# Credit Risk Modeling and Scoring System

In this project, we designed a full end-to-end credit risk assessment pipeline using machine learning. The aim is to predict the creditworthiness of individuals and generate an interpretable credit score. This system combines advanced data preprocessing, model ensemble techniques, feature engineering, and scorecard generation to deliver reliable predictions and human-understandable results.

# Data Loading and Cleaning

The raw dataset was loaded from a CSV file. Given the nature of financial data, extensive cleaning was necessary. We addressed several common data quality issues:

- Replaced various formats of missing or corrupted values (e.g., 'NA', 'unknown', special characters).

- Cleaned numeric columns that contained non-numeric characters.

- Converted Credit\_History\_Age into a unified numeric format (months).

- Normalized text data (e.g., Payment\_Behaviour, Credit\_Mix) to consistent lowercase representations.

- Dropped columns with more than 30% missing values.

- Applied smart imputation strategies: categorical values were filled with 'missing', while numeric ones were filled with the median along with a new missing indicator column.

# Target Variable Engineering

Instead of using an existing label, we created a custom target (target) based on multiple behavioral factors:

- Payment behavior patterns and credit mix were scored using domain-informed mappings.

- The number of delayed payments contributed non-linearly to the score.

- A binary label was generated by comparing the composite score against the 40th percentile—ensuring class balance.

- An explanatory text column was added to improve interpretability.

# Feature Engineering

New features were created to enhance model performance:

- Debt-to-Income Ratio, Utilization Ratio, Delay Ratio, and EMI-to-Income to quantify financial behavior.

- Age and credit history were binned into meaningful categories.

- Combined features like Credit\_Products\_Count gave insight into an individual’s credit activity.

# Data Preprocessing

- Categorical features were label encoded.

- Numerical features were scaled using Min-Max normalization.

- The data was split using stratified sampling to maintain class proportions.

# Model Training

Three diverse models were trained individually:

- Logistic Regression – a linear baseline with class balancing.

- Random Forest – a non-linear tree-based model with interpretability.

- Support Vector Machine – wrapped in a calibration layer for probability outputs.

A soft-voting classifier was then created to combine their strengths.

# Model Evaluation

Each model was evaluated using:

- Confusion Matrix – to visualize true vs. predicted outcomes.

- ROC AUC Score – to measure classification performance.

- Precision-Recall Curve – to assess trade-offs in prediction quality.

Performance Highlights:

- All models showed solid AUC scores, with the ensemble voting model outperforming individual ones.

- The Random Forest and Logistic Regression models provided useful insights into feature importance.

# Interpretability and Feature Importance

- Logistic Regression coefficients revealed linear relationships with the target.

- Random Forest importance highlighted key non-linear influencers.

- SHAP values gave a global overview of top contributing features to model decisions, ensuring transparency and trust in the system.

# Credit Scorecard Generation

We created an interpretable scorecard using the Logistic Regression model:

- Used industry-standard metrics like PDO (Points to Double Odds) and base score scaling.

- Mapped each feature’s impact into score points, then summed to a base score to produce a final numeric score per user.

- Features were ranked by their coefficient magnitudes to highlight their relative importance.

# Credit Score Calculation and Rating

Each customer’s score was computed and mapped into credit categories:

- Excellent (800+)

- Very Good (740–799)

- Good (670–739)

- Fair (580–669)

- Poor (<580)

This provides a straightforward, industry-aligned rating system.

A violin plot was generated to visualize the distribution of scores across these categories, clearly highlighting separability.

# Conclusion

This system showcases a powerful and interpretable approach to credit risk modeling:

- It integrates strong data cleaning, smart feature engineering, diverse modeling, and explainability.

- The voting classifier ensures robustness while the scorecard enhances trust with stakeholders.

This project can be extended with:

- More temporal features from credit history.

- Model monitoring pipelines for production.

- Integration with a real-time dashboard or app.

8. Model Interpretability and Explainability  
  
Understanding how a credit risk scoring model makes decisions is crucial for gaining trust from stakeholders and ensuring regulatory compliance. Here are several techniques and strategies used in this project to enhance interpretability:  
  
 8.1 Feature Importance  
  
We examined the most influential features using various methods depending on the model:  
- For tree-based models like Random Forest and XGBoost, we used built-in feature importance scores based on Gini impurity or information gain.  
- For logistic regression, we interpreted the coefficients to understand how features positively or negatively influence the prediction.  
  
This analysis helps stakeholders understand which attributes most affect the credit risk prediction, such as income, loan amount, and credit history.  
  
8.2 SHAP (SHapley Additive exPlanations)  
  
SHAP values provide a unified framework to interpret any machine learning model by quantifying the contribution of each feature to a particular prediction. In our project, SHAP plots (summary, force, and dependence plots) highlighted how individual features influenced the model’s output for each applicant.  
  
 8.3 Partial Dependence Plots (PDP)  
  
PDPs were used to visualize the average predicted response as a function of one or two features. These plots helped to understand how changes in a particular feature (like loan amount or income) impact the model’s risk prediction, holding all other features constant.  
8.4 Global vs Local Interpretability  
  
- \*\*Global interpretability\*\* provides an overall view of feature importance and behavior across the entire dataset.  
- \*\*Local interpretability\*\* focuses on explaining individual predictions. We used tools like LIME and SHAP to break down single predictions and highlight which factors led to a loan being classified as high or low risk.  
8.5 Transparency and Fairness  
  
To ensure the model does not exhibit bias, we tested for fairness across different demographic groups. We also documented the model assumptions and limitations clearly, especially in cases where decisions may impact users’ financial opportunities.  
  
These explainability techniques ensure that the credit scoring model is not only accurate but also transparent, accountable, and trustworthy.