Project Proposal
Data Analysis Project: Personal Loan Campaign

Overview:

Personal loans have lower fundraising costs and are more adaptability to individual needs [1]. Bank plans to expand their personal loan business by launching the campaign to specially offer personal loans to their target customers. The last campaign showed a 9% success rate, so bank needs to improve their campaign and better identify potential customers to increase the success rate of their campaign.

Objective:

The goal of this project is to determine whether or not the customers will respond to bank's personal loan campaign by accepting the loan offered during the campaign period. Additionally, several statistical models will be developed to predict campaign outcomes and to identify the important characteristics of bank's customers who are likely to accept the personal loans offered by the bank.

Dataset:

The "Bank Personal Loan" dataset contains 5,000 observations, retrieved from https://www.kaggle.com/itsmesunil/bank-loan-modelling.

The dataset has 13 independent attributes including 7 numeric attributes such as age, income, and the average monthly spending on credit cards, and 6 categorical attributes such as education level and binary attributes. The target attribute is a binary variable (yes/no), which identifies whether or not the customers accepted personal loan offered in the last campaign.

Proposed Methodology:

The classification algorithms will be used to solve this supervised learning problem to predict target (binary) attribute. I plan to utilize several Python libraries including Pandas, Matplotlib, and NumPy. The models will be developed using k-nearest neighbors algorithm (KNN) and Decision Tree. The following steps will be implemented for this project:

- Explore the distribution and basic statistics and visualize data of all attributes
- Preprocess data by creating bins, normalizing numeric attributes, and creating dummy variables for categorical attributes
- Perform correlation analysis to identify the association among customers' characteristics and target attribute
- Split data into a training set and a test set
- Develop KNN model and Decision Tree
- Perform model validation to compare different models and select the final model to make a prediction

Reference [1]: " Financial instruments working with personal loans " Retrieved from https://www.fi-compass.eu/sites/default/files/publications/factsheet_personal_loans_0.pdf