Week 4: Predict Car sales using Pyspark

Lecturer: Nang Nguyen

Data: /data/Car_Purchasing_Data.csv

PROBLEM STATEMENT

You are working as a car salesman and you would like to develop a model to predict the total dollar amount that customers are willing to pay given the following attributes:

- Customer Name
- Customer e-mail
- Country
- Gender
- Age
- Annual Salary
- Credit Card Debt
- Net Worth

The model should predict:

- Car Purchase Amount

Small guide

- *Read the file, look at attributes*
- *Brainstorm how to solve the problem statement*
- *Feature selection*
- *Model selection*
- *How to evaluate the model*
- *Predict*