

Aprendizaje Automático

Trabajo Práctico 3

Para la realización de los ejercicios, se consideran los datos del archivo `Estudiantes.csv` los cuales fueron descargados del enlace <https://www.kaggle.com/datasets/amrmaree/student-performance-prediction>. Este conjunto de datos contiene información sobre estudiantes que rindieron un examen.

Ejercicio 1. Se desea conocer el rendimiento de un estudiante al rendir un examen teniendo en cuenta la cantidad de horas de estudio por semana, las calificaciones obtenidas previamente y el porcentaje de asistencia a las clases.

1. Dividir aleatoriamente el conjunto de datos para utilizar una parte de los mismos como conjunto de entrenamiento (el 80 %) y otro como conjunto de prueba (el 20 %).
2. Hallar la ecuación del modelo de regresión lineal múltiple y calcular el coeficiente de determinación R^2 .
3. Hallar la ecuación del modelo de regresión lineal logística, construir la matriz de confusión para el conjunto de prueba y calcular el valor de *accuracy* y el F1-score. (Este modelo puede construirse con bibliotecas específicas.)
4. Utilizar cada uno de los modelos hallados para predecir la calificación obtenida y la condición de aprobación de un estudiante que le dedicó 25 horas de estudio a la semana, que asistió un 58 % a las clases y obtuvo una calificación de 68 puntos en exámenes previos.

Ejercicio 2. Utilizando los mismos conjuntos de entrenamiento y de prueba que en el Ejercicio 1., se pide lo siguiente.

1. Aplicar el método de SVM para clasificar los ejemplos del conjunto de prueba, utilizando diferentes *kernels* y diferentes valores del parámetro C , costo de violación de restricción. (Este punto puede realizarse con bibliotecas específicas.)
2. Para cada uno de los modelos construidos en el inciso anterior, calcular la matriz de confusión para el conjunto de prueba, el valor de *accuracy* y el F1-score.

Ejercicio 3. ¿Cuál de todos los métodos implementados en este trabajo se recomienda utilizar para predecir si un estudiante aprobará el examen o no? Justificar la respuesta.