Credit Card Behaviour Score Prediction Using Classification and Risk-Based Techniques

Aditya Kumar Verma
22116008
Btech (Electronics and Communcation)
Email: aditya_kv@ece.iitr.ac.in

***Abstract*—This project presents a binary classification model to predict credit card default in the upcoming billing cycle, leveraging behavioral and demographic features. By optimizing for recall and F2 score, the solution aligns with Bank A's goal of proactive risk management and early warning activation.**

***Index Terms*—credit risk, default prediction, LightGBM, F2-score, threshold tuning, financial behavior modeling**

I. Introduction

Financial institutions need predictive tools to flag likely defaulters early. This work develops a behavior-based classification system trained on anonymized customer data, targeting high recall and interpretability to minimize credit risk.

II. Dataset Description

The dataset consists of:

- **Train:** 25,000 customers with labeled default outcomes.
- **Validation:** 5,000 customers without labels.
- **Target:** `next_month_default` (1 = default, 0 = no default)

Features include demographic details, credit limits, past payment status, bill amounts, and repayments.

III. Exploratory Data Analysis

A. Age and Credit Limit

Customers in the younger age groups and those with lower credit limits show higher default tendencies.

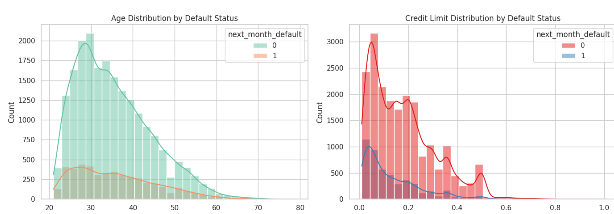

Fig. 1: Age and Credit Limit Distribution by Default Status

B. Payment History Trends

Defaulters display multiple late payments (payment status ≥ 1) across several months, unlike non-defaulters who consistently pay on time.

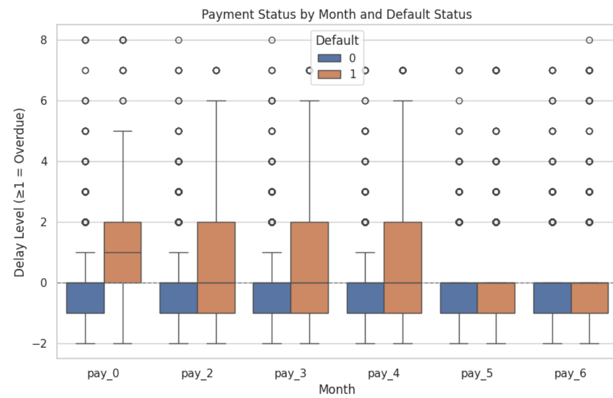

Fig. 2: Boxplot of PAY Variables by Default

IV. Feature Engineering

To enhance interpretability and predictive power, we created:

- **AVG_Pay_amt:** Mean of 6-month repayments.
- **Utilization_Ratio:** Average bill divided by credit limit.
- **Delinquency_Streak:** Count of months with PAY ≥ 1.
- **Repayment_Std:** Std deviation of past repayments.

V. Model Training and Evaluation

Four classifiers were evaluated:

- Logistic Regression
- Random Forest
- XGBoost
- **LightGBM (best performer)**

Metrics prioritized: F2 Score, Recall, F1 Score, Accuracy

LightGBM provided the best trade-off between recall and false positives. Class weights were applied to handle imbalance.

VI. Threshold Tuning

Since the default threshold of 0.5 yielded suboptimal recall, we performed threshold optimization. The best threshold was found to be 0.36, maximizing the F2 score.

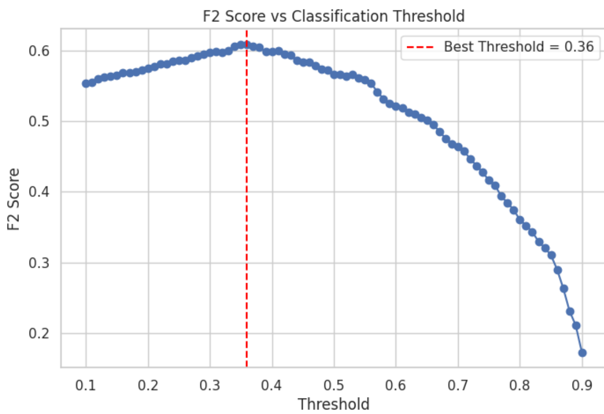

Fig. 3: F2 Score vs Classification Threshold

A. *Optimized Performance*

- **F2 Score:** 0.6092
- **Recall:** 0.78
- **Precision:** 0.32

VII. Final Submission and Output

- **Model Used:** LightGBM (class-weighted)
- **Threshold:** 0.36
- **Submission File:** `submission_22116008.csv`

VIII. Business Implications

High-risk customers can be targeted for:

- Credit exposure reduction
- Monitoring
- Early payment reminders

The use of interpretable behavioral metrics supports defensible credit decisions.

IX. Conclusion

This solution meets the financial institution's goals by delivering a recall-focused, interpretable credit default predictor. It achieves strong performance on the F2 metric while enabling actionable insights.

References

[1] UCI Credit Card Default Dataset, [Online]. Available: https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients