

Aprendizaje Automático

Examen del 7.11.2024

Clasificación

- 1) Importa los datos del fichero `dataset_classification.csv` y realiza el análisis exploratorio de los datos. Describe en el cuadernillo los resultados de este análisis.

Deposita el código Python en Aula Virtual en el fichero `answer1.ipynb`

- 2) Implementa un modelo de clasificación para este conjunto de datos usando el método SVM. Usa el 70 % de los datos para el entrenamiento y el 30 % para las pruebas.
- a. ¿Es recomendable estandarizar los datos antes de usar este método de clasificación? (contesta en el cuadernillo)
 - b. Determina la función kernel y los parámetros que proporcionan los mejores resultados en términos de precisión. Indica la precisión que se ha alcanzado. (contesta en el cuadernillo)
 - c. ¿Cuál es el valor de la variable de salida si el vector de características de entrada es $[4, 3, 2, 4, 2, 2]$? (contesta en el cuadernillo)

Deposita el código Python en Aula Virtual en el fichero `answer2.ipynb`

- 3) Implementa un modelo de clasificación para este conjunto de datos usando el método Decision Tree. Usa el 75 % de los datos para el entrenamiento y el 25 % para las pruebas.
- a. ¿Es recomendable estandarizar los datos antes de usar este método de clasificación? (contesta en el cuadernillo)
 - b. Determina el criterio de división en los nodos y los parámetros que proporcionan los mejores resultados en términos de precisión. Indica la precisión que se ha alcanzado. (contesta en el cuadernillo)
 - c. Determina la característica de los datos que el algoritmo ha seleccionado para la división en el nodo raíz. (contesta en el cuadernillo)
 - d. ¿Cuál es el valor de la variable de salida si el vector de características de entrada es $[1, 3, 1, 2, 3, 2]$? (contesta en el cuadernillo).

Deposita el código Python en Aula Virtual en el fichero `answer3.ipynb`

Regresión

- 4) Importa los datos del fichero `dataset_regression.csv` y realiza el análisis exploratorio de los datos. Describe en el cuadernillo los resultados de este análisis.

Deposita el código Python en Aula Virtual en el fichero `answer4.ipynb`

- 5) Implementa un modelo de regresión lineal para este conjunto de datos. Usa el 75 % de los datos para el entrenamiento y el 25 % para las pruebas.
- a. ¿Qué rendimiento se alcanza con este modelo? (contesta en el cuadernillo)
 - b. ¿Cuál es el valor de la variable de salida si el vector de características de entrada es $[180.0, 10.4, 120.0, 28, 162.0, 765.0, 830.0, 275.0]$? (contesta en el cuadernillo)
 - c. Determina el valor de los coeficientes del modelo y comprueba el resultado obtenido en el punto anterior. (contesta en el cuadernillo)

Deposita el código Python en Aula Virtual en el fichero `answer5.ipynb`

6) Implementa un modelo de regresión polinomial cuadrática para este conjunto de datos. Usa el 70 % de los datos para el entrenamiento y el 30 % para las pruebas.

- a. ¿Qué rendimiento se alcanza con este modelo? (contesta en el cuadernillo)
- b. ¿Cuál es la dimensión del vector de características transformadas? (contesta en el cuadernillo)
- c. ¿Cuál es el valor de la variable de salida si el vector de características de entrada es [207.0, 5.0, 161.0, 28, 179.0, 736.0, 867.0, 132.0]? (contesta en el cuadernillo)

Deposita el código Python en Aula Virtual en el fichero answer6.ipynb

7) Implementa un modelo de regresión cúbica para este conjunto de datos. Usa el 70 % de los datos para el entrenamiento y el 30 % para las pruebas.

- a. ¿Qué rendimiento se alcanza con este modelo? (contesta en el cuadernillo)
- b. ¿Cuál es la dimensión del vector de características transformadas? (contesta en el cuadernillo)
- c. ¿Cuál es el valor de la variable de salida si el vector de características de entrada es [153.0, 8.0, 194.0, 28, 192.0, 623.0, 935.0, 149.0]? (contesta en el cuadernillo)

Deposita el código Python en Aula Virtual en el fichero answer7.ipynb