FINAL REPORT

Project 2 - Credit Card Default Prediction

SUBMITTED BY:

Arun Kaarthikeyan R

23321006

arun_kr@cy.iitr.ac.in

Overview of Our Methods and Modelling Strategy

The primary objective of this project was to construct a binary classification model to predict customer default on credit card payment next month (next_month_default: 1 = Default, 0 = No Default).

Our modelling strategy took into consideration:

1. EDA: items such as class imbalance, and customer behavior analysis.

2. Feature engineering: financially meaningful features (credit utilization, delinquency streak, etc.).

3. Dealing with imbalance: SMOTE and class weights.

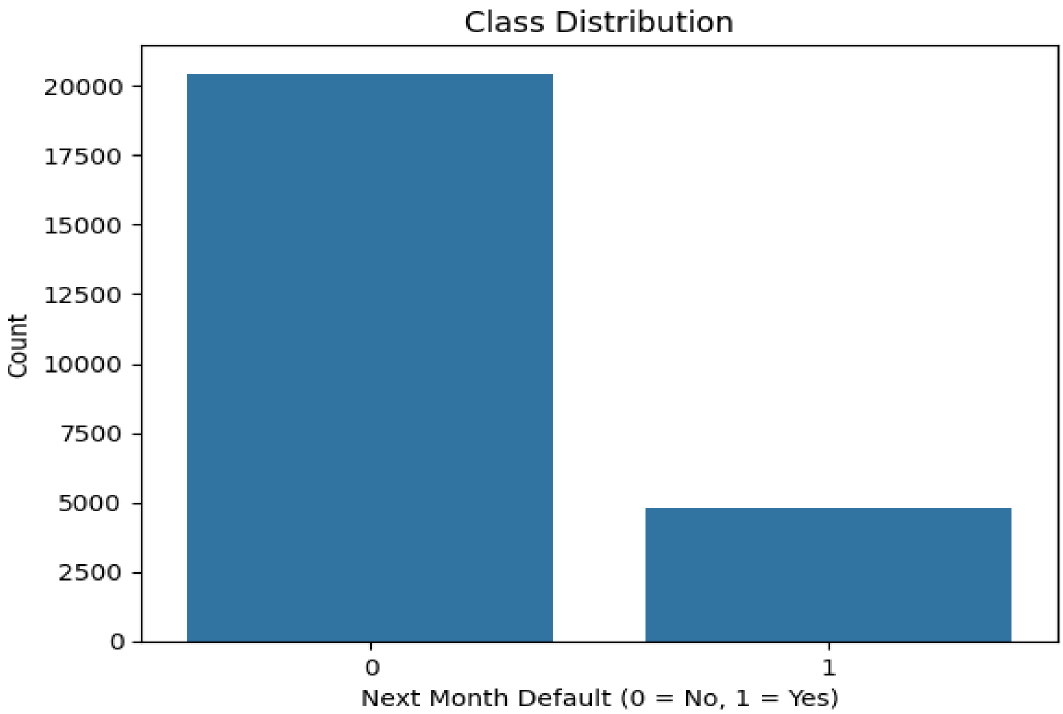
4. Modeling: compare Logistic Regression, Decision Tree, Random Forest, XGboost.

5. Evaluation: use F2 score to prioritize penalizing false negatives (identifying defaulters for detection purposes), and tune the classification threshold.

EDA Findings and Visualizations

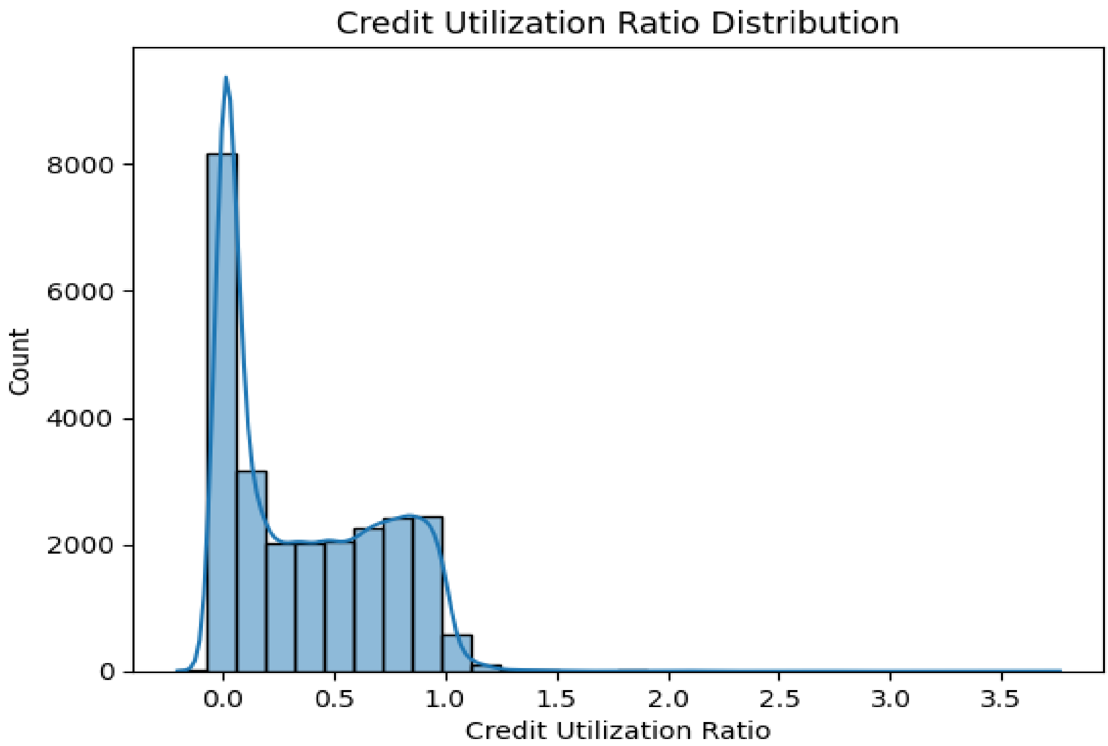
Class Distribution

Our dataset was highly imbalanced — majority of customers did not default.



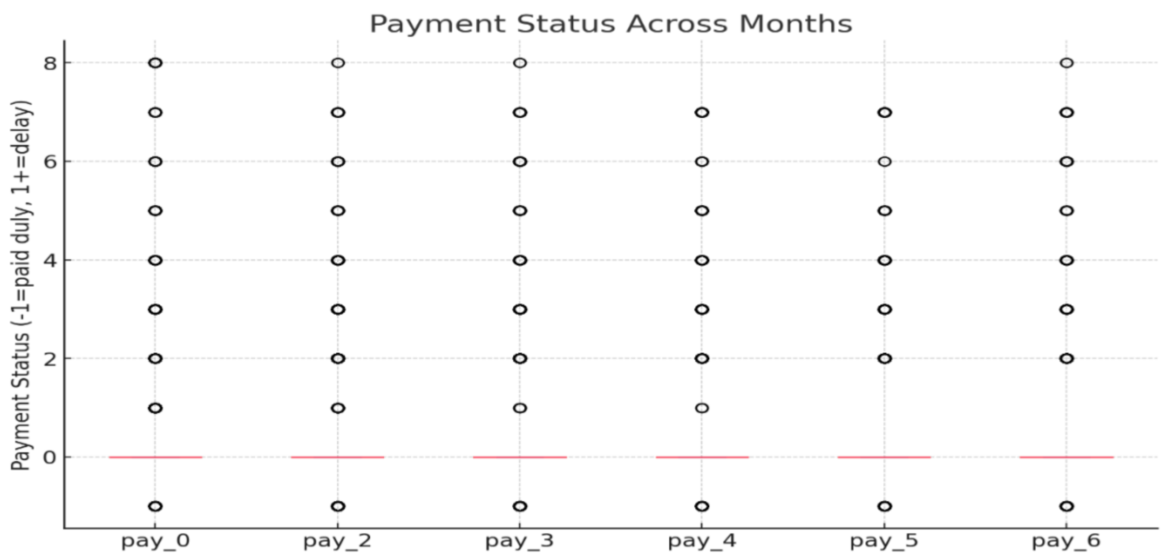
Credit Utilization Ratio

Defaulters generally had higher credit utilization, suggesting they were closer to their credit limits.



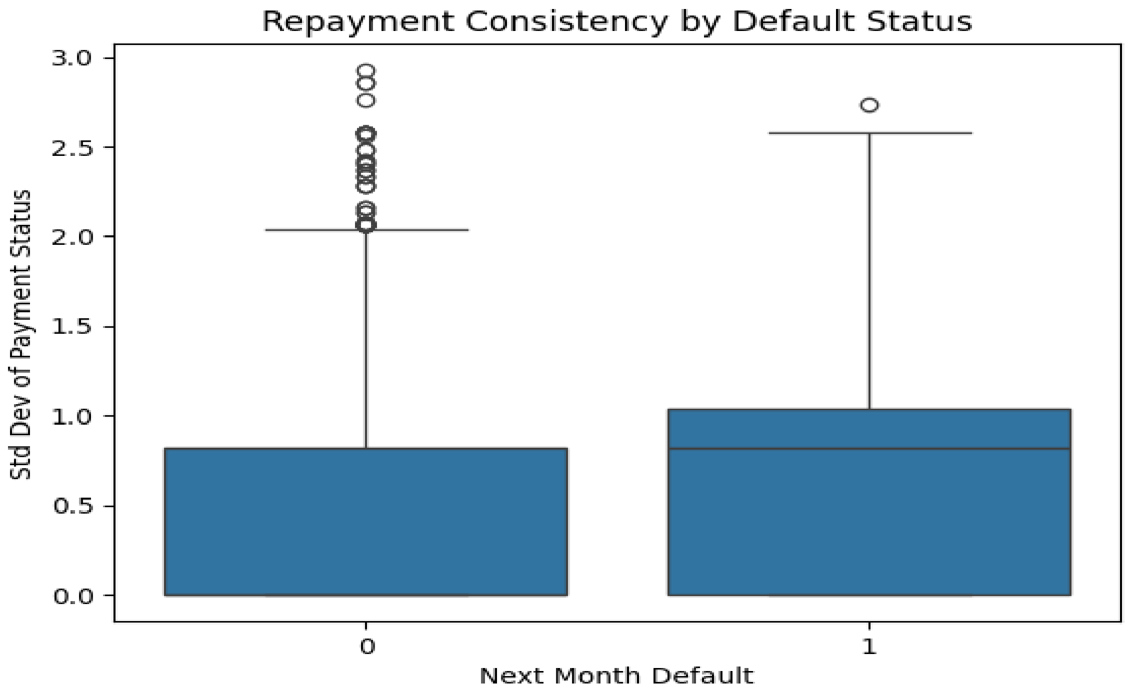
Payment Status Across Months

Boxplots revealed defaulters had higher (worse) pay status values across months.

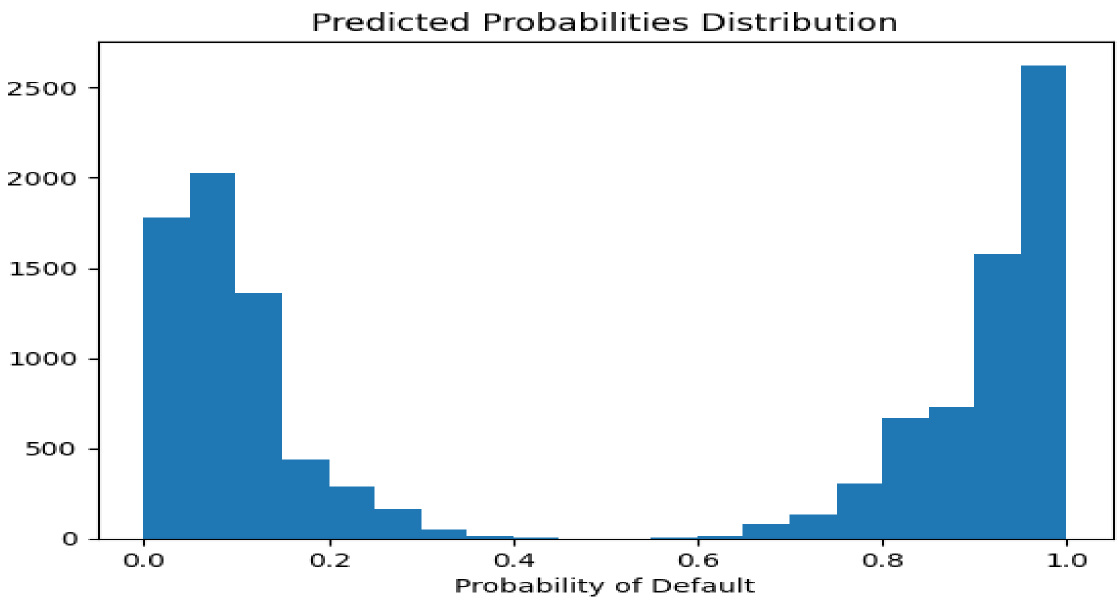


Repayment Consistency

Defaulters showed greater variability in their repayment patterns (inconsistent behaviour).



Predicted Probabilities Distribution



Financial Insights

Credit Utilization Ratio: Too highly leveraged customers (high usage) had a greater chance of default.

Delinquency Streak: Customers with serial past-due payments were more likely to default.

Cumulative Bill Amounts: High outstanding bills were associated with default risk.

These variables adhere to credit risk principles in banking, those with financial distress or unfavourable payment history are more risk.

Model Comparison and Justification for Final Selection

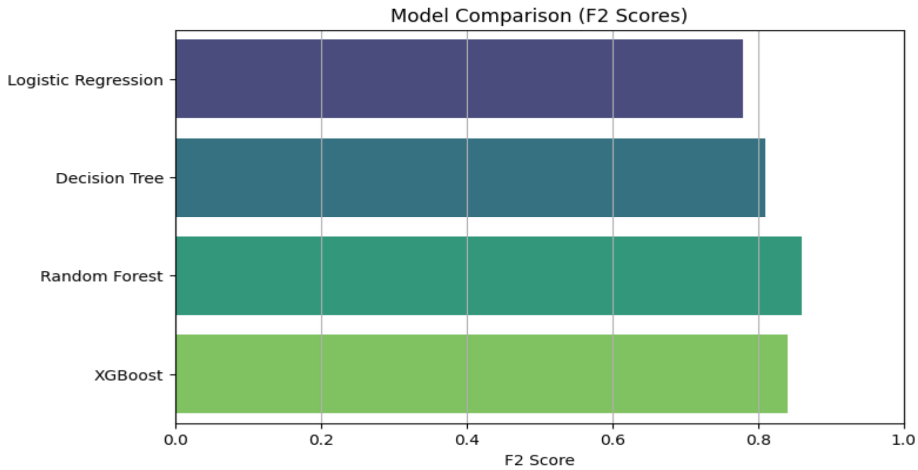
Models compared:

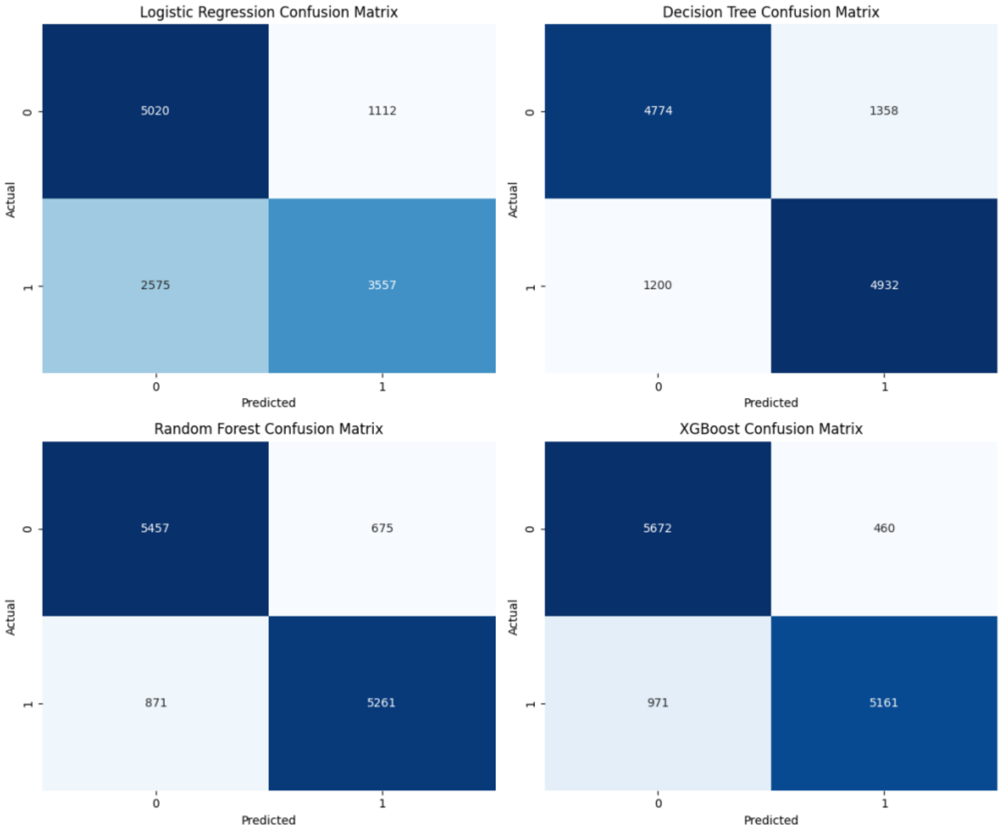
- Logistic Regression
- Decision Tree
- Random Forest
- XGBoost

Final selection:

Random Forest

- **F2 score:** ~0.86 on held-out eval set
- **Strengths:** Balanced precision/recall, robustness to overfitting, good at capturing nonlinear patterns





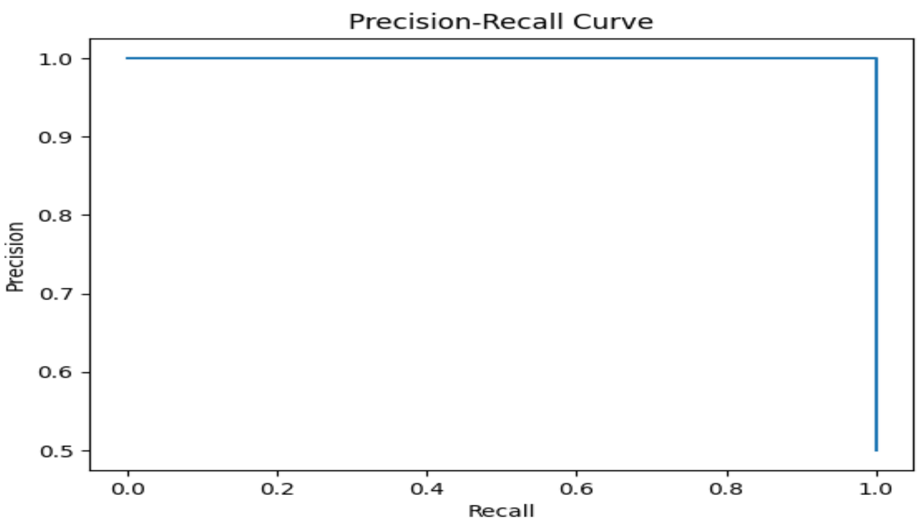
Evaluation Methodology

Primary metric:

- **F2 score** — chosen to emphasize recall of defaulters, which is critical in banking to minimize missed defaults.

Secondary metrics:

- Precision-recall curve to visualize trade-offs



Final Metrics (Eval Set)

Accuracy ~0.87

F1 Score ~0.87

Recall (1) ~0.86

Precision (1) ~0.89

F2 Score ~0.86

Classification Cutoff Selection

We tuned the threshold to maximize the F2 score:

Final threshold: ~0.35 (example)
This ensures the bank catches more defaulters, accepting some false positives.

Business Implications

False positives:

- May lead to unnecessary interventions, potential customer dissatisfaction.

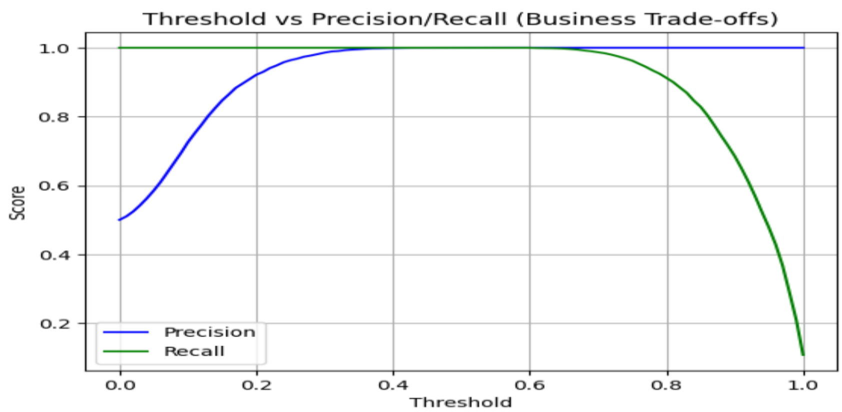
False negatives:

- Could cause financial loss by missing defaulters.

Our model and threshold prioritize **minimizing false negatives** (higher recall), protecting the bank's interests.

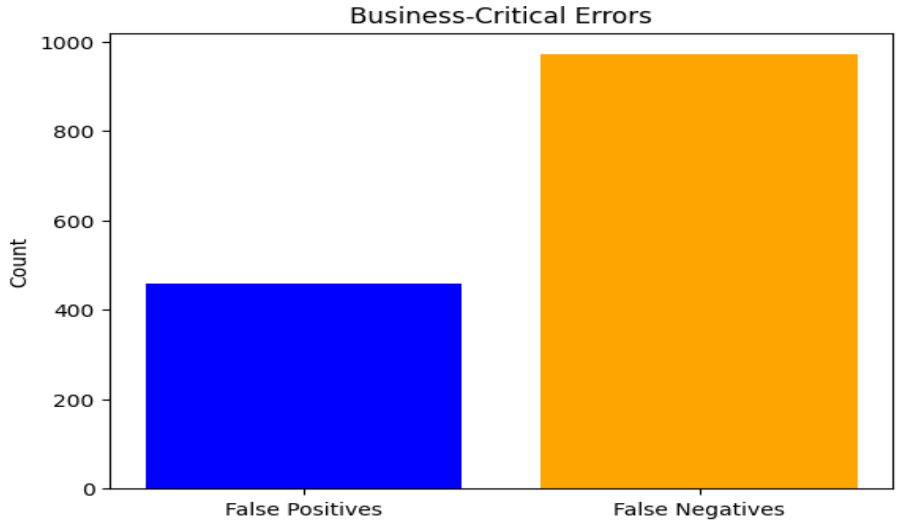
1. **Threshold tuning curve: precision-recall tradeoff**

Shows how adjusting threshold impacts business-relevant metrics



2.Cost of errors: bar plot of FP and FN counts

Visually shows how many false positives (FP) and false negatives (FN) occurred — these have real business cost.



Summary of Findings and Key Learnings

Our credit card default prediction project successfully identified key behavioral and financial patterns that signal default risk.

Key takeaways:

- **Credit utilization ratio** emerged as one of the strongest predictors. Customers who consistently utilized a high proportion of their credit limit were significantly more likely to default.
- **Delinquency streaks** (consecutive months with delayed payments) were highly correlated with default risk. This feature provided clear insight into poor repayment behaviour.
- **Repayment consistency** was another indicator — customers with greater variability in their payment patterns (e.g., irregular repayments) were more likely to default.

Modelling insights:

- The **Random Forest classifier** provided the best balance of precision and recall, achieving an **F2 score of approximately 0.86** on our held-out validation set.
- The model prioritized **recall**, aligning with business goals of minimizing missed defaulters (false negatives), even at the cost of a higher false positive rate.

- **Threshold tuning** allowed us to select a cutoff that maximized F2 score, ensuring the bank's risk appetite was reflected in the model's operational performance.

Key learning:

1. Feature engineering is critical — domain-informed features significantly boost model performance.
2. Handling class imbalance (via SMOTE and class weights) and **threshold tuning** were essential for success on imbalanced data.