**Problem Description**

Once upon a time, there was a bank offering services to private persons. The services include managing of accounts, offering loans, etc. The bank wants to improve their services by finding interesting groups of clients (e.g. to differentiate between good and bad clients). The bank managers have only vague idea, who is good client (whom to offer some additional services) and who is bad client (whom to watch carefully to minimize the bank loses). Fortunately, the bank stores data about their clients, the accounts (transactions within several months), the loans already granted, the credit cards issued So the bank managers hope to find some answers (and questions as well) by analyzing this data.

***Task 1: You’re required to build a machine learning/deep learning model to predict customers who are more likely to default on their loan using the data provided.***

As a DS, you will also need to perform the following tasks to ensure that your model is robust:

1. **Exploratory data analysis (EDA)**

* Some data insights incl. data visualization eg. Target distributions
* Imputation
* Outlier detection
* Sampling

1. **Feature selection & feature engineering**

* Feature selection
* Multicollinearity
* Feature engineering incl normalization/scaling/binning etc

1. **Model development, performance & interpretation**

* Build at-least two ML/DL models & cross (validate) the models.
* Model performance & stability.
* Model interpretation/explainability (Responsible AI)

***Task 2: Please create a brief ppt presentation on the use-case. Also include a slide on post model development analysis (deployment, Monitoring & documentation).***

***Notes:***

* *Share both the presentation and the code used in the above exercise.*
* *Use either R/Python/Pyspark/Spark-R to complete the case study.*
* *Use the 1st column of the app data (called Def\_target) as the target/response of the models.*

*Good luck*