

Pauta Control 1 - Máquinas de Aprendizaje (INF-393)

Semestre II 2017 - Viernes 13.10.17

Respuestas:

1. El conjunto de entrenamiento permite al algoritmo de aprendizaje seleccionar la hipótesis que modela el mapeo entre X e Y . Medir el desempeño de la hipótesis seleccionada en un conjunto de validación (diferente al conjunto de entrenamiento) permite ajustar los hiperparámetros del modelo y el conjunto de pruebas permite evaluar el desempeño del algoritmo de aprendizaje sobre datos que no han sido vistos en la etapa de entrenamiento.
2. Selección de atributos escoge un subconjunto de los atributos de entrada, en cambio reducción de dimensionalidad se refiere a cualquier técnica que escoja un número de características de entrada menor a la del espacio de entrada original. Por ejemplo, los atributos de entrada podrían ser un conjunto de nuevas características de entrada que son obtenidas a partir de las originales. Un modelo obtenido con selección de atributos es fácil de interpretar, sin embargo, una desventaja vital de feature selection es que el problema es de naturaleza combinatorial y por lo tanto se vuelve infactible rápidamente con la dimensionalidad. En cambio, si se utilizan otro tipo de características para modelar los datos de entrada, el problema puede ser más fácil de resolver, pero podría perderse la interpretabilidad de la solución.
3. Para la matriz de covarianzas, usaremos el supuesto que $E[\epsilon\epsilon^T] = \sigma^2\mathbf{I}$:

$$\begin{aligned}\text{Cov}(\hat{\beta}^{\text{ls}}) &= E \left[\left(\hat{\beta}^{\text{ls}} - E \left[\hat{\beta}^{\text{ls}} \right] \right) \left(\hat{\beta}^{\text{ls}} - E \left[\hat{\beta}^{\text{ls}} \right] \right)^T \right] = E \left[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \epsilon \epsilon^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right] \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E [\epsilon \epsilon^T] \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}\end{aligned}$$