REPORT

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

1. *Overview of Approach and Modeling Strategy*

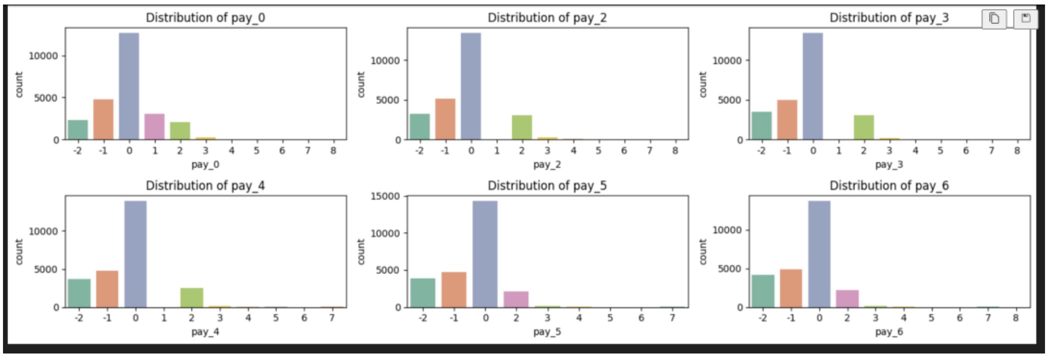
The objective of this project was to build a predictive model to identify customers likely to default on their credit obligations next month. Given the real-world nature of the dataset, special care was taken to address class imbalance, financial interpretability, and performance evaluation that aligns with business risk.

Key steps:

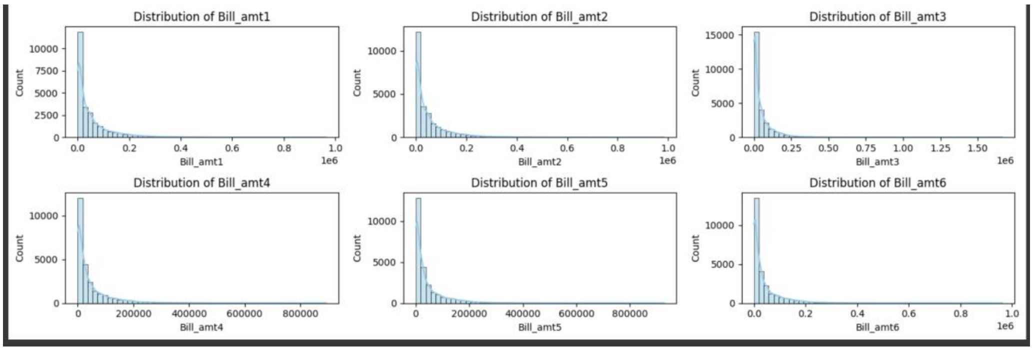
- Data cleaning, outlier handling, and missing value imputation.
- Financially meaningful feature engineering.
- Tackling class imbalance using SMOTE.
- Baseline modeling and advanced modeling (Random Forest, XGBoost, LightGBM).
- Threshold tuning
- Final predictions on a held-out validation dataset.

2. *EDA Findings and Visualizations*

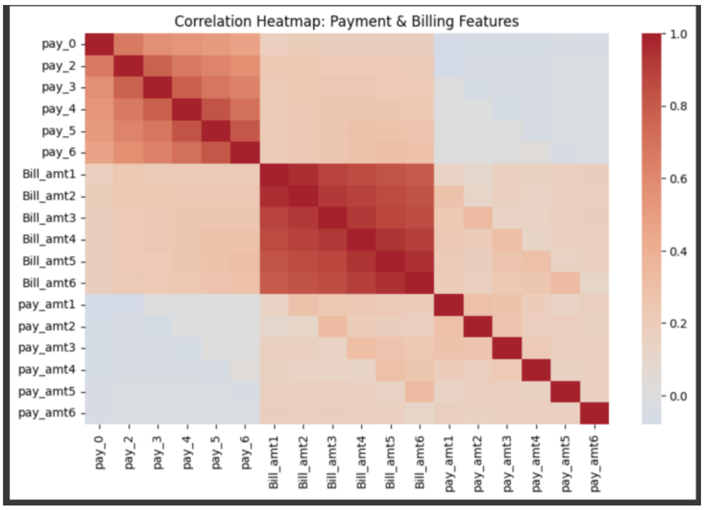
- Female customers have a slightly higher default rate than males comparing the percentage.
- educated ones are more defaulter than others
- singles are more defaulters than married and others.
- Defaulters have significantly lower average credit limits than non-defaulters. This suggests that customers with lower credit limits are more likely to default.
- Customers in lower credit limit bands have a higher default rate, indicating that low-limit customers are more financially at risk.
- Defaulters have a lower payment-to-bill ratio, meaning they are less likely to pay off their bills, a strong financial stress indicator.
- Highly imbalanced target variable (~22% default rate).
- Strong correlation between delayed payments (pay_0 to pay_6) and default.
- Credit limit (LIMIT_BAL) and past bill amounts had right-skewed distributions.
- Customers with zero repayments over several months had significantly higher default rates.
- Variables like avg_pay_amt, dues, and months_delayed clearly separated defaulting vs. non-defaulting behavior.



- ➢ PAY = 0 (Partial Payments) This is the most frequent bar across all months. Second Most Common: PAY = -1 (Fully Paid on Time)

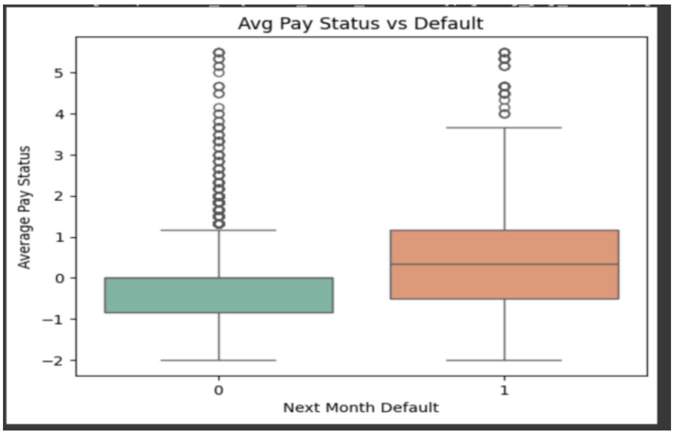


- ➢ Most values are concentrated on the left side (low amounts).
- ➢ A long tail extends to the right, indicating a small number of customers with very large bills.
- ➢ Majority of customers have monthly bills < 100,000.
- ➢ This suggests most users are using moderate credit amounts.

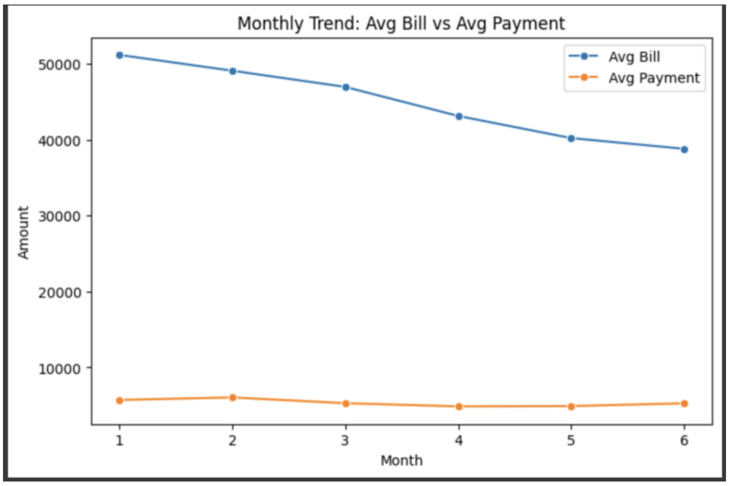


- ➢ Bill Amounts (Bill_amt1 to Bill_amt6)
- ➢ Strong positive correlations among all months (deep red).
- ➢ Implies customers with higher bills tend to carry that trend month-to-month.

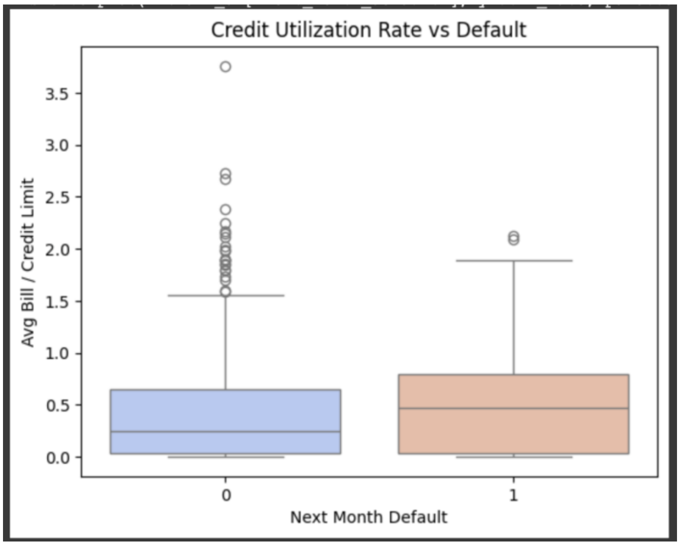
- Suggests stable credit usage.
- Payment Amounts (pay_amt1 to pay_amt6)
- Weak-to-moderate correlations among each other.
- Means people don't always pay the same amounts, even if bills are similar.
- Could be due to partial payments, minimum dues, or missed payments.
- Payment Status (pay_m) vs Bill or Pay Amounts
- Weak or slightly negative correlations with pay_amtX.
- Suggests that late payments are not always tied to high bills or payment amounts.



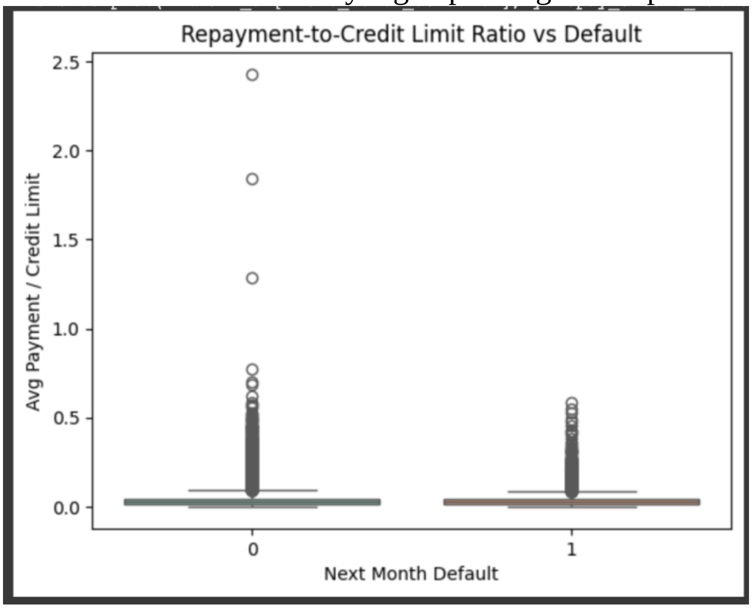
- Defaulters generally have a higher average pay status, indicating more frequent and severe payment delays
- Defaulters have a lower payment-to-bill ratio, meaning they are less likely to pay off their bills, a strong financial stress indicator.



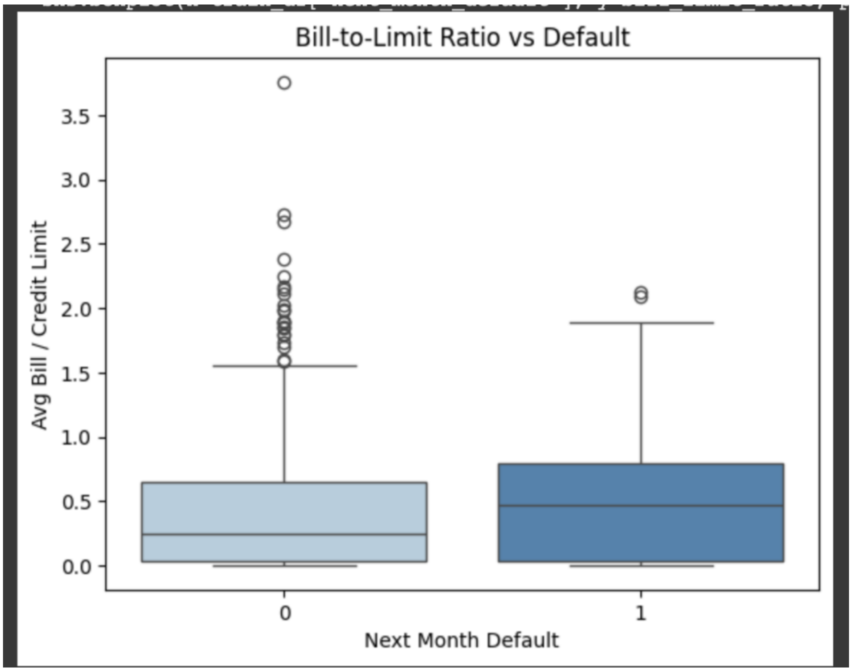
- Across all months, average bill amounts are consistently higher than payments, suggesting that many customers are underpaying or carrying balances.



- Defaulters (1) have a higher median utilization rate than non-defaulters (0). Indicates they are using more of their credit limit, which is a classic risk signal. over half of defaulters consistently carry higher credit usage. More outliers in non-defaulters Possibly high-spending but responsible customers



- Non-defaulters (0): Higher median — repaying more relative to their credit limit Defaulters (1): Lower median — repaying less, sometimes far below what they should The IQR (middle 50%) is narrower for defaulters Suggests they consistently under-repay, with less variation in behavior Some non-defaulters pay more than their credit limit (ratio > 1) Likely high-income or responsible users repaying aggressively.



- Defaulters (1) have a higher median bill-to-limit ratio than non-defaulters (0). This means defaulters consistently use more of their credit limit. The distribution for defaulters is more spread, showing that some users fluctuate wildly in billing.

3.*Financial Analysis of Default Drivers*

Feature	Description
dues	Total billed amount minus total paid amount
dues_to_limit_ratio	Measures over-utilization
zero_pay_count	Count of months with no payments
months_delayed	Number of delayed payment months
pay_trend	Linear trend in repayment over 6 months
recent_late, last_two_months_late	Recency of delay

Customers with consistently delayed or zero payments and high utilization showed much higher risk of default.

The graphs for many features , has been displayed above.

4.Model Comparison and Final Model Justification:

Models Trained:

- Logistic Regression
- Decision Tree
- LightGBM
- XGBoost
- **Random Forest (Final Chosen Model)**

Why Random Forest?

- Highest **F2 Score**: ~0.89 after threshold tuning.
- Strong **ROC AUC** and generalization performance.
- Robust to outliers and overfitting.
- Balanced recall and precision.

5. *Evaluation Methodology*

The dataset was **highly imbalanced**, with only about **22% of customers defaulting**. This posed a serious challenge because most models could easily achieve high accuracy by predicting only the majority class (non-defaulters), yet perform poorly where it really matters: **identifying actual defaulters**.

To address this:

- **SMOTE** (Synthetic Minority Oversampling Technique) was applied **only to the training set**, ensuring the model learned from a more balanced view of both default and non-default cases.
- Evaluation focused on metrics that highlight the model's ability to catch defaulters.

The performance was evaluated using multiple metrics:

- **F2 Score** (prioritized): More weight on recall.
- Accuracy, Precision, Recall, F1 Score.
- ROC AUC.

Why F2 Score?

- The **F2 score** gives more weight to **recall** than precision, which aligns with business needs in credit risk.
- In financial lending:

- ✓ A **false negative** (missing a real defaulter) could lead to **significant financial loss**.
- ✓ A **false positive** (flagging a safe customer) may result in **temporary friction**, but it's far less costly than default.

By optimizing the **F2 score**, the model becomes more sensitive to **catching defaulters**, even if it sacrifices a bit of precision.

6.*Classification Threshold Selection*

Instead of using the default 0.5 cutoff, threshold tuning was done from 0.1 to 0.9 using F2 optimization.

- **Optimal threshold for Random Forest:** `0.31`
- Boosted recall significantly while maintaining acceptable precision.
- Used `fbeta_score` (beta=2) for evaluation during tuning.

7.*Business Implications*

Implementing this model at Bank A can deliver multiple strategic benefits:

1. Early Risk Detection

- Identifying customers with high likelihood of default before they miss payments enables the bank to take proactive actions (e.g., calls, restructuring, soft collections).

2. Reduced Credit Losses

- Minimizing false negatives (missed defaulters) helps the bank avoid non-performing assets (NPAs) and associated write-offs.

3. Personalized Risk Communication

- The bank can design targeted outreach campaigns or offer custom repayment plans for at-risk customers based on their SHAP-based explanations.

4. Better Credit Limit Management

- Customers predicted as high risk can have their credit limits frozen, reduced, or more closely monitored.

5. Improved Portfolio Quality & Compliance

- Helps the bank stay within risk-weighted asset thresholds set by regulators and maintain a healthier loan book.

8.*Summary of Findings and Key Learnings*

This project addressed the challenge of predicting credit card defaults using a structured machine learning pipeline. Key takeaways include:

- **Class imbalance** (only ~22% defaulters) was handled effectively using **SMOTE**, improving model learning.
- Financially driven **feature engineering** (e.g., dues, utilization ratio, payment trend) improved both performance and interpretability.
- **F2 score** was chosen to prioritize recall, reflecting the business need to catch as many defaulters as possible. Threshold tuning (optimal at 0.31) enhanced this.
- **Random Forest** emerged as the best model, achieving a strong F2 score (~0.89) and ROC AUC (~0.93).
- The model has clear business value: enabling early risk detection, reducing credit losses, and supporting regulatory compliance.