Proposal

Background

The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it as their advantage by becoming a defaulter. When the company receives a loan application, two types of risks are associated with the bank's decision: a. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company. b. If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

Research question

The objective is to identify patterns which indicate if a client has difficulty paying their installments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable of repaying the loan are not rejected.

Dataset

To help the bank to identify each customer's risk on whether they could pay back the loan, a dataset found in Kaggle(https://www.kaggle.com/datasets/arkapravasen/bank-loan-default) that includes a wide range of variables related to personal, financial, employment, and loan characteristics. This enables a holistic assessment of an applicant's profile, encompassing factors known to influence loan repayment behaviors. The detailed financial attributes (like income, credit amount, loan annuity, and previous credit inquiries) allows for a clearly understanding of an individual's financial health and behavior. Also, Socio-demographic variables (such as age, gender, family status, and number of children) provide insights into the applicant's life stage and responsibilities, which can affect their loan repayment capability. By using machine learning algorithms, we can analyze this data to predict the likelihood of default, which could help bank to improve the decision-making process for loan approvals.

Why it is worthwhile?

Installments as an important part that may influence the operation of the economic system. However, clients may face credit risks that may have difficulties on their installments that cannot be applied to installments, which may increase the workload for banks. This projects' objective is to identify patterns which indicate if a client has difficulty paying their installments.

By predicting clients who may face difficulties in paying their installments, the enterprise can better work on risk management associated with loan defaults. This early determination can

decrease the workload for front-end staff, but also allows for the implementation of restructuring the loan, offering finance consulting, and then reducing the financial losses. Through machine learning models, we can automate the process of identifying at-risk clients to increase operational frequency. It reduces the need for manual review of each client's account to check whether they are under credit risks, and allows the enterprise to allocate resources and services more frequently and focus on the decision-making of other non-risk clients and approve loan applications for them. In other ways, by understanding the clients difficulties and then addressing their problem on the installments not only aids in risk mitigation but also enhances clients' satisfaction and customer retention. By being sympathetic to the customers' requirements and offering support to customers, it can better improve the relationship between the customer and enterprise to win more respect and loyalty. Also by predicting the difficulties, the enterprise can get date-driven decision-making based on the machine learning model analysis. The enterprise can make adjustments on credit policies, make research and development on new financial products to fit more clients, especially those with different level difficulties for improvement of the loan approval process. Leveraging advanced analytics and machine learning for risk management can provide a competitive edge. It demonstrates to stakeholders and clients alike that the enterprise is forward-thinking, capable of innovating, and committed to maintaining financial stability and customer satisfaction.

Based on the machine learning models and availability and quality of data, we will have successful project outcomes. With access to comprehensive datasets on client behavior, credit information, and target variables on difficulties, there's a strong foundation for building accurate predictive models. By using Logistic Regression model, Decision Tree Model, Random Forest Tree model, we can capitalize on the strengths of each to improve overall predictive performance. Also Machine learning projects benefit from an iterative approach, where models are continuously refined and improved based on feedback and new data. This adaptability enhances the project's chances of achieving its objectives.

Conclusion

The strategic importance of using machine learning to identify clients at risk of payment difficulties, combined with the technical feasibility and the potential for significant positive impact, justifies the investment in this project. With careful planning, execution, and ongoing refinement, the project stands a good chance of succeeding and delivering substantial value to the enterprise.