

Práctica 2.2: Métodos alternativos de clasificación

Objetivo: El objetivo de esta práctica es introducir el uso de otros métodos de clasificación estudiados en teoría

Realice los siguientes ejercicios usando el módulo scikit-learn de Python y cualquier otro módulo adicional que considere utilizando los mismo conjuntos de datos de la práctica 2:

1. Seleccione al menos tres conjuntos de datos de **2 classes**. Es preferible conjuntos pequeños con variables nominales.
2. Para cada uno de estos conjuntos aplique el método RIPPER del módulo “*wittgenstein*” (ejemplo en `ripper.py`). Evalúe los modelos obtenidos y compárelos con un árbol de decisión.
3. Trate de interpretar los modelos obtenidos.
4. Seleccione al menos 10 conjuntos de datos. Genere un clasificador SVM con un kernel lineal y un valor fijo de $C=1$. Compare el rendimiento con un árbol de decisión.
5. Utilice el método *GridSearchCV* para obtener los mejores hiper-parámetros usando los siguientes valores:

```
# Set the parameters by cross-validation
parameters = [
    {"kernel": ["rbf"], "gamma": [0.01, 0.1, 1.0], "C": [1, 10, 100, 1000]},
    {"kernel": ["linear"], "C": [1, 10, 100, 1000]},
]
```

Compare el rendimiento de los resultados usando validación cruzada de hiper-parámetros y los métodos usando valores fijos.

6. (*) Seleccione un conjunto de datos con al menos miles de atributos. Compare el rendimiento de un modelo SVM frente a un árbol de decisión. Realice una comparación adicional mediante el test de McNemar.

NOTAS:

- Los ejercicios marcados con un * son opcionales.