

Inteligencia Computacional

Guía de trabajos prácticos 3

Métodos básicos de aprendizaje automático

Scikit-learn¹ es una biblioteca de aprendizaje automático que incluye varios algoritmos de clasificación y regresión, entre otras herramientas útiles en el desarrollo de modelos de inteligencia computacional. Se trata de una librería de software libre para Python que está diseñada para interactuar con las bibliotecas numéricas y científicas NumPy y SciPy. Para resolver los ejercicios propuestos en esta guía de trabajos prácticos aprenderemos a utilizar Scikit-learn.

Trabajos prácticos

Ejercicio 1: La biblioteca dispone de la función `train_test_split`, que permite generar una partición de entrenamiento y prueba, y `KFold`, que genera varias particiones de validación cruzada.

Estudie el uso de estas funciones y utilícelas para generar particiones a partir del conjunto de datos *Digits* que en Scikit-learn puede cargarse usando la función `load_digits` del módulo `datasets`.

Mediante `MLPClassifier` construya un perceptrón multicapa adecuado para este conjunto de datos y analice su desempeño considerando:

- 1 única partición de datos generada mediante `train_test_split`,
- 5 particiones generadas mediante `KFold`,
- 10 particiones generadas mediante `KFold`.

Para el análisis tenga en cuenta la media y la varianza de la tasa de acierto obtenida en cada partición.

Ejercicio 2: Mediante el esquema de cinco particiones generadas con `KFold`, compare el desempeño del perceptrón multicapa con los siguientes clasificadores:

- Naive Bayes
- Análisis discriminante lineal
- K vecinos más cercanos
- Árbol de decisión
- Máquina de soporte vectorial.

Ejercicio 3: Utilizando las funciones provistas por Scikit-learn, implemente los métodos de ensambles de clasificadores *Bagging* y *AdaBoost*. Compare el desempeño de estos modelos empleando 5 particiones con el conjunto de datos *Wine*.

¹<https://scikit-learn.org/>