

PRÁCTICA 1

Métodos de Aprendizaje Estadístico y Diseño Experimental.

En esta primera práctica, nos centraremos en los Temas 2, 3 y 8. Para ello trabajaremos con distintas técnicas estadísticas sobre dos problemas de clasificación aplicando el diseño experimental adecuado.

Sobre las técnicas a emplear:

Se corresponden con las vistas en el Temas 3. En concreto, se aplicarán el discriminante lineal de Fisher y el discriminante cuadrático.

Sobre los conjuntos de datos:

Definimos la dimensión de un conjunto de datos en función del número de ejemplos y el número de variables de entrada. A lo largo de las prácticas iremos tratando conjuntos de datos con diversas combinaciones de estos dos factores. Así, nos interesará incluir tanto conjuntos con pocos ejemplos (menos de 300) como con muchos, en ambos casos con un número de entradas razonable; conjuntos con muchas variables de entrada (50 o más) y número de ejemplos en los dos extremos, de modo que más adelante sea interesante aplicar técnicas de reducción de la dimensionalidad de entrada, etc.

En esta primera práctica nos centraremos en dos conjuntos disponibles en el UCI Machine Learning Repository (<https://archive.ics.uci.edu/ml/datasets.php>).

Nombre conjunto	Muestras	Variables	Clases
Iris	150	4	3
QSAR biodegradation	1055	41	2

Sobre la metodología:

Los experimentos se deberán realizar aplicando las técnicas explicadas en el Tema 8 dedicado a Metodología tanto en lo que se refiere a preprocesado de los datos, medidas de rendimiento, métodos para la estimación real y determinación del mejor modelo entrenado para un problema determinado. En esta práctica no se aplicará NINGÚN método para la reducción de la dimensión de los datos de entrada.

Entrega de la práctica:

Consultad CampusVirtual para fechas, instrucciones y método de entrega.

Presentación de resultados:

En las fechas indicadas en el calendario se realizará una breve presentación de resultados. Se dispondrá de documento modelo en Campusvirtual.