

Introducción a la Inteligencia Artificial – LCC – 2022

Trabajo Práctico: Aprendizaje Automatizado en Python

Utilizar el conjunto de datos de imágenes de dígitos integrado dentro de la librería Scikit-Learn ([link al dataset](#)):

1. Realizar un breve análisis del mismo indicando: cantidad de muestras, cantidad de clases, cantidad de muestras por clase, entre otras. Grafique muestras de las imágenes.
2. Árboles de decisión:
 - a. Entrenar un árbol de decisión con los parámetros de parada por defecto. ¿Qué efectos observa sobre ese árbol? ¿Cómo los explica?
 - b. Analizar cómo varían los resultados ante modificaciones de algunos de los parámetros de parada del entrenamiento. Realizar gráficas del accuracy sobre el conjunto de entrenamiento y evaluación para todos los casos, y determinar los valores óptimos para cada uno de ellos.
 - c. ¿Es posible entrenar un árbol que supere un accuracy del 80% sin sufrir de sobreentrenamiento? (Podemos considerar que no existe sobreentrenamiento cuando la diferencia entre el accuracy sobre el conjunto de evaluación y entrenamiento es menor al 5%) ¿Según las características de este tipo de modelos, le parece razonable el resultado obtenido para el conjunto de datos dado?
3. Redes neuronales:
 - a. Entrenar una red neuronal con sus hiperparámetros por defecto. ¿Considera que se obtiene un resultado satisfactorio?
 - b. Graficar cómo varía la métrica de accuracy cuando se modifica el parámetro de learning rate entre los valores: 10, 1, 0.1, 0.01, 0.001, 0.0001 y 0.00001 ¿Cómo explica el comportamiento de la red neuronal en este rango de valores?

Realizar un informe con la descripción de dataset y el desarrollo de todos los experimentos realizados junto con las conclusiones de cada punto.