

AAI - Tema 7: Máquinas de vectores de soporte

- Teoría

- Ficha
 - Necesidad de escalado: sí
 - Uso: clasificación, regresión
- Características
 - Asignación probabilística
- Fortalezas
 - Trabaja bien con pocos datos de entrenamiento
 - Trabaja bien con muchas dimensiones (truco del kernel, primal/dual)
 - Resistente a ruido/outliers con soft margin
 - Bajo uso de memoria
- Inconvenientes
 - Trabaja mal con muchas instancias/big data
- Clasificación
 - Tipos según margen
 - * Clasificación de margen suave (*soft margin*)
 - * Clasificación de margen duro (*hard margin*)
 - Hiperparámetros
 - * C : penalización de errores (+C, +penalización, +overfitting)
 - * γ (gamma): radio de influencia de instancia (+gamma, -influencia, +overfitting)
 - * Estrategias estándar multiclas
 - *One versus one* (OvO)
 - *One versus rest* (OvR)
 - * Implementaciones Scikit-learn
 - `LinearSVC`
 - Formulación primal: escalable a instancias
 - `SVC`
 - Formulación dual: NO escalable a instancias
 - Regresión
 - * Hiperparámetros
 - ϵ (épsilon): ancho de la calle
 - C
 - γ (gamma)
 - * Implementaciones Scikit-learn
 - `LinearSVR` (~`LinearSVC`)
 - `SVR` (~`SVC`)
- Tipos de formulación
 - Primal
 - * Complejidad memoria: $O(n)$
 - * Complejidad de crecimiento: $O(m \times m)$
 - * m es lineal (núm. muestras) \Rightarrow escalable a muestras
 - * depende de n (núm. características) \Rightarrow NO escalable a características

ticas

- Dual
 - * No lineal, *kernel trick*
 - * Complejidad memoria: $O(m^2)$
 - * Complejidad crecimiento: $\in O(m^2 \times n), O(m^3 \times n)$
 - * Depende de $m \times m$ (núm. muestras) \Rightarrow NO escalable a características
 - * No depende de n (núm. características) \Rightarrow escalable a muestras
- Tipos de soluciones
 - Lineal
 - No lineal
 - * Tipos de kernels
 - Linear
 - Polinomial
 - Función de base radial gaussiana / *Gaussian RBF*
 - Sigmoide
 - String

La distancia a la frontera de decisión es una medida natural de confianza en SVM (aunque no es una probabilidad).