



Enunciado

Los datos objeto de análisis se corresponden con resultados de las pruebas de selectividad. Los datos que describen cada alumno contienen la siguiente información: año, convocatoria, localidad del centro, opción cursada (de 5 posibles), calificaciones parciales obtenidas en lengua, historia, idioma y las tres asignaturas opcionales, así como la designación de las asignaturas de idioma y las 3 opcionales cursadas, calificación en el bachillerato, calificación final y si el alumno se presentó o no a la prueba. Por tanto, puede comprobarse que la cabecera del fichero de datos, "selectividad.arff", sigue el formato mencionado anteriormente en el tutorial.

En este caso, el objetivo es relacionar los resultados obtenidos en las pruebas y determinar si pueden establecerse relaciones claras entre los atributos numéricos, utilizando algoritmos de modelos de relaciones numéricas:

1. ¿Qué características tienen los datos y como se distribuyen? Visualizar histogramas de cada uno de los atributos y realizar operaciones de filtrado
2. ¿qué relaciones entre variables son las más significativas? (utilizar opciones de filtrado y visualización para analizar dependencias entre atributos)
3. ¿se puede predecir la calificación del alumno con alguna variable conocida? Explorar métodos de aprendizaje de predicción numérica sobre las variables disponibles