Modulo 5
Analítica de datos y extracción de conocimiento mediante técnicas de IA

Task 6

Description:

Your company wishes to better understand the evolution of our sales globally.

More specifically, it is unclear whether we should create a VIP program for our clients, focus on particular customer segments or which markets, if any, we should pay more attention to. We also do not know if higher category shipments are being profitable or we are losing money instead and should drop them.

Try to help your company improve its situation.

Submit a pdf file applying machine learning techniques to try and extract insights from the sales data. **For each technique**, you should include (i) a **screenshot of the implementation in jupyter**, (ii) **a visualization of the model or its results (if it is possible),** (iii) **an interpretation of the results obtained** by the model (accuracy, coefficients)

Remember that the file submitted must **include the full name(s) of the student(s) involved**.

To do:

Using the data from **the previous session** OR from **the ORACLE database**, add a new column 'VIP' with two values: 0 and 1. The value for VIP will be 1 for those in the list of top 50 customers according to the profit made thanks to them, 0 for the rest.

Train and interpret the output of the following models:

- Decision Tree with target to predict: VIP
- Decision Tree Regressor with target to predict: Profit
- Linear Regression with target to predict: Profit
- Kmeans with scaler