

Pauta Control 2 - Máquinas de Aprendizaje (INF-393)

Semestre II 2017 - Viernes 24.11.17

Respuestas:

1. (15 puntos) El parámetro C controla el *tradeoff* entre el margen de clasificación y el error sobre el conjunto de entrenamiento, permitiendo que algunos datos sean parcialmente ignorados en la definición del primero. Debido a que un mayor margen ofrece garantías teóricas de una mayor capacidad de generalización, usar un valor de $C \ll \infty$ podría permitir obtener un clasificador con menor error de pruebas aún si se obtiene un error de entrenamiento mayor.
2. (15 puntos) Se puede evitar el sobreajuste en árboles de decisión controlando el nivel de crecimiento del árbol, o podándolo luego de realizar el entrenamiento.
3. (25 puntos) Utilizando la ayuda, tenemos que

$$\arg \max_{\mathbf{w}} \mathbf{w}^T S_B \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^T S_W \mathbf{w} = 1. \quad (1)$$

Entonces el Lagrangiano es

$$\mathcal{L}(\mathbf{w}, \lambda) = \mathbf{w}^T S_B \mathbf{w} - \lambda(\mathbf{w}^T S_W \mathbf{w} - 1)$$

Derivando el lagrangiano respecto de λ tenemos que:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = S_B \mathbf{w} - \lambda S_W \mathbf{w}$$

Igualando esta derivada a cero tenemos que:

$$\begin{aligned} S_B \mathbf{w} &= \lambda S_W \mathbf{w} \\ S_W^{-1} S_B \mathbf{w} &= \lambda \mathbf{w} \end{aligned} \quad (2)$$

Por lo tanto, \mathbf{w} es un vector propio de $S_W^{-1} S_B$ y λ es su valor propio.

Reemplazando (2) en el Lagrangiano tenemos que encontrar λ que maximice

$$\mathbf{w}^T S_B \mathbf{w} - \mathbf{w}^T S_W \lambda \mathbf{w} + \lambda = \mathbf{w}^T S_B \mathbf{w} - \mathbf{w}^T S_W S_W^{-1} S_B \mathbf{w} + \lambda = \mathbf{w}^T S_B \mathbf{w} - \mathbf{w}^T S_B \mathbf{w} + \lambda = \lambda$$

Por lo que λ debe ser el mayor valor propio y por lo tanto \mathbf{w} es su vector propio asociado.