Actividad: Detección de anomalías y técnicas de agrupamiento

Objetivos

Mediante este trabajo se pretende que pongas en práctica la aplicación de los algoritmos de detección de anomalías (*outliers*) y las técnicas de agrupamiento. El objetivo es que comprendas de forma práctica, con un problema determinado, los pasos que hay que realizar para la detección automática de valores inusuales y, por otro lado, para analizar los clúster o grupos resultado de aplicar un algoritmo de agrupamiento.

Pautas de elaboración

El conjunto de datos con el cual vamos a trabajar se encuentra en el siguiente enlace:

<https://www.kaggle.com/shubhamjoshi2130of/abstract-data-set-for-credit-card-fraud-detection>

* 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 y analiza los resultados.
* Aplica una técnica de detección de anomalías.
* Aplica una técnica de agrupamiento.
* 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.

Extensión y formato

Extensión máxima de la actividad: 20 páginas. Formato: Calibri 12, interlineado 1,5 puntos.

Rúbrica

|  |  |  |  |
| --- | --- | --- | --- |
|  | Descripción | Puntuación máxima  (puntos) | Peso  % |
| Criterio 1 | Realiza el análisis descriptivo de los datos de manera adecuada | 2 | 20 % |
| Criterio 2 | Aplica técnicas de clúster sobre los datos | 3 | 30 % |
| Criterio 3 | Aplica técnicas de detección de anomalías | 3 | 30 % |
| Criterio 4 | Los comentarios sobre los resultados son pertinentes y adecuados | 2 | 20 % |
|  |  | **10** | **100 %** |