

# PRÁCTICA 3

## Aprendizaje basado en *Kernels* y Diseño Experimental. Parte 1.

En esta práctica, nos centraremos en los **Temas 5 y 8** de la asignatura. Para ello trabajaremos sobre dos problemas de clasificación y, en una segunda parte, con un problema de regresión **aplicando el diseño experimental adecuado**.

### **Sobre las técnicas a emplear:**

En práctica trabajaremos con aprendizaje basado en *kernels* y, en esta primera parte, usando **conjuntos de clasificación** que compararemos con las técnicas vistas en las Prácticas anteriores (discriminante lineal de Fisher, discriminante cuadrático y árboles de decisión).

### **Sobre los conjuntos de datos:**

Utilizaremos los mismos DOS conjuntos de datos que en las prácticas anteriores

### **Sobre la metodología:**

Los experimentos se deberán realizar aplicando las técnicas explicadas en el Tema dedicado a Metodología. En este caso se requiere que se hagan **3 repeticiones** de la técnica de validación escogida<sup>1</sup>. Tanto en lo que se refiere a preprocesado de los datos, medidas de rendimiento, métodos para la estimación real y determinación del mejor modelo para un problema determinado. En este caso, no se aplicará ningún método para la reducción de la dimensión de los datos de entrada.

### **Resultados:**

Se entrenarán varias versiones (mínimo 3) de las máquinas de vectores soporte modificando los distintos parámetros que la herramienta software utilizada nos permita. Por ejemplo, en MatLab podemos hacer versiones:

- Modificando el *kernel* y los parámetros asociados. Por ejemplo, en un *kernel* polinomial podemos especificar el grado del polinomio.
- Modificando el parámetro C (boxConstraint).

Estas condiciones deberán ser pensadas en el marco del problema a resolver (número de datos disponibles, existencia de clases linealmente separables, etc.). Al finalizar, tendremos resultados (medidas de rendimiento) de varios modelos entrenados utilizando las cuatro técnicas vistas hasta ahora: discriminante lineal, discriminante cuadrático, árboles de decisión y máquinas de vectores soporte (varias versiones). Utilizando las metodologías adecuadas se deberá decidir cuál de estos modelos es el mejor para resolver cada uno de los dos problemas.

---

<sup>1</sup> Algunos modelos consumen mucho tiempo de entrenamiento, es recomendable probar el modelo con una única repetición antes de ejecutar las 3 que se requieren.

**Entrega de la práctica:**

Consultad Campus Virtual 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 un documento con las instrucciones en el Campus Virtual.