Predicting the Acceptance of Personal Loans

Haotian Zhang

March 2025

1 Introduction

In today's rapidly evolving and highly competitive financial environment, banks face mounting pressure to innovate their customer engagement strategies, diversify revenue streams, and enhance operational efficiency (Krasnikov et al., 2009). A critical challenge for many financial institutions is optimizing the composition of their customer portfolios—specifically, striking an effective balance between liability customers, who maintain deposits, and asset customers, who generate interest income through lending products (Nikiel and Opiela, 2002). Thera Bank, a U.S.-based, growth-oriented retail bank, epitomizes this challenge. Currently, the majority of its customers are liability holders with varying levels of deposits. While these customers contribute to funding stability, their direct impact on revenue is relatively limited when compared to asset customers, who are more directly linked to income-generating loan products (Machauer and Morgner, 2001).

In response to this structural imbalance, Thera Bank has embarked on a strategic shift focused on expanding its asset customer base. Specifically, the bank is pursuing the conversion of existing liability customers into personal loan holders—an approach that not only leverages established customer relationships, thereby reducing acquisition costs, but also has the potential to significantly increase customer lifetime value (Moro et al., 2014). Central to this strategy is the need to determine which liability customers are most likely to respond positively to personal loan offers. This question lies at the heart of the present research, which aims to help the bank identify high-potential target customers in order to more effectively tailor its marketing strategies and optimize promotional resource allocation. Accurately predicting these conversion probabilities is essential not only for optimizing marketing spend and improving response rates, but also for increasing the return on promotional investments (Fenton, 2007).

Initial results from Thera Bank's targeted campaigns have been encouraging. Over the past year, the bank recorded a personal loan conversion rate exceeding 9%, a figure that surpasses internal benchmarks and validates the promise of targeted marketing. Building on this momentum, the retail marketing team is now seeking more advanced methods to enhance campaign precision and scale outreach. The team contends that machine learning offers a robust and scalable framework for data-driven targeting—one capable of accurately predicting which customers are most likely to adopt a loan product (Ala'raj et al., 2021).

To this end, the current study develops a predictive classification model that estimates the likelihood of liability customers converting to personal loan users, based on variables such as demographics, financial behavior, and previous engagement with the bank. By uncovering patterns that distinguish likely converters from non-responders, the model is designed to support more efficient customer segmentation and tailored marketing interventions. In addition to enhancing campaign performance, this approach also advances personalization efforts by aligning offers with customer needs and preferences.

Broadly speaking, this research contributes to the expanding literature on predictive analytics in financial services by demonstrating how data-driven modeling can inform customer segmentation and marketing strategy in real-world banking contexts. The findings are relevant not only to practitioners seeking operational insights but also to scholars exploring the applications of machine learning in consumer financial behavior, marketing analytics, and strategic decision-making.

2 Literature Review

In an era of increasing data availability and technological advancement, retail banks are undergoing a paradigm shift toward data-driven marketing and decision-making (He et al., 2022). Predictive analytics has become central to understanding customer behavior, particularly in customer acquisition, product cross-selling, and loan adoption (Boustani et al., 2024). One critical challenge lies in identifying which liability customers, who contribute deposit-based capital, are most likely to convert into asset customers through products such as personal loans (Chang et al., 2024). Accurate prediction of such conversions can significantly enhance marketing efficiency, reduce acquisition costs, and increase customer lifetime value (CLV) (Ulug and Akyüz, 2025).

The application of predictive analytics in retail banking has been widely studied in areas such as credit scoring, churn prediction, and loan default forecasting (Singh et al., 2024). However, research focusing on loan product adoption from existing customers—especially liability to asset transition—is relatively nascent (Văduva et al., 2024). Traditional marketing practices, such as rule-based segmentation and demographic targeting, are increasingly giving way to machine learning techniques that leverage complex, non-linear customer behaviors (de Waal et al., 2024).

In the context of personal loan marketing, predictive modeling allows banks to create highly targeted campaigns by identifying customers with the highest likelihood of responding to offers (Rahman et al., 2024). As noted by Bharathi S et al. (2022), the integration of customer-level behavioral features (e.g., transaction patterns, credit card usage) has dramatically improved classification accuracy in banking environments. These developments underline the importance of choosing the right algorithm—not only for predictive performance, but also for interpretability, scalability, and deployability (Al-Quraishi et al., 2025).

In the following part of Literature review, I will show some literatures of pros and cons of logistic regression, tree model and knn methods and decide which is the better model for my project.

3 Data

Including some descriptive statistics and some visualization pictures of the data set.

4 Methods

1. Logistic regression 2. Tree Model (optional) 3. knn methods (optional)

5 Findings

Use confusion matrix to evaluate the models

6 Conclusion

References

Al-Quraishi, T., Albahri, O., Albahri, A., Alamoodi, A., and Sharaf, I. M. (2025). Bridging Predictive Insights and Retention Strategies: The Role of Account Balance in Banking Churn Prediction. *AI*, 6(4):73.

Ala'raj, M., Abbod, M. F., and Majdalawieh, M. (2021). Modelling customers credit card behaviour using bidirectional LSTM neural networks. *Journal of Big Data*, 8(1):69.

Bharathi S, V., Pramod, D., and Raman, R. (2022). An Ensemble Model for Predicting Retail Banking Churn in the Youth Segment of Customers. *Data*, 7(5):61.

Boustani, N., Emrouznejad, A., Gholami, R., Despic, O., and Ioannou, A. (2024). Improving the predictive accuracy of the cross-selling of consumer loans using deep learning networks. *Annals of Operations Research*, 339(1-2):613–630.

Chang, V., Hahm, N., Xu, Q. A., Vijayakumar, P., and Liu, L. (2024). Towards data and analytics driven B2B-banking for green finance: A cross-selling use case study. *Technological Forecasting and Social Change*, 206:123542.

de Waal, H., Nyawa, S., and Wamba, S. F. (2024). Consumers' Financial Distress: Prediction and Prescription Using Interpretable Machine Learning. *Information Systems Frontiers*.

Fenton, E. M. (2007). Visualising Strategic Change:. *European Management Journal*, 25(2):104–117.

He, W., Hung, J.-L., and Liu, L. (2022). Impact of big data analytics on banking: a case study. *Journal of Enterprise Information Management.*

Krasnikov, A., Jayachandran, S., and Kumar, V. (2009). The Impact of Customer Relationship Management Implementation on Cost and Profit Efficiencies: Evidence from the U.S. Commercial Banking Industry. *Journal of Marketing*, 73(6):61–76.

Machauer, A. and Morgner, S. (2001). Segmentation of bank customers by expected benefits and attitudes. *International Journal of Bank Marketing*, 19(1):6–18.

Moro, S., Cortez, P., and Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62:22–31.

Nikiel, E. M. and Opiela, T. P. (2002). CUSTOMER TYPE AND BANK EFFICIENCY IN POLAND: IMPLICATIONS FOR EMERGING MARKET BANKING. *Contemporary Economic Policy*, 20(3):255–271.

Rahman, M. H., Chandra Das, A., Shak, M. S., Uddin, M. K., Alam, M. I., Anjum, N., Bony, M. N. V. A., and Alam, M. (2024). TRANSFORMING CUSTOMER RETENTION IN FINTECH INDUSTRY THROUGH PREDICTIVE ANALYTICS AND MACHINE LEARNING. *The American Journal of Engineering and Technology*, 06(10):150–163.

Singh, P. P., Anik, F. I., Senapati, R., Sinha, A., Sakib, N., and Hossain, E. (2024). Investigating customer churn in banking: a machine learning approach and visualization app for data science and management. *Data Science and Management*, 7(1):7–16.

Ulug, B. and Akyüz, S. (2025). Optimized churn prediction using ensemble-based feature selection via second-order cone programming. *Annals of Operations Research*.

Vădupa, A.-G., Oprea, S.-V., Niculae, A.-M., Bâra, A., and Andreescu, A.-I. (2024). Improving Churn Detection in the Banking Sector: A Machine Learning Approach with Probability Calibration Techniques. *Electronics*, 13(22):4527.