Actividad: Clasificación con máquina de vectores de soporte y redes de neuronas

Objetivos

Mediante esta actividad se pretende que pongas en práctica la creación de modelos basados en máquinas de vector de soporte y redes de neuronas. **El objetivo** es que comprendas de forma práctica, con un problema determinado, **las diferencias que existen a la hora de entrenar los diferentes modelos.**

Pautas de elaboración

El conjunto de datos con el cual vamos a trabajar está adjunto en la actividad. Para mayor información, revisar el siguiente enlace:

<https://www.kaggle.com/iabhishekofficial/mobile-price-classification#train.csv>

Contiene la información necesaria sobre los datos; la variable respuesta es «price\_range»

* Análisis descriptivo de los datos:
  + De las variables numéricas, halla datos estadísticos.
  + De las variables categóricas, lista las diferentes categorías y halla la frecuencia de cada una de ellas.
  + Crea matriz de correlaciones existentes entre las variables numéricas del conjunto de datos, analiza los resultados.
  + Considerando la variable de respuesta, verifica si el dataset está balanceado.
* Aplica máquina de vectores de soporte y redes neuronales.
* Compara, mediante las medidas que te parezcan adecuadas, la capacidad predictiva de ambos métodos.
* Comenta las ventajas y desventajas de cada modelo. De acuerdo con los resultados, ¿son realmente útiles los modelos creados para el conjunto de datos propuesto?
* Otros comentarios que consideres adecuados.

Formato de entrega

Comenta todos los resultados obtenidos, exporta el *notebook* a un fichero con extensión .PDF y comenta con todos los puntos que se solicitan. Es importante asegurarse que la visualización es correcta y todos los apartados del ejercicio son legibles (se puede obtener directamente desde los *notebooks* «imprimiendo» en PDF).

Rúbrica

|  |  |  |  |
| --- | --- | --- | --- |
|  | Descripción | Puntuación máxima  (puntos) | Peso  % |
| Criterio 1 | Realiza el análisis descriptivo de los datos de manera adecuada | 2 | 20 % |
| Criterio 2 | Aplica máquina de vectores de soporte sobre los datos de entrenamiento | 3 | 30 % |
| Criterio 3 | Aplica adecuadamente la técnica de redes neuronales | 3 | 30 % |
| Criterio 4 | Conclusiones sobre los resultados. | 2 | 20 % |
|  |  | **10** | **100 %** |