



Default Prediction Model PROJECT PROPOSAL

PREPARED BY:

Pralad Kadel

Data Analyst Apprentice



Chakupat-10, Lalitpur
Nepal

coderush.com.np

6th Sep, 2022

Introduction

- One of the main income earning assets for a bank is loan
- Default on loans diminishes asset quality of a bank
- Important for banks to invest their assets in safe hands

Objective

- To build a default prediction model based on historical data that can predict whether it is safe to issue a loan to a particular applicant or not.

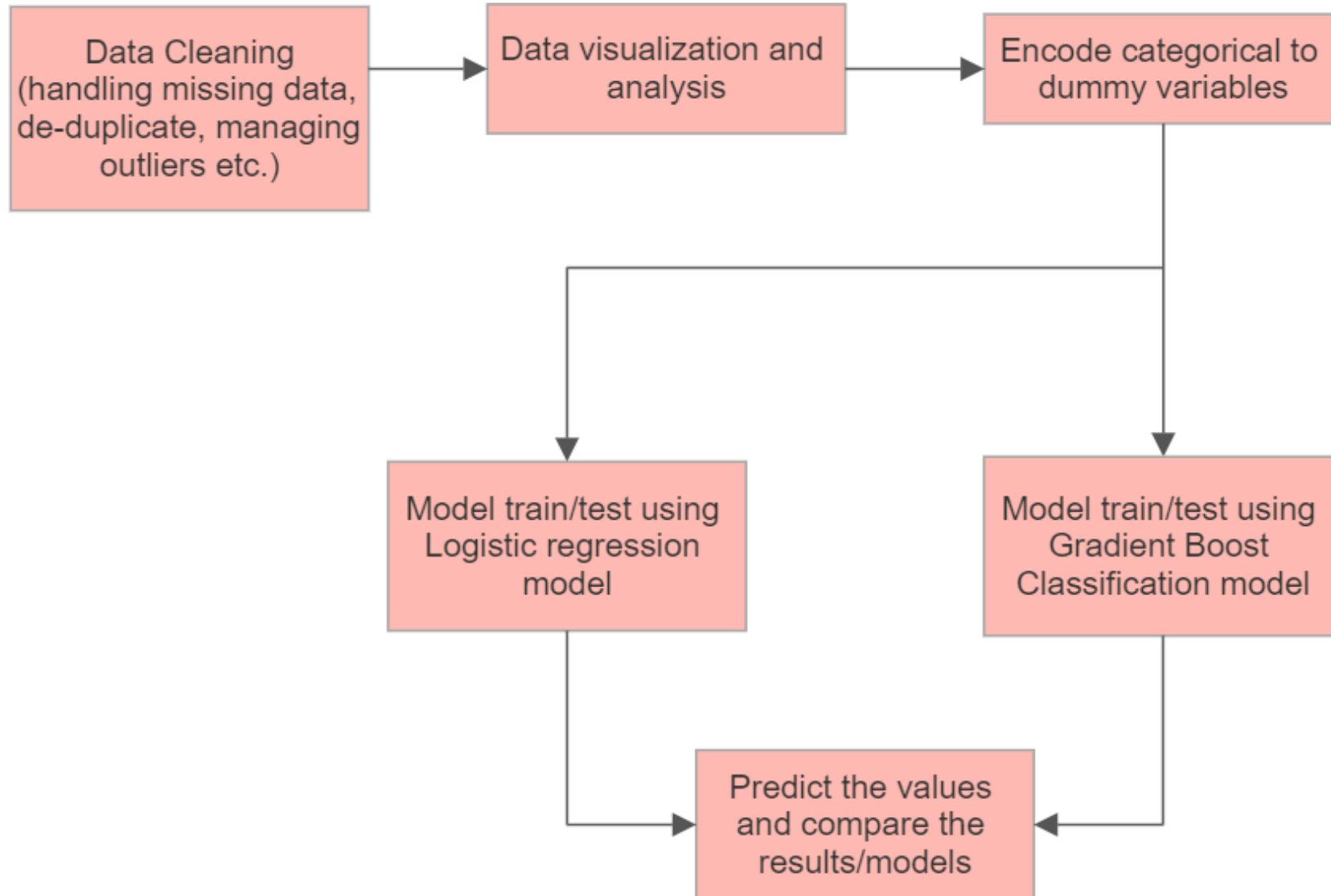
Dataset

- From Kaggle website
- Dataset contains columns simulating credit bureau data
- No. of cases: 32,582
- No. of Defaults: 7,108 (21.82%)
- No. of variables/features: 11
- Total loan amount : 312,431,300
- Total loan amount for default : 77,125,375 (24.68%)

Related Work

- **Reddy and Kavitha (2010)** showed that using Neural Networks through attribute relevance analysis to build a prediction model increases the speed of Neural Network and feasible accuracy.
- **Shoumo et al. (2019)** focused on applying appropriate dimensionally reduction approach using Recursive Feature Elimination with Cross-Validation (RFECV) and Principal Component Analysis (PCA), noise handling, parameters tuning, using a grid search with cross-validation and on handling the imbalanced data problem.

Methodology



Limitations

- Cannot generalize the model as the data is taken from only one company/source
- Other various factors/features/variables can affect the default process that are not included in the dataset

Conclusion

- Being able to determine whether the customer are likely to become default or not is very beneficial.
- We will pre-process the data, perform EDA and then build prediction models, plus, also evaluate the performance of the models.

Reference

- A. Jeremy Mahoney (2020, Sep 9), Credit risk modeling with machine learning. towardsdatascience site, Last accessed 3rd Sep 2022:
<https://towardsdatascience.com/credit-risk-modeling-with-machine-learning8c8a2657b4c>

THANK YOU