Credit Score Classification Analysis Report

Executive Summary

This report presents an analysis of credit score classification using machine learning techniques. The study aims to forecast credit scores based on historical financial data, segment creditworthiness into distinct categories (Bad, Average, and Good), and provide actionable insights to stakeholders such as lenders and financial institutions. Using extensive exploratory data analysis (EDA) and predictive modeling, key factors influencing credit scores were identified, and the best classification model was selected based on performance metrics.

Business Understanding: Problem Statement

The objective of this project is to develop a predictive model that can accurately classify individuals' credit scores based on various financial indicators. Financial institutions and lenders require robust risk assessment models to make informed credit decisions. By leveraging machine learning, we aim to enhance risk evaluation, reduce defaults, and support data-driven lending strategies.

Background

Credit scores are a crucial metric in determining an individual's financial reliability. Various factors, including payment history, outstanding debt, and credit mix, influence credit scores. Traditional scoring methods often lack transparency and adaptability to evolving financial patterns. This study employs machine learning techniques to address these limitations by identifying the most impactful features affecting credit scores and improving prediction accuracy.

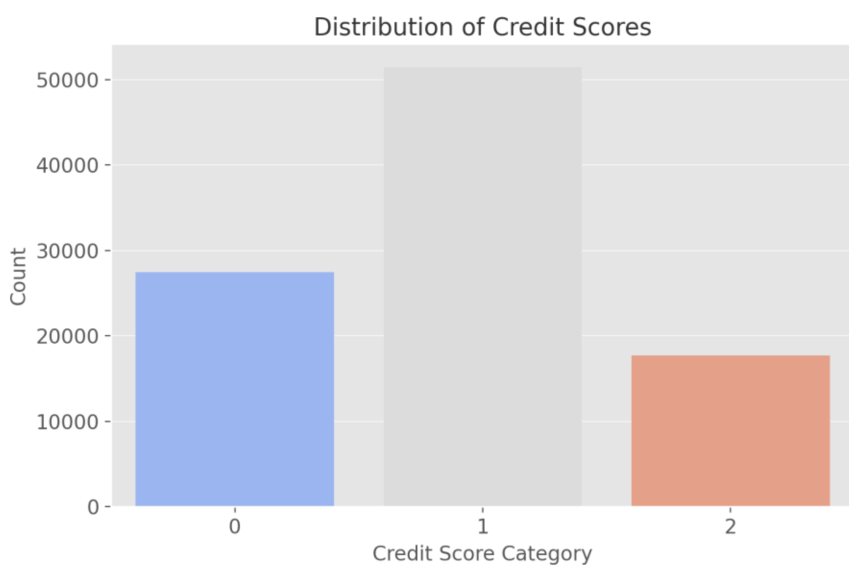
Data Preparation & Analysis

The following steps were performed for data cleaning and transformation:

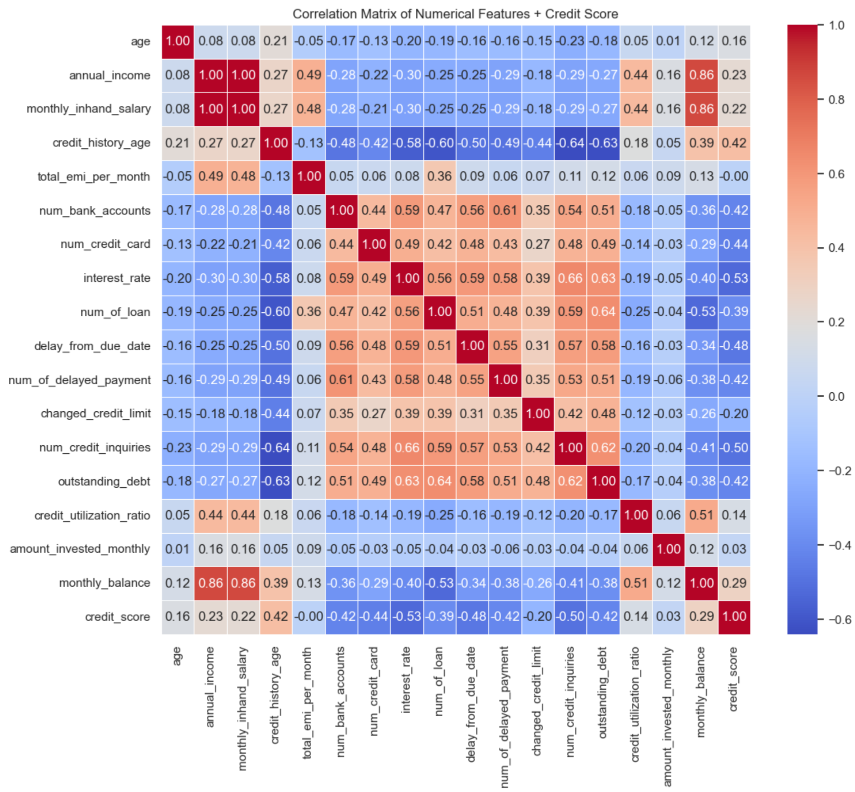
- **Handling Missing Values –** Missing data was imputed using the average for numerical columns and mode for categorical columns.
- **Aggregation –** Eight rows of customer data were aggregated into single records for better model performance.

Exploratory Data Analysis

- Most customers fall under Category 1 (Average Credit Score).
- Fewer customers have Bad (0) or Good (2) Credit Scores, indicating potential class imbalance.



- Correlation analysis revealed key relationships between financial factors and credit scores.



Dropped Variables from the Correlation Matrix

- **Monthly In-Hand Salary:** Removed due to its perfect correlation (1.0) with annual income, indicating redundancy. Keeping both could cause multicollinearity issues, reducing model stability.
- **Amount Invested Monthly:** Excluded due to its weak correlation (0.03) with credit score, showing minimal predictive value. Dropping this variable streamlines the dataset for better model focus.

Model Selection & Evaluation

We do the following steps prior to fitting our model & hyperparameter tuning.

- **Feature Scaling** – Normalization of numerical features using Standard Scaler to ensure model stability.
- **One-hot Encoding** – Transformation of categorical variables into numerical representations.
- **Feature Selection** – Dropping irrelevant columns to reduce noise.
- **Stratified Train-Test Split –** In order to account for imbalanced target.

We use 2 metrics to ensure we properly predict the ordinal nature of our imbalanced target

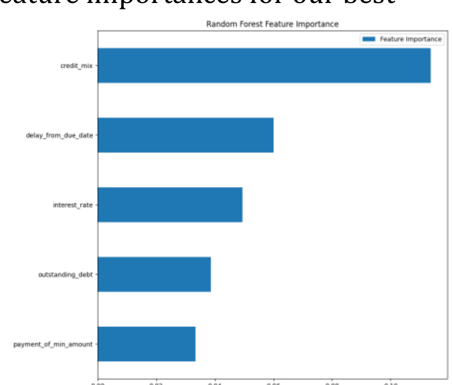
- **Weighted Cohen's Kappa**
- **F1 Score**

	F1 Score	Weighted Kappa
Logistic Regression	0.709565	0.597992
Random Forest	0.767050	0.683618
XGBoost	0.758447	0.670593
CatBoost	0.762795	0.661908

After hyperparameter tuning, the best model was selected based on a balance of F1-score and Weighted Kappa Score via cross-validation.

Model Inference

We use **Permutation feature importance** to measure the feature importances for our best model. Essentially it measures the drop in model performance by randomly shuffling values for one column. This gives us the following top 5 features (highest to lowest importance):



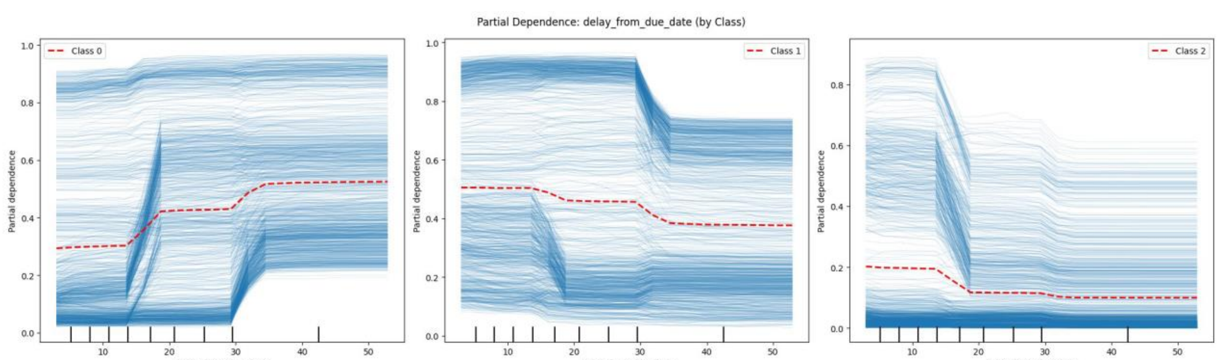
- *Credit_mix (Bad, Standard, Good)*
- *Delay_from_due_date*

- *Interest_rate*
- *Outstanding_debt*
- *Payment_of_min_amount (Yes, No)*

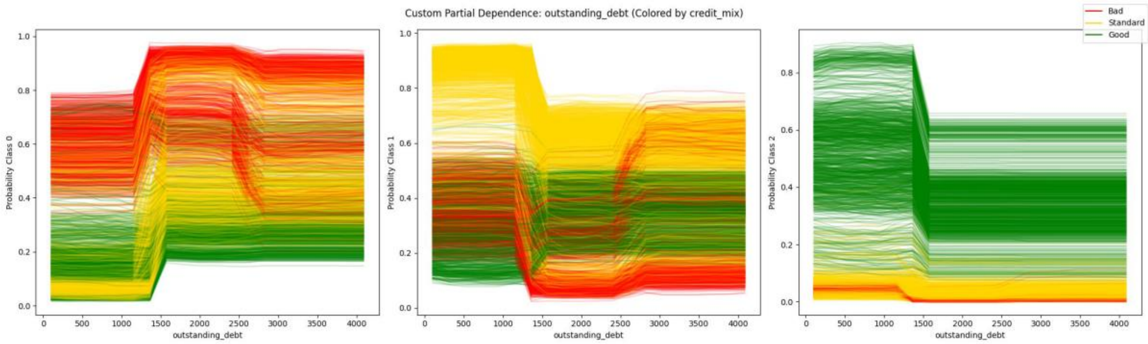
Next, we use Partial Dependence (PD) & Individual Conditional Expectation (ICE) plots for our top 5 features to better understand their impact.

- PDP shows the average marginal effect of a feature (or features) on the predicted outcome by marginalizing over other variables.
- ICE shows the response of a model for individual instances by varying the feature of interest while keeping other features fixed.

An example of the graphs is shown below. PDP is essentially the red line showing the average effect of a feature over the probability of being in one of the 3 classes (Credit Score). The blue lines are the ICE plot essentially showing the effect for every individual. The main takeaway would be that more the delay, more probability of having lower credit score, with certain delays like 15, 30 days resulting in more jumps.



We also developed a custom ICE plot to see interactions between categorical and numerical variables. Again, the main takeaway is that having a Good Credit Mix is critical, however large outstanding debt weighs down probability of good credit score.



A few more of these interpretation plots are also shown in the appendix for more interpretation of the model.

Conclusion

Our machine learning-based credit score prediction model effectively categorizes customers into Bad, Average, and Good credit classifications with high predictive accuracy. The Random Forest algorithm delivered superior performance by optimally balancing F1-Score and Weighted Kappa metrics.

Key predictive factors identified include credit mix, payment timeliness, interest rates, outstanding debt, and minimum payment behavior. These insights provide actionable value to financial institutions for risk assessment and to consumers seeking credit improvement.

This model establishes a solid foundation for ongoing refinement as financial markets evolve. Addressing the identified limitations will further enhance accuracy, fairness, and stakeholder value throughout the financial ecosystem.