Assignment 4

Objective
Identify Fraudulent Credit Card Users.

Dataset
Data: https://www.kaggle.com/mlg-ulb/creditcardfraud?select=creditcard.csv

Possible Work to be done
1. How can the problem statement help the industry?
2. State the different types of variables in the dataset.
3. Identify the key independent variables critical for classification.
4. Which variables are ideal candidates for feature transformation?
5. Transform any one variable which is interval/ratio to ordinal variable
6. Is there any association between any two interval/ratio variable? How will you go about it?
7. Which graph type would you prefer to use for visualizing each variable.
8. Do you think decision tree is an ideal classifier for this problem?
9. Investigate and compare the performance of decision tree and Naive Bayes Classifier.
10. What are the challenges in deploying the solution in real time?