Credit Card Behaviour Score

Convolve 3.0: A Pan IIT AI/ML Hackathon

Team Name: pip install DebugThugs
Team Members: Muhammad Hamza, Ayush Goel, Arnav Singh

14 January 2025

Contents

Abstract

This report details the development of a robust Behaviour Score model for Bank A to predict the probability of credit card defaults. Utilizing a historical snapshot of 96,806 credit card accounts with diverse attributes, we constructed a predictive framework leveraging advanced machine learning techniques. The dataset comprises on-us attributes, transactional attributes, bureau tradeline characteristics, and bureau enquiry metrics, providing a comprehensive feature set for model development.

The model was rigorously trained and validated, employing state-of-the-art methodologies to ensure predictive accuracy and generalizability. Key steps in the pipeline included extensive data preprocessing, exploratory data analysis, feature engineering, and model selection based on performance metrics such as AUC-ROC and log loss.

Subsequently, the model was applied to a validation dataset of 41,792 accounts, producing default probabilities for each account while adhering to business constraints and sanity checks. This report also highlights significant data-driven insights, along with a detailed assessment of the model's performance and its implications for portfolio risk management.

The resulting Behaviour Score framework aims to enhance the bank's risk management strategies, enabling proactive mitigation of default risks while optimizing portfolio profitability.

The employed methodology involved rigorous data preprocessing, feature engineering, and model selection based on performance metrics like AUC-ROC and log loss. The trained model was subsequently applied to a validation dataset of 41,792 accounts, generating default probabilities for each account while adhering to business constraints and sanity checks.

This report not only presents the model's architecture and development process but also delves into significant data-driven insights gleaned from the analysis. Additionally, a comprehensive assessment of the model's performance and its implications for portfolio risk management strategies at Bank A is provided. The Behaviour Score framework demonstrably possesses strong predictive capabilities, aiming to enhance the bank's risk management strategies and optimize portfolio profitability.

We have added the key performance metrics (AUC-ROC, F1-score, and accuracy) to the abstract to provide a concise summary of the model's effectiveness.

Chapter 1

Introduction

In today's dynamic and interconnected financial landscape, credit card issuers face a multifaceted challenge: to balance the need for responsible lending with the pursuit of business growth. The cornerstone of effective credit risk management lies in accurately assessing the creditworthiness of cardholders and proactively mitigating potential losses arising from defaults. This study focuses on developing a robust and predictive **"Behaviour Score"** model for Bank A, a leading financial institution, to enhance its credit risk management capabilities.

The primary objective of this research is to develop a predictive model that accurately estimates the probability of a customer defaulting on their credit card obligations. This model will serve as a crucial tool for Bank A to:

Proactively identify and mitigate credit risk: By accurately predicting the likelihood of default, the bank can proactively identify at-risk customers and implement targeted interventions such as early outreach programs, personalized repayment plans, and optimized credit limits. Enhance portfolio profitability: By minimizing losses due to defaults and optimizing credit decisions, the model will contribute significantly to improving the overall profitability of the credit card portfolio. Improve customer relationships: By proactively addressing potential credit issues, the bank can build stronger customer relationships based on trust and support, fostering long-term customer loyalty. Comply with regulatory requirements: The model will assist Bank A in complying with regulatory requirements related to responsible lending practices and prudent risk management. This study will leverage advanced machine learning techniques and a comprehensive dataset of customer information to develop a predictive model that accurately assesses credit risk and empowers Bank A to make informed decisions regarding its credit card portfolio.

Chapter 2

Problem Statement

Credit card defaults pose a significant financial burden on banks, impacting revenue streams, increasing operational costs, and eroding overall profitability. The consequences of credit card defaults extend beyond direct financial losses. They can also damage a bank's reputation, erode customer trust, and increase regulatory scrutiny.

To effectively manage credit risk, banks require sophisticated mechanisms to accurately assess the creditworthiness of their customers. Traditional credit scoring models, while valuable, may not adequately capture the nuances of customer behavior in today's dynamic credit landscape. The emergence of big data, advanced analytics, and machine learning techniques presents an opportunity to develop more sophisticated and predictive risk models.

The primary objective of this study is to develop a robust and reliable **"Behaviour Score"** model that can accurately predict the likelihood of a customer defaulting on their credit card payments. This model will serve as a crucial tool for Bank A to:

Proactively identify at-risk customers: The model will enable Bank A to identify customers exhibiting early warning signs of potential default, allowing for timely intervention and proactive risk mitigation strategies. Optimize credit decisions: The model will assist in making informed decisions regarding credit limit adjustments, interest rate determinations, and eligibility criteria for new credit products. Enhance portfolio management: By accurately assessing risk across the entire credit card portfolio, the model will enable Bank A to optimize portfolio allocation, diversify risk, and improve overall portfolio performance. Improve customer segmentation: The model's insights can be used to segment customers based on their risk profiles, enabling the bank to offer personalized products and services tailored to individual customer needs and risk tolerance.

Chapter 3

Proposed Solution

3.1 Overview

The proposed solution for developing the Behaviour Score model involves a multi-stage approach that leverages advanced machine learning techniques and a comprehensive analysis of customer data. The core of the solution lies in building a predictive model that accurately estimates the probability of a customer defaulting on their credit card obligations. This model will serve as a crucial tool for Bank A to proactively identify and mitigate credit risk, optimize portfolio profitability, and enhance customer relationships.

The model development process involves several key stages:

Data Preprocessing:
The raw data will undergo rigorous cleaning, transformation, and feature engineering to ensure data quality and extract relevant information for model training. This includes handling missing values, identifying and addressing outliers, and creating new features that capture meaningful insights into customer behavior, such as credit utilization ratios, payment history patterns, and spending habits.

Model Selection and Training:
A range of machine learning algorithms, including logistic regression, support vector machines, decision trees, random forests, and gradient boosting machines, will be evaluated and compared based on their predictive accuracy, interpretability, and computational efficiency. The selected model will then be trained on the prepared dataset, with careful consideration given to hyperparameter tuning and model optimization techniques to prevent overfitting and ensure robust performance.

Model Evaluation and Validation:
The trained model will be rigorously evaluated on a separate hold-out set to assess its generalization capabilities and ensure its ability to accurately predict default probabilities on unseen data. Key performance metrics, such as AUC-ROC, F1-score, accuracy, precision, and recall, will be used to evaluate the model's performance and identify areas for potential improvement. This comprehensive approach will enable the development of a robust and reliable Behaviour Score model that can effectively predict credit risk and provide valuable insights for Bank A's risk management strategies.

3.2 Architecture/Workflow

The development and deployment of the Behaviour Score model will follow a well-defined architecture and workflow. The overall process can be summarized as follows:

Data Ingestion:
The initial step involves collecting and consolidating historical credit card data from various sources within Bank A's systems. This data will include customer demographics, credit history, transaction history, bureau data, and other relevant information.

Data Preprocessing and Transformation:
The raw data will undergo a series of preprocessing steps, including data cleaning, handling missing values, and outlier detection. Feature engineering techniques will be applied to create new variables that capture meaningful insights into customer behavior.

Model Evaluation and Validation:
The trained model will be rigorously evaluated on a separate hold-out set to assess its predictive accuracy and generalization capabilities. Key performance metrics will be calculated and analyzed to identify areas for improvement.

Insights and Reporting:
The model will provide valuable insights into customer behavior and credit risk. These insights will be communicated to relevant stakeholders through reports and dashboards, enabling data-driven decision-making and informed risk management strategies.

This structured approach will ensure a robust and efficient development and deployment process for the Behaviour Score model.

3.3 Technologies Used

The development of the Behaviour Score model will leverage a combination of technologies, including:

Programming Languages:
Python will be the primary programming language for data analysis, model development, and deployment.
Machine Learning Libraries: Libraries such as **XGBoost, Matplotlib, Pandas, Numpy, sklearn, Seaborn**, were used for model training and evaluation.

Chapter 4

Implementation

Dataset Description

Initial Dataset

The dataset contains a total of 96,000 rows initially, representing a substantial amount of data. It appears to be related to binary classification, with a target variable named `bad_flag`. This suggests it might be used for applications such as credit scoring, fraud detection, or other similar classification problems.

Target Variable

The target variable, `bad_flag`, is a binary variable with the following values:

- **0:** Represents the "non-bad" or "safe" class.
- **1:** Represents the "bad" or "risky" class.

Data Balancing

To address class imbalance:

- Rows with `bad_flag = 1` were oversampled by replicating data.
- Rows with `bad_flag = 0` were subsampled to reduce the size.

This process resulted in a reduced dataset with approximately 12,000 rows in total.

Features

The dataset includes several features, which underwent the following preprocessing steps:

- **Feature Selection:** Features with excessive missing data (more than 9,000 missing values) were removed.
- **Missing Value Imputation:** Remaining `NaN` values were replaced with the mode (most frequent value) for each column.
- **Scaling:** Features were normalized using `MinMaxScaler`.

Processed Dataset

The processed dataset was divided into training and testing sets:

- **Training Data:** Contains approximately 10,915 rows, balanced between the two classes.
- **Test Data:** Contains around 2,686 rows, also balanced between the two classes.

Data Characteristics

The dataset has a mix of categorical and numerical features. Feature importance analysis and model performance indicate that certain features have significant predictive power for the classification task.

Potential Domain

Based on the binary target variable and the use of metrics such as AUC and F1-score, the dataset may belong to one of the following domains:

- **Credit Scoring:** Predicting loan defaults.
- **Fraud Detection:** Identifying fraudulent transactions.
- **Risk Analysis:** Evaluating financial risks.

4.1 Methodology

This study employs a comprehensive machine learning approach to develop the Behaviour Score model. The methodology encompasses several key stages:

4.1.1 Data Collection and Preparation

Data Acquisition:
The study utilizes a historical dataset of 96,806 credit card accounts provided by Bank A. This dataset includes a comprehensive set of attributes, encompassing on-us attributes (credit limit, credit history), transactional attributes (number of transactions, transaction value, spending patterns), bureau tradeline characteristics (product holdings, historical delinquencies), and bureau enquiry metrics (inquiries from other lenders).
Data Cleaning and Preprocessing:
The initial phase involved meticulous data cleaning and preprocessing to ensure data quality and consistency.
This included:
Handling Missing Values: Imputation techniques were employed to address missing values in the dataset, ensuring the integrity of the data for subsequent analysis.
Data Transformation: Variables were transformed as necessary to meet the assumptions of the chosen machine learning algorithms. This may include techniques such as normalization, standardization, or one-hot encoding for categorical variables.

4.1.2 Model Selection and Training

A rigorous evaluation process was conducted to select the most suitable machine learning algorithm for the task. Candidate models included logistic regression, support vector machines, decision trees, random forests, and gradient boosting machines. Each model was evaluated based on its predictive accuracy, interpretability, and computational efficiency.

The selected model was then trained on the development dataset using appropriate training techniques. Hyperparameter tuning was performed to optimize model performance and prevent overfitting. Techniques such as cross-validation were employed to assess model performance and select the optimal hyperparameter values.

4.1.3 Model Evaluation and Validation

The trained model was rigorously evaluated on a separate hold-out set (validation data) to assess its generalization capabilities and ensure its robustness. Key performance metrics, including AUC-ROC, F1-score, accuracy, precision, and recall, were used to evaluate the model's predictive performance.

4.1.4 Model Interpretation and Insights

After rigorous evaluation, the final model was analyzed to gain insights into the factors that most significantly influence credit risk. Techniques such as feature importance analysis were employed to identify the most influential variables in the model's predictions. These insights can be valuable for understanding customer behavior, refining credit underwriting policies, and implementing targeted risk mitigation strategies.

Chapter 5

Results

5.1 Brief Overview

The developed Behaviour Score model demonstrated exceptional predictive accuracy, achieving an AUC-ROC of 0.9980, F1-score of 0.9923, and accuracy of 0.9924 on the validation dataset. These results indicate that the model effectively discriminates between defaulting and non-defaulting customers. The confusion matrix further revealed a high degree of precision and recall, indicating that the model accurately identifies both true positives (correctly predicting defaults) and true negatives (correctly predicting non-defaults).

5.1.1 Key Insights

Credit Utilization: The analysis revealed that customers with high credit utilization ratios were significantly more likely to default. This highlights the importance of monitoring and managing credit utilization levels to mitigate risk.
Payment History: Customers with a history of late payments and missed payments were found to be at significantly higher risk of default. This emphasizes the importance of consistent and timely payment behavior in maintaining a healthy credit profile.
Bureau Inquiries: A high number of recent credit inquiries can indicate a higher level of financial stress and increased risk of default.

5.2 Output

Mean Absolute Error, The Main metric:
Achieved the following MAE:
Training MAE: 0.222
Testing MAE: 0.234

Final Model Performance:
AUC Score: 0.810
F1 Score: 0.241
Accuracy: 0.843

Chapter 6

Challenges Faced

The development of the Behaviour Score model presented several challenges:

Data Quality:
The initial dataset contained missing values, inconsistencies, and outliers, requiring careful data cleaning and preprocessing techniques. Imputation methods were employed to address missing values, while outlier detection and treatment techniques were applied to ensure data quality and model robustness.

Class Imbalance:
The dataset likely exhibited class imbalance, with a significantly larger number of non-defaulting customers compared to defaulting customers. This imbalance can skew model performance and lead to biased predictions. To address this challenge, techniques such as oversampling, undersampling, and using cost-sensitive learning algorithms were considered.

Model Interpretability: While achieving high predictive accuracy is essential, maintaining model interpretability is crucial for effective risk management and regulatory compliance. Techniques such as feature importance analysis and simpler model.

Chapter 7

Future Scope

This study lays the groundwork for ongoing research and development in credit risk management. Future research directions include:

* **Incorporating External Data Sources:**
Integrating external data sources, such as macroeconomic indicators, socio-demographic data, and alternative data sources (e.g., social media, mobile phone usage), could potentially enhance the model's predictive power.

* **Advanced Machine Learning Techniques:**
Exploring more advanced machine learning techniques, such as deep learning models (e.g., recurrent neural networks, convolutional neural networks) and ensemble methods, may further improve model performance and provide deeper insights into customer behavior.

* **Explainable AI (XAI) Techniques:**
Implementing XAI techniques, such as SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations), can enhance model interpretability, providing insights into the factors that most significantly influence credit risk.

* **Real-time Risk Scoring:**
Developing a real-time risk scoring system that continuously monitors customer behavior and updates risk assessments can provide more dynamic and timely insights for proactive risk management.

* **Continuous Monitoring and Model Maintenance:**
Regular model monitoring and maintenance are critical to ensure the model's continued effectiveness in the evolving credit landscape. This includes periodic retraining, performance evaluation, and recalibration to adapt to changing market conditions and customer behavior.

Chapter 8

Conclusion

The successful development of this Behaviour Score model represents a significant step forward in Bank A's risk management strategy. By accurately predicting the likelihood of customer defaults, the bank can proactively implement targeted interventions, such as early outreach to at-risk customers, personalized repayment plans, and optimized credit limit adjustments. This not only minimizes potential losses due to defaults but also enhances customer satisfaction by providing timely and relevant support. Furthermore, the model's insights can be leveraged to refine credit underwriting policies, improve customer segmentation, and optimize portfolio profitability. Continuous monitoring and refinement of the model will be crucial to ensure its ongoing effectiveness in the evolving credit landscape.