bigML?

2. Pregunta: Presenta y describe las características del dataset red wine quality?

3. Pregunta: Investiga y describe cómo se pueden eliminar valores nulos o atípicos en un dataset que tengas cargado en bigML antes de entrenarlo.

4. Pregunta: ¿De qué tipo es la variable objetivo en este ejercicio? ¿Cómo la seleccionamos y establecemos como la variable objetivo?

5. Pregunta: ¿Qué hiperparámetros se pueden ajustar al crear un árbol de decisión en BigML?

6. Pregunta: ¿Qué métricas proporciona BigML para evaluar el modelo?

7. Pregunta: ¿Qué variables influyen más en la predicción de la calidad del vino según el modelo generado? (Buscar la sección "*Field Importance*" (*Importancia de los Campos*))

8. Pregunta: ¿Cómo se puede mejorar la precisión del modelo, ajustando hiperparámetros o aplicando técnicas como la poda?

9. Pregunta: ¿Cómo se puede utilizar el modelo para predecir la calidad de nuevos vinos en BigML?

10. Pregunta: ¿Qué opciones ofrece BigML para exportar el modelo o integrarlo en una aplicación externa?

11. Pregunta: Investiga y describe el formato JSON