

Control Practico 1

Implementación de técnicas de Clusters

Profesor: Fredy Troncoso

Objetivo

Aplicar técnicas de clúster a una base de datos real, analizando los resultados concluyendo respecto a estos.

Desarrollo

- Pre procese la base de datos de manera que esta pueda ser aplicable a técnicas de clusters.
- Aplique las técnicas de Clusters vistas en clases a la base de datos BlackFriday, determinados el número de clusters óptimo, según las medidas vistas en clases.
- Interprete los resultados e indique toda la información que pueda ser útil para la toma de decisiones en el negocio.
- Investigue otra técnica de cluster, diferente a las vistas en clases. Deberá investigar y describir esta técnica en forma teórica, describiendo los principios que la fundamentan y su implementación el RapidMiner o R.
- El informe debe ser entregado en formato Power Point, además de los archivos de respaldo relacionado con el pre procesamiento de la base de datos y la implementación de los cluster.
- El trabajo debe ser desarrollado en forma individual o en parejas. La fecha de entrega es el miércoles 10 de abril en la hora de clases y ese mismo día se deberá realizar una presentación de más 10 minutos, sin público.

Descripción base de datos

A retail company “ABC Private Limited” wants to understand the customer purchase behaviour (specifically, purchase amount) against various products of different categories. They have shared purchase summary of various customers for selected high volume products from last month.

The data set also contains customer demographics (age, gender, marital status, city_type, stay_in_current_city), product details (product_id and product category) and Total purchase_amount from last month.

Now, they want to build a model to predict the purchase amount of customer against various products which will help them to create personalized offer for customers against different products

<https://www.kaggle.com/mehdidag/black-friday>