INTERNATIONAL
UNIVERSITY OF
APPLIED SCIENCES

Credit Card Routing for Online Purchases via Predictive Modelling

University of Applied Science - Online

Masters in Data Science (MSDS60ECTS)

Case Study: Model Engineering(DLMDSME01)

Gummadi Sai Dheeraj

Matriculation: 3210935

gummadi.saidheeraj@iu-study.org

Advisor: Sahar Qaadan

Delivery date: October 09, 2024

Contents

I List of Figures

1 Abstract

This case study addresses the issue of increasing online credit card payment failure rates and decrease the transaction fees of a major online shop. CRISP-DM has been applied while designing the solution as learned through this research machine learning model identifies the best Payment Service Provider (PSP) using success likelihood with transaction fee. The experimentation includes several machine learning algorithms, including Logistic Regression, Decision Tree, and Random Forest Classification, to construct and test the models. The XGBoost classifier outstands as the best model because of its good success probability prediction and PSP selection. The quality assessment of the data; feature selection and processing imbalanced data are relevantly put in focus by the paper. Final deployment plan presents a holistic approach toward integration with the business environment to smoothen and solidify the payment service. Concluding the study with some possible directions for research and further enhancement needed in the credit card routing system.

2 Introduction

In the evolving world of e-commerce, credit card processing must flow effortlessly and efficiently in order to ensure both a safe customer experience and financial gain. The critical challenge is a sudden and significant increase in online credit card payment failure rates over the past year. This has translated into substantial monetary loss to the company and frustrated customers about their online shopping. The existing credit card routing system that is running on payment service providers (PSPs) is based on a set of predefined rules have not been appropriate enough to resolve this issue so far. Regarding this issue, predictive modeling is now being sought after by the business stakeholders.

The primary objective is to develop an automated credit card payment service provider routing system that maximizes success rates for payments, but does this in a strategic manner that also reduces transaction costs. It is, essentially, a project following a full-cycle approach towards applying CRISP-DM. It introduces a strategy with machine learning to meet the twin objectives of improving success rates and cutting the transaction cost. The intention of transforming the online retail industry's credit card routing by way of investigation, data analysis, preparation, modeling, evaluation, and deployment that this objective takes place. At the same time, the business owes these to a broader objective of financial efficiency and customer satisfaction.

2.1 Problem Statement

A leading retailer has an online payment department that is experiencing a trend of unexplained increases in failed credit card transactions on the Web, which leads to great monetary losses and dissatisfied customers. The current manual practice of routing the credit cards through PSPs is inefficient and provides an impetus for finding a data-driven solution to routing a credit card. The major task is that of developing predictive models that can help automate the process of routing payment service providers and raise success rates for transactionswith minimal transaction fees. This assignment will be based on choosing one PSP depending on feasibility and prospects as well as contractual arrangements with four PSPs. Finally, the purpose of this is to design an effective predictive model system that appropriately serves the organizational mission toward business financial success as well as customer satisfaction.

2.2 Business Context

Online credit card payments experience a high failure rate, thus resulting in losses and dissatisfaction among customers in an online payment department of a retail company. Four PSPs operate at different transaction costs taking part in online payments and credit card routing is made manually, which is expensive, hence requiring automation using predictive modeling.

This achieves better performance in terms of financial performance due to the regulation of losses emanating from unfruitful transactions. Improve customer satisfaction with a better online shopping experience.

3 Methodology

The widely approved framework for data science projects is used in this case study, namely the Cross-Industry Standard Process for Data Mining (CRISP-DM), which will guide the development of a predictive model for credit card routing for online retail. The CRISP-DM methodology makes it possible to achieve a structured and iterative approach toward successful projects; hence, it indicates cooperation between different stakeholders in regard to the project being faced to meet objectives - business experts all the way up to data scientists.

3.1 Data Understanding

This is a collection and exploratory phase, aiming to understand the quality and nature of the data. Source identification and preliminary exploration have thrown light upon inherent properties of the dataset.

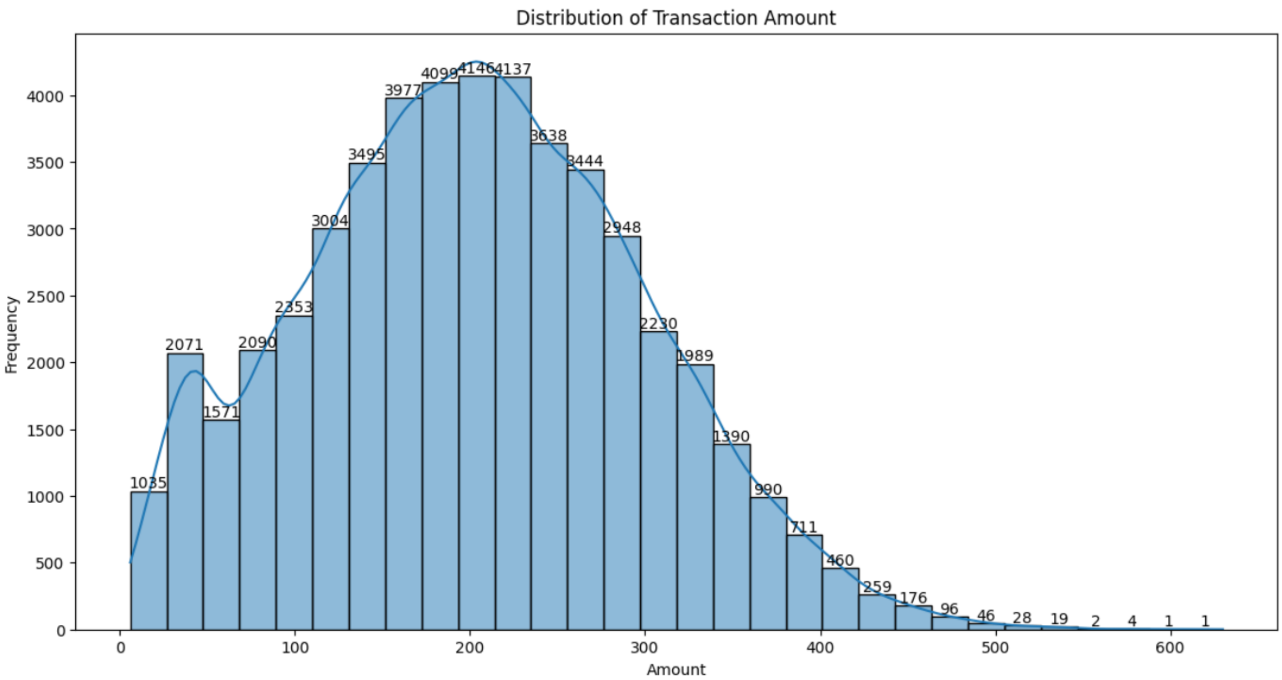


Figure 1: Amount Distribution

Fig. 1 plots the amount distribution; the data is observed to be provided in the form of right skewed distribution. The top position of the curve is found to be around 200, which means most of the transactions lie within this range. It means that most of the transactions are close to 200 with fewer transactions below or at significantly higher values. This is important information as it will help understand average transaction values and easily identify outliers. For instance, when these amounts run far above 200, further action could be considered to ascertain whether they are valid transactions or not.

Fig 2 visualizes the distribution of successful and unsuccessful transactions. The graph shows the successful and unsuccessful transactions clearly showing a significant difference here as the unsuccessful ones are much more than the successful ones. This shows the pain the payment failure is causing the company, as the lost revenuers and frustration of the customers are very significant in this regard. The gap in success against failure

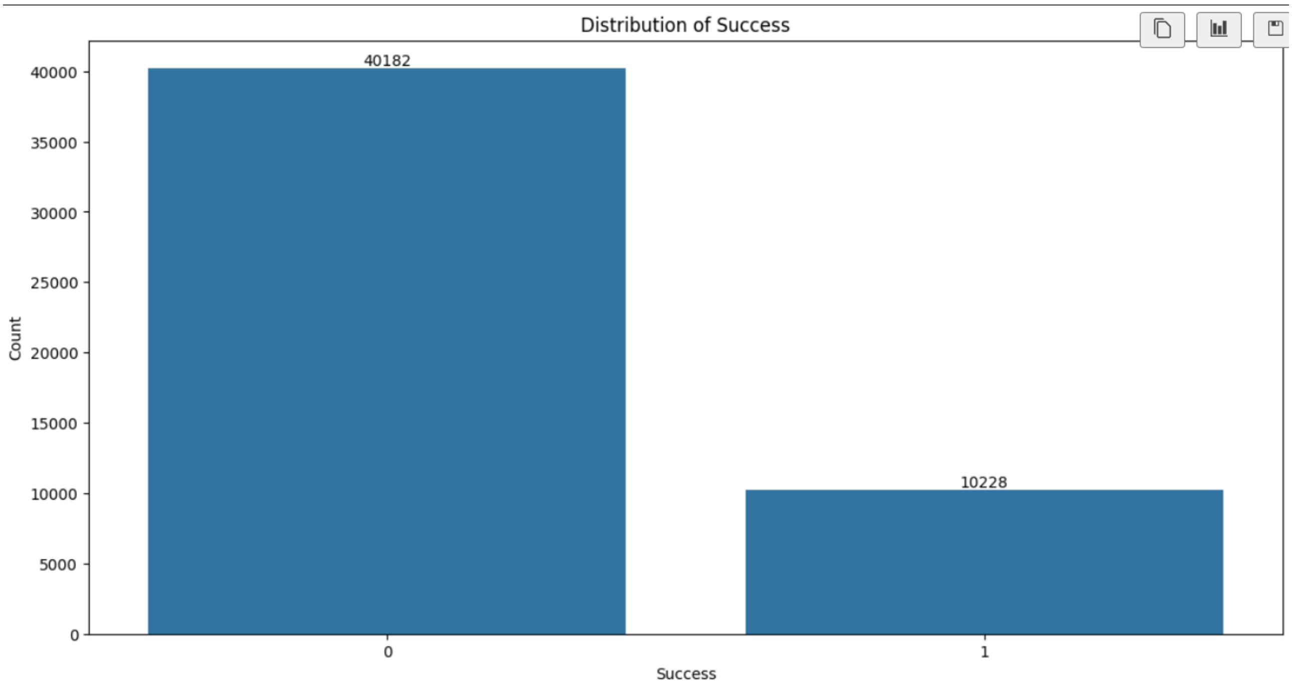


Figure 2: Success Distribution

rates truly indicates that an answer has to be found to elevate the success ratio and lower the rate of failure in payments. This could be through the implementation of firmer fraud detection tools, streamlining the process of payment processing, or even collaboration with a better provider of payment services.

Fig 3 shows the percentage of successful transactions for each type of PSP. Goldcard stands high in this regard, followed by Moneycard, UK Card, and Simplecard. This may be interpreted to mean that Goldcard is the most trustworthy payment service provider in terms of payment since it provides more successful transactions on a regular basis. The company can use this information to route as many transactions as possible directly to Goldcard so as to maximize the payment success rate and minimize the failure risk. A rule-based system or a predictive model can thus be developed to automatically route transactions with the highest likelihood of success with Goldcard for processing.

Fig 4 shows country-wise transaction count. The country with the maximum number of transactions was Germany, followed by Switzerland and Austria. This would tell a lot about the geographical spread of the business and trend in a region. For instance, with the higher transaction count on the German site, it could be said that there are more customers in Germany or even more shopping activity taking place online in Germany. It might be very useful when an organization plans to use tailoring marketing campaigns or to optimize logistics for specific regions.

Fig 5 box plot depicts the transaction amount by each of the countries' distribution. The median for both countries is the same, so the average value of the transactions is similar in all of them regardless of their location. However, box plots indicate differences in data points spread, which might imply that the transaction patterns were different for countries. For instance, the wider dispersion in values for Austria could be expressed as a greater variety of sums being tied up in transactions, possibly originating in that country through the provision of a wider variety of products or services. This is useful in an understanding of the patterns of purchasing by customers and, accordingly, the adaptation of marketing schemes.

Fig 6 illustrates the transaction success rate trend of time. The success rates keep on oscillating throughout the period and are not found to be in any upward or downward fashion. This implies that the success rates have been highly influenced by different factors that do not come out from the data, for instance, seasonal factors or change

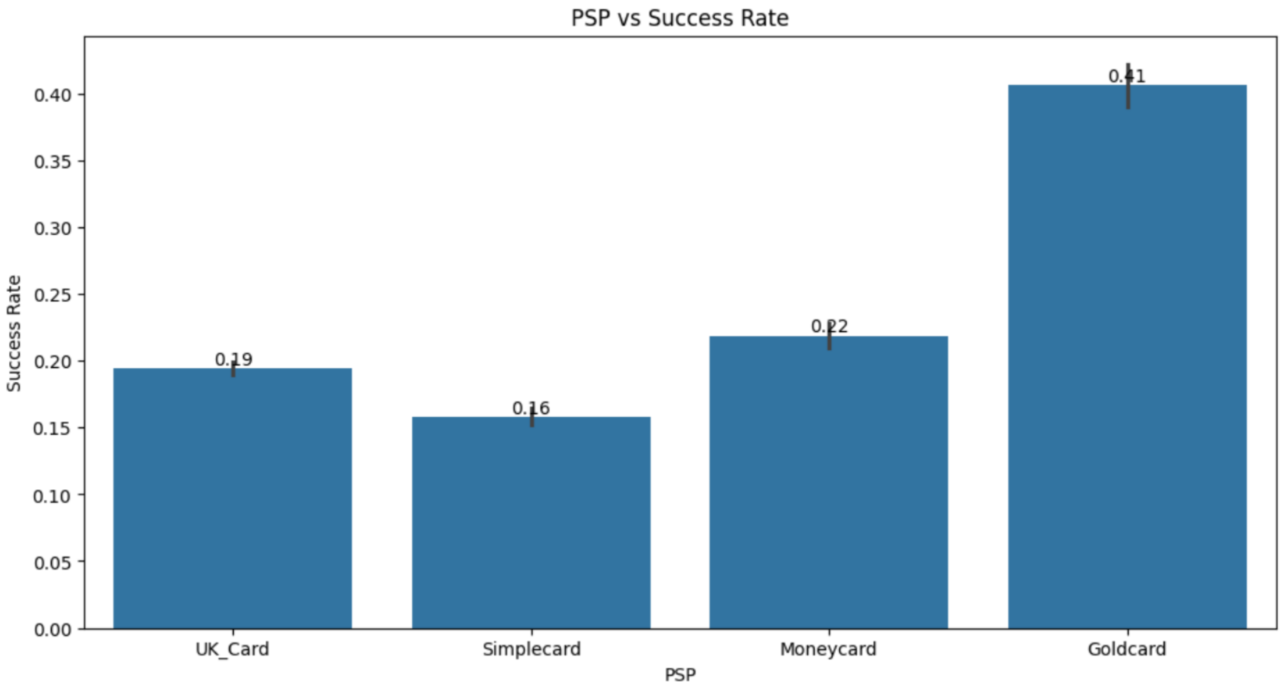


Figure 3: PSP vs Success Rate

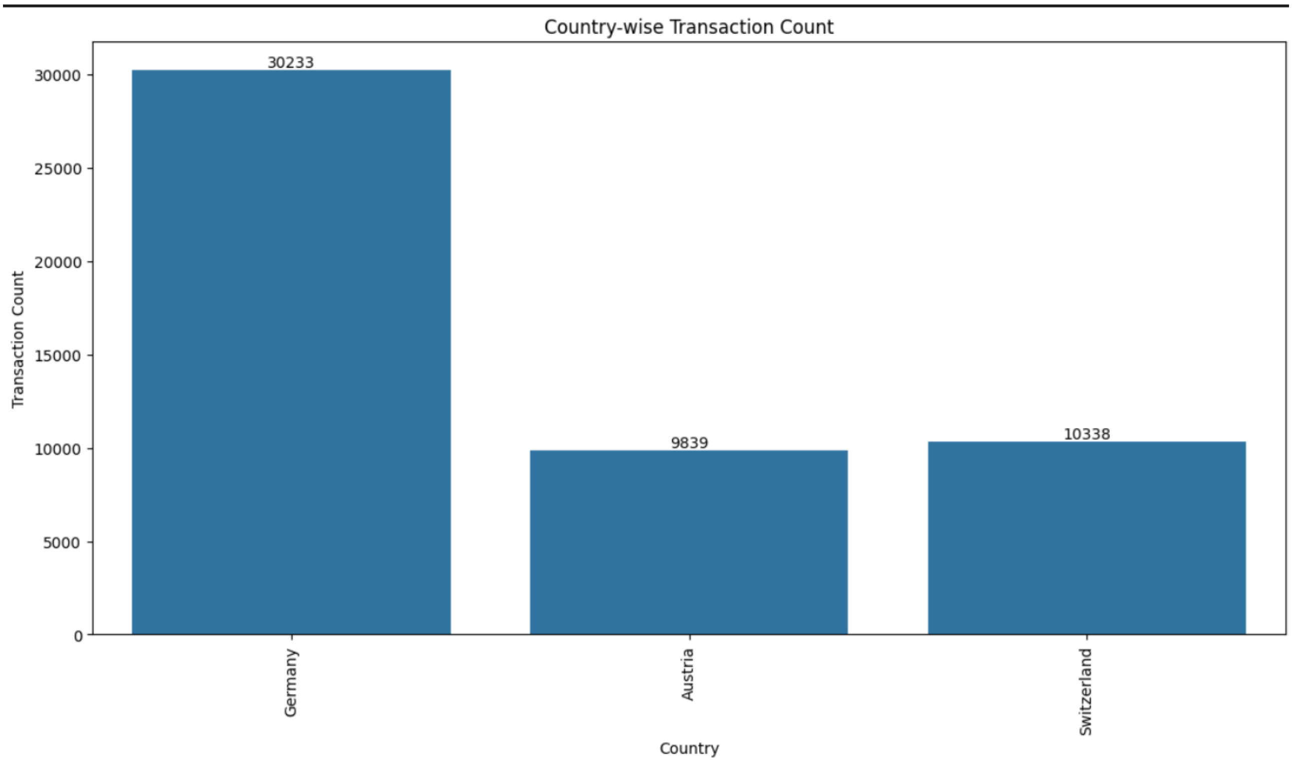


Figure 4: Country vs Transactoin count

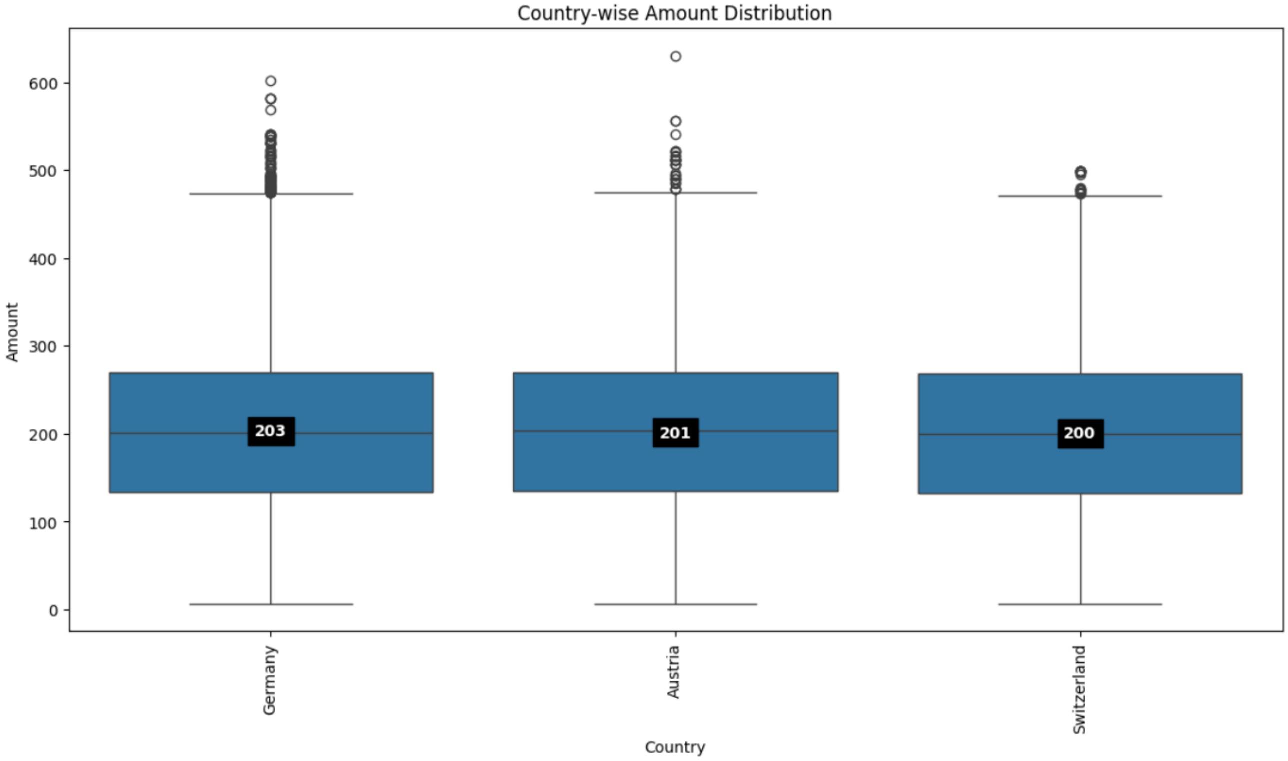


Figure 5: Country wise amount distribution

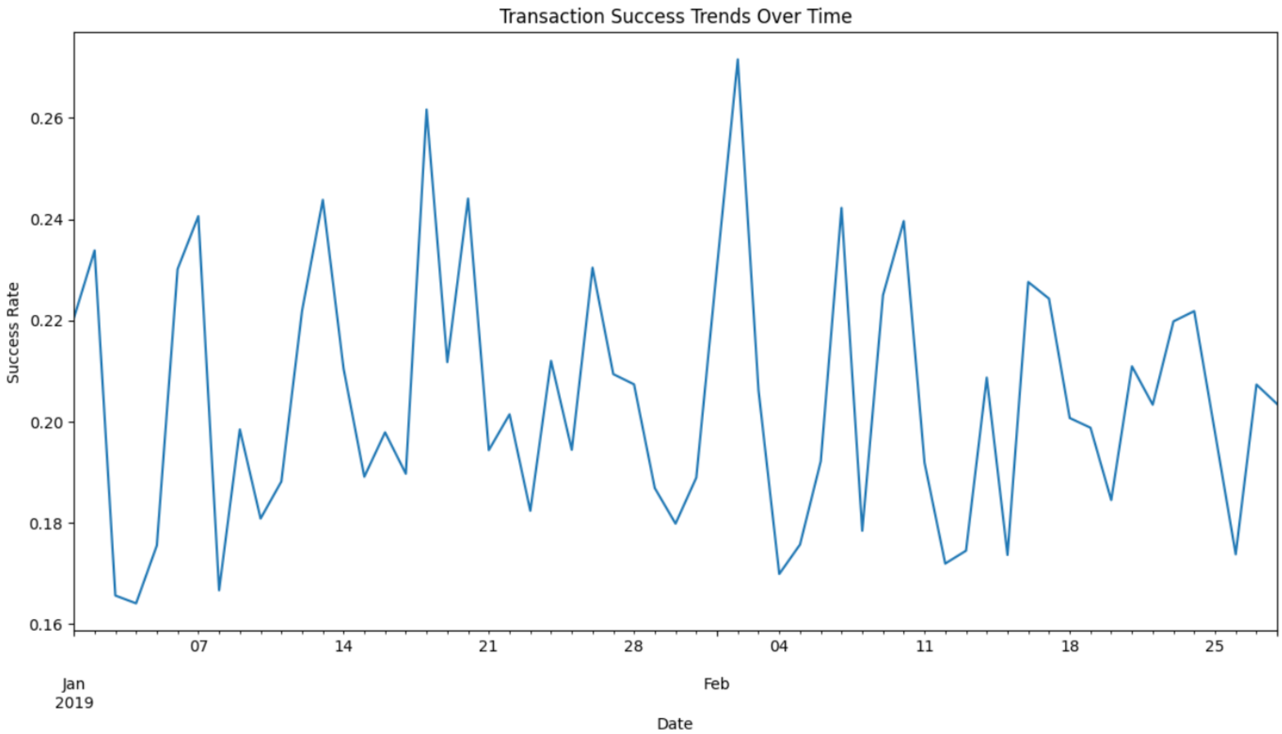


Figure 6: Transaction success trends

in customer behavior. For example, an improvement in the success rate for a particular month may be attributed to the load on the cards server or less number of transactions in that time period.

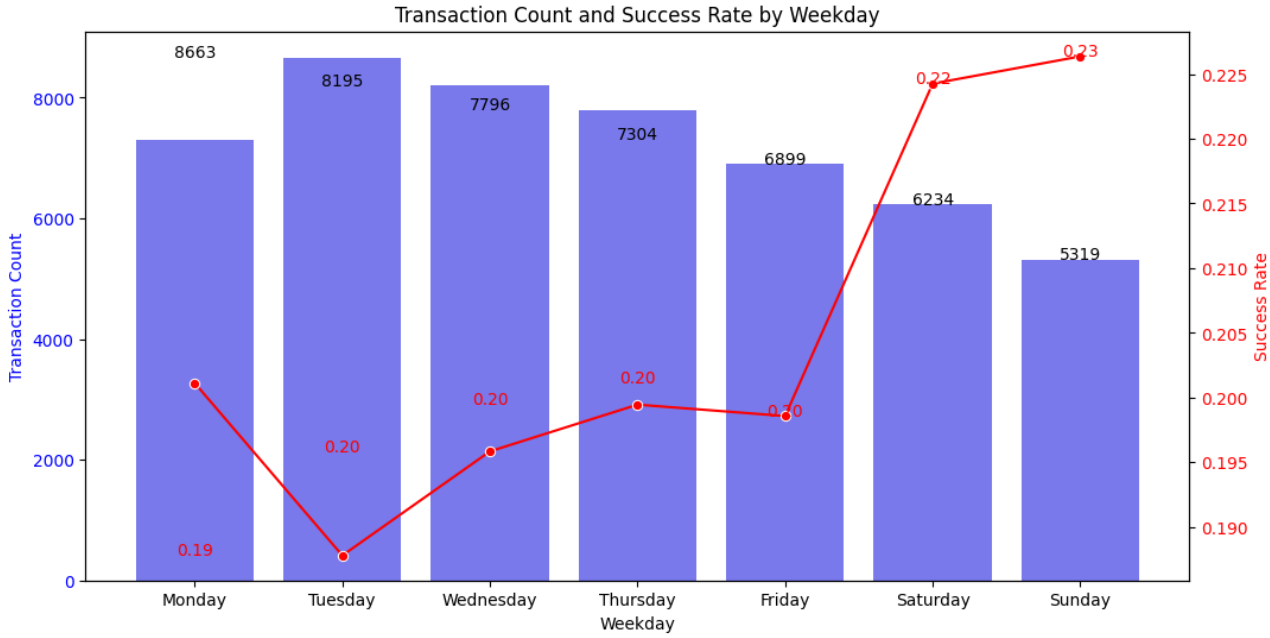


Figure 7: Transaction count vs Success rate by Weekday

Fig 7 shows the highest transaction counts happen on Tuesday, and Wednesday, going gradually down till the weekend. The success rate is lower on Tuesday, when there is a high number of payments which indicate payments may be failing since there is heavy loads on the servers. This portrays the customer's behavior is not so inconsistent in terms of its success ratio within weeks, but may probably increase its success rate by weekends.

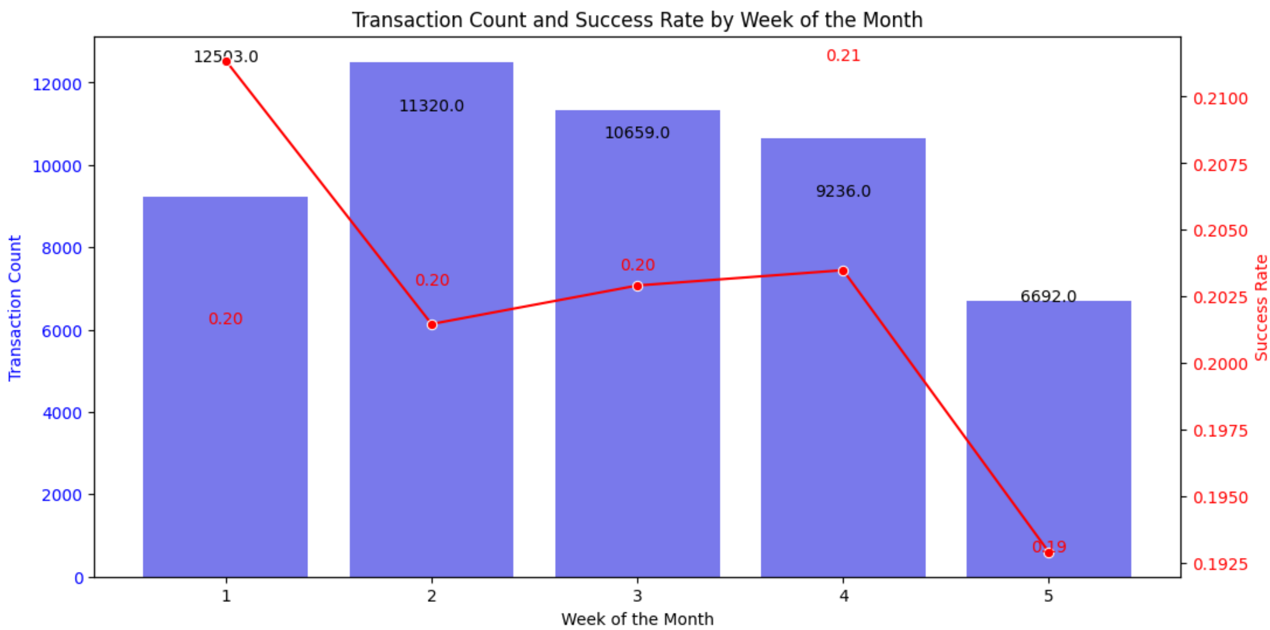


Figure 8: Success Rate by Week of the Month

As can be seen in Fig 8 success rate generally increase from the second week up to the fourth week of the month, then dips a bit into fourth and very significantly into fifth. This indicates there may possibly be some patterns associated with the week of the month that relate to the customer behavior and the payment process.

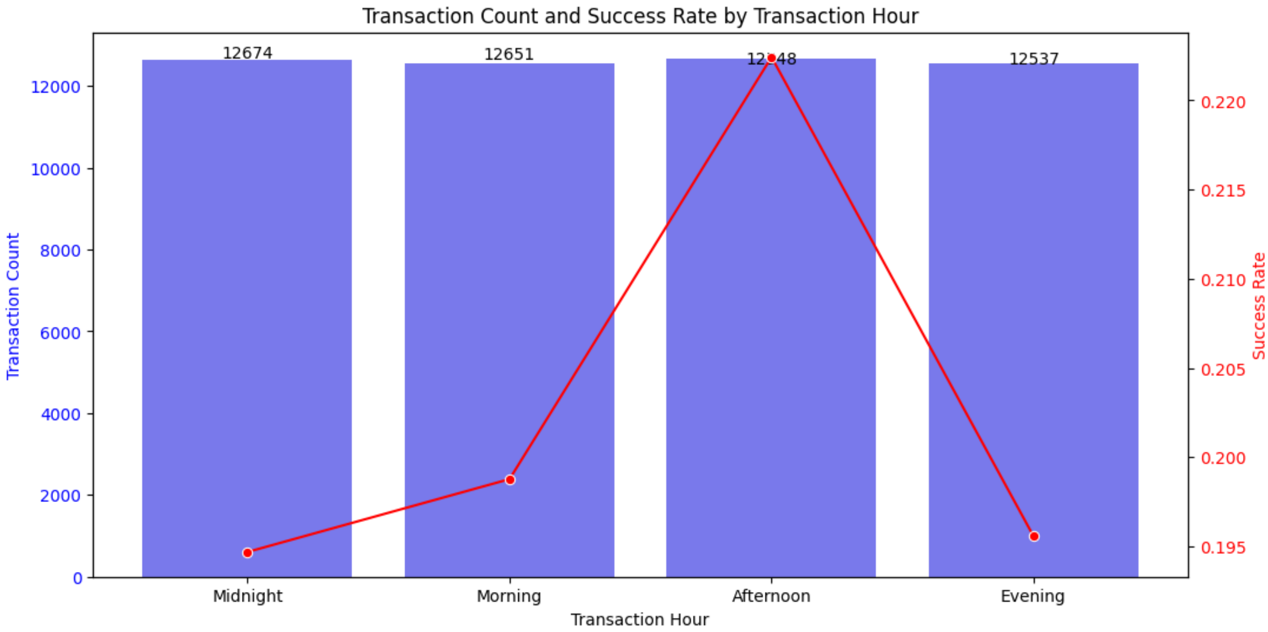


Figure 9: Transaction Count and Success Rate by Transaction Hour

Fig 9 depicts that the transaction count in all four time periods have been fairly consistent with a small peak in the Afternoon. This seems to denote the fair activity of customers throughout the day as their percent of success rate is also found to be relatively stable with a small dip in the Evening. This means that the likelihood of a transaction being accomplished is not strong at all and likely not affected much by the time of day. The graph itself doesn't appear to indicate the presence of strong time-based patterns in transaction success. This would suggest that those factors to do with the success or lack thereof of a transaction might be more complex than this and not so affected by the time of day in particular.

Fig 10 is trend for total amount of transactions over time. It varies for the period with no apparent uptrend or downtrend. This suggests that the total amount of transactions seems somehow responsive to the effects of different other factors like seasonal variations and shift in customers' consumption patterns. In simple words, if the total transaction amount has grown in the beginning of the month, then it must be due to a campaign or changing customer spendings. This will enable the firm to find sources of revenues that would otherwise remain hidden and optimize their marketing efforts.

Fig 11 shows success rate for each category of the transaction amount. Success rate would be the highest when the amount of the transactions is low, followed by medium amount and finally high amount. It simply means that the smaller the transaction, the higher will be its chance of success, which may be due to some bigger customers or even strict fraud checking. This data can be used to facilitate an understanding of the relationship between transaction amount and success rate, allowing for appropriate strategies to be devised to increase the success rate for those transactions with higher amounts.

Fig 12 shows the success rate for each category of amount and card type. Success rate is highest for low amount transactions and for Diner cards, followed by low amounts and Master cards, and lastly low amounts and Visa cards. It actually suggests that the smaller transactions made using Diner cards are more likely to go through, probably because they have higher customer values or because of stricter fraud checks. This may be very useful to know and hence develop strategic suggestions for optimizing the payment process based on the card type.

Fig 13 shows the success rate by transaction hour. The success rate is relatively uniform in all hours indicating no critical effect of the time of the day on the chances of a transaction being successful. This would indicate that other factors, namely the chosen PSP, the amount of the transaction or payment method chosen by the customer

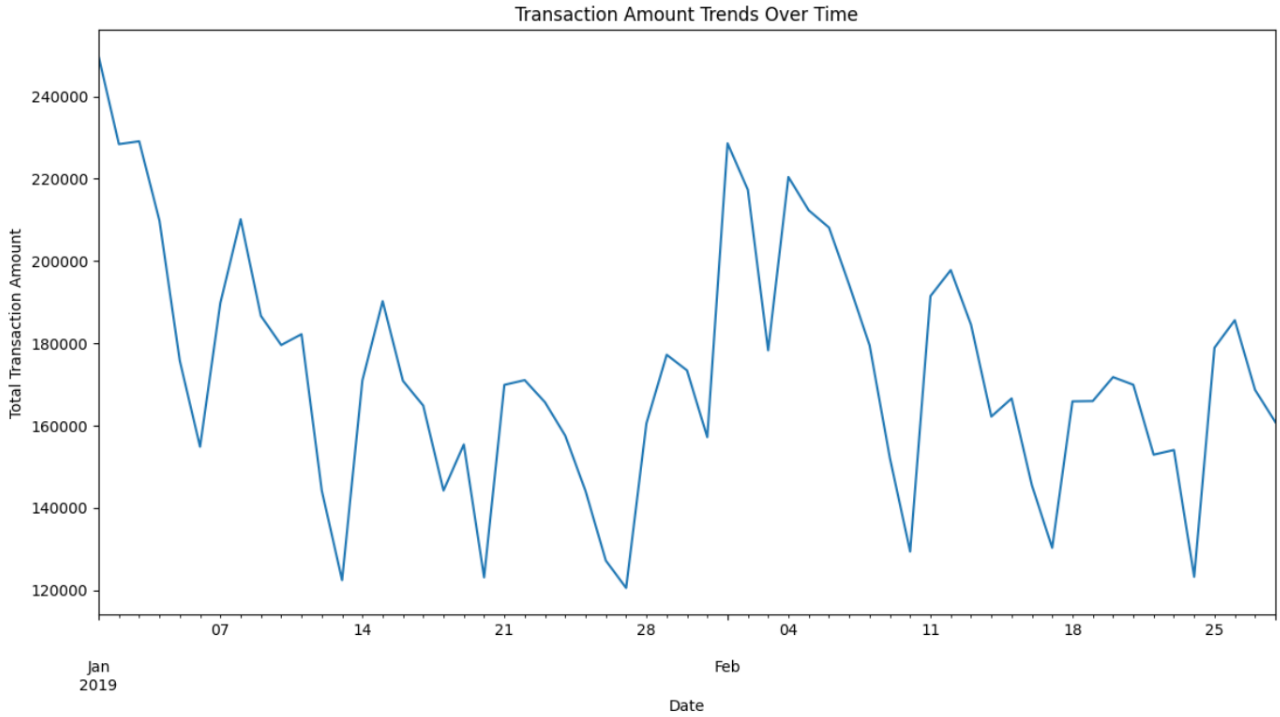


Figure 10: Transaction amount trend overtime

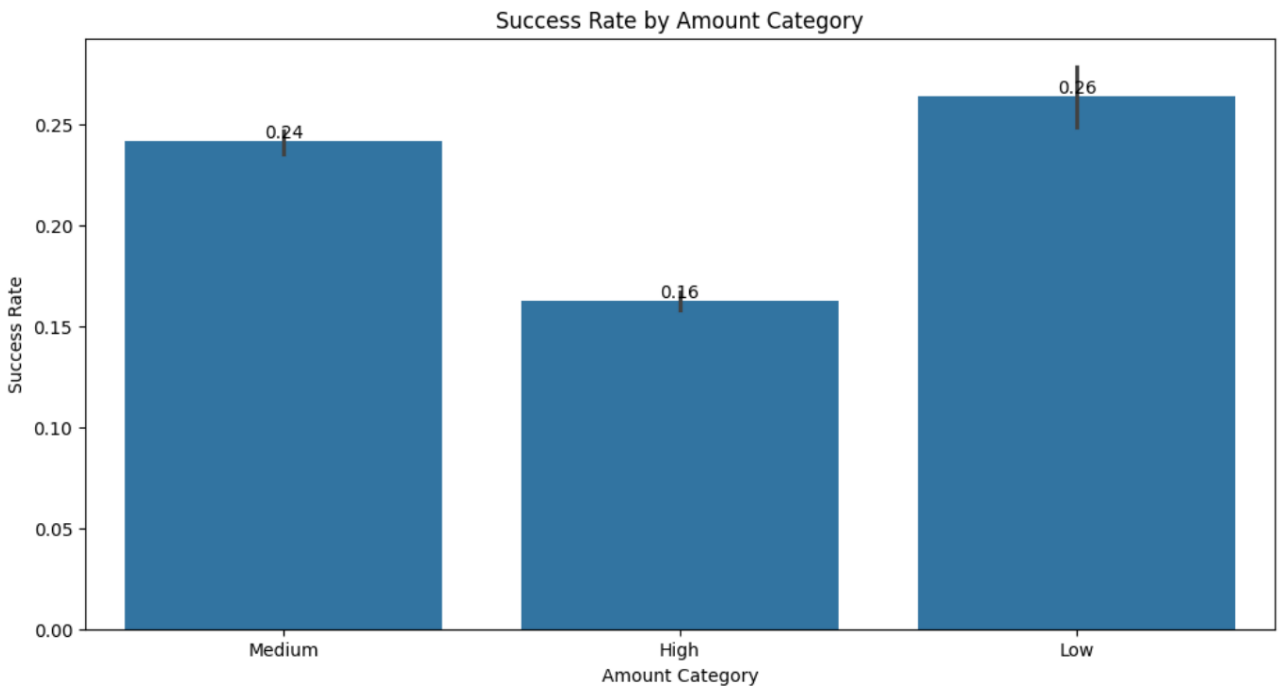


Figure 11: Success Rate by amount category

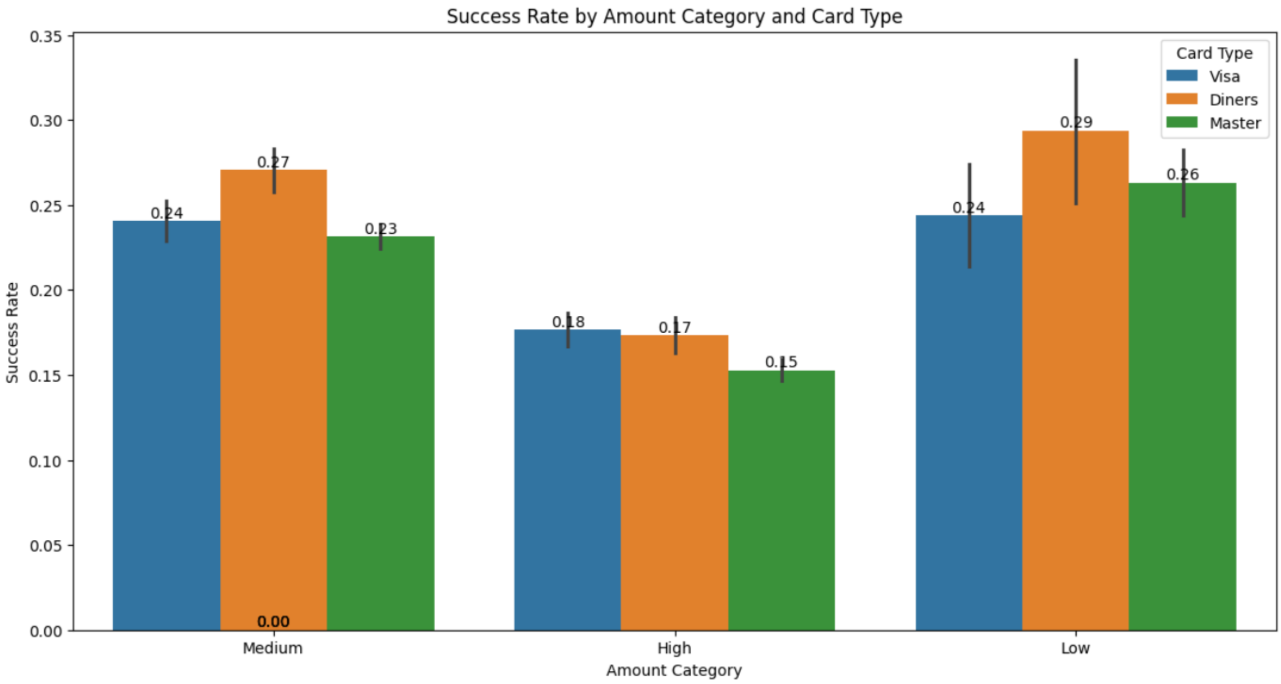


Figure 12: Success Rate by amount category and Card Type

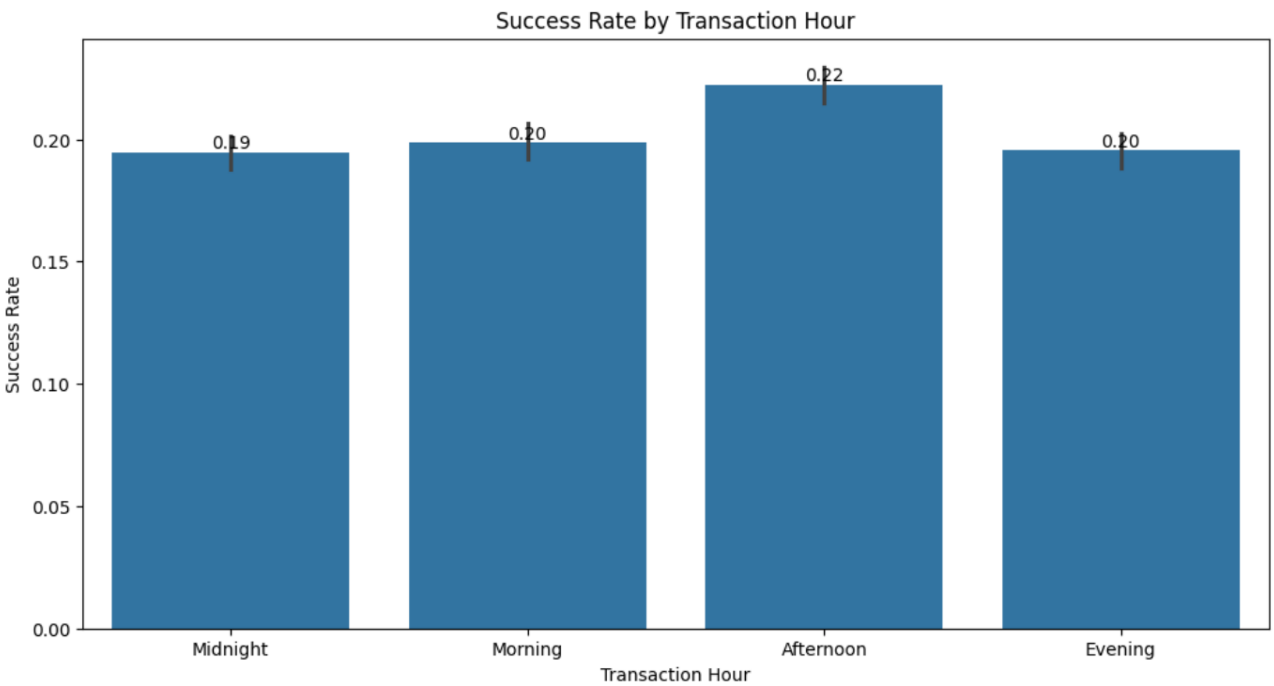


Figure 13: Success Rate by Transaction Hour

play the leading role. This information is useful in understanding the factors that contribute to payment success and in the devising of strategies in order to improve the overall success rate.

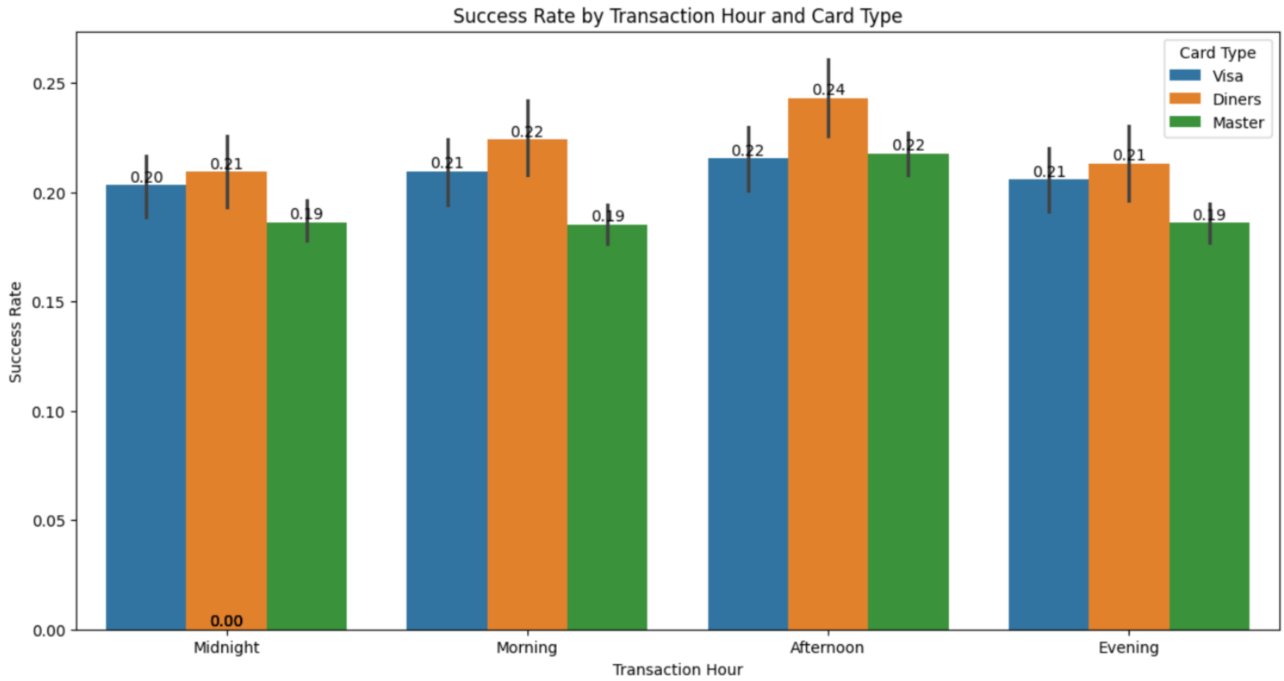


Figure 14: Success Rate by Transaction Hour and Card Type

Fig 14 is a graph that demonstrates the success rate for each transaction hour and card type. Here, the success rate is reasonably constant at all hours and card types, meaning that neither time of day nor type of card are strong predictors of a successful transaction. This, therefore, implies that the success rate is influenced by other variables such as the chosen PSP, the amount of transaction, or any form of payment a customer offers. It can be useful for understanding factors that may help increase the overall success rate.

Fig 15 depicts the success rate of the respective PSP and card types. This gives the impression that the success rate is the highest for Gold Diners card, then comes Gold Master cards followed by Gold Diners cards. This means Gold Diner cards seem to be the best combination that will more or less process the payment without failing. This may prove useful for a strategy on the way to handle transactions where the selection of Gold masters cards could be given preference, trying to have the maximum number of successful payments and a minimum number of failed payments.

Fig 16 displays the success rate for each PSP and for each weekday. The highest success rate is obtained using Gold cards for any weekday, followed by Money cards, followed by UK cards and Simple cards. From this, we can infer that the combination of Gold cards along with Money cards are the safest options used to successfully make payments at any weekday. This information would hence be utilized to formulate an even more informative strategy of routing transactions in consideration to the PSP, type of card, and a day of the weekday so as to maximize the levels of successful payments spread all over the days of the week.

3.2 Feature Engineering

The week of month feature captures potential cyclical patterns in transaction behavior within a month. For example, a particular week could be busier compared to others or have a different transaction success rate simply because it's the weekly payday, holiday, or season. Analysis of the transaction success rate by seeing first week as opposed to last week could provide leads on customer expenditure patterns and possibly changes in

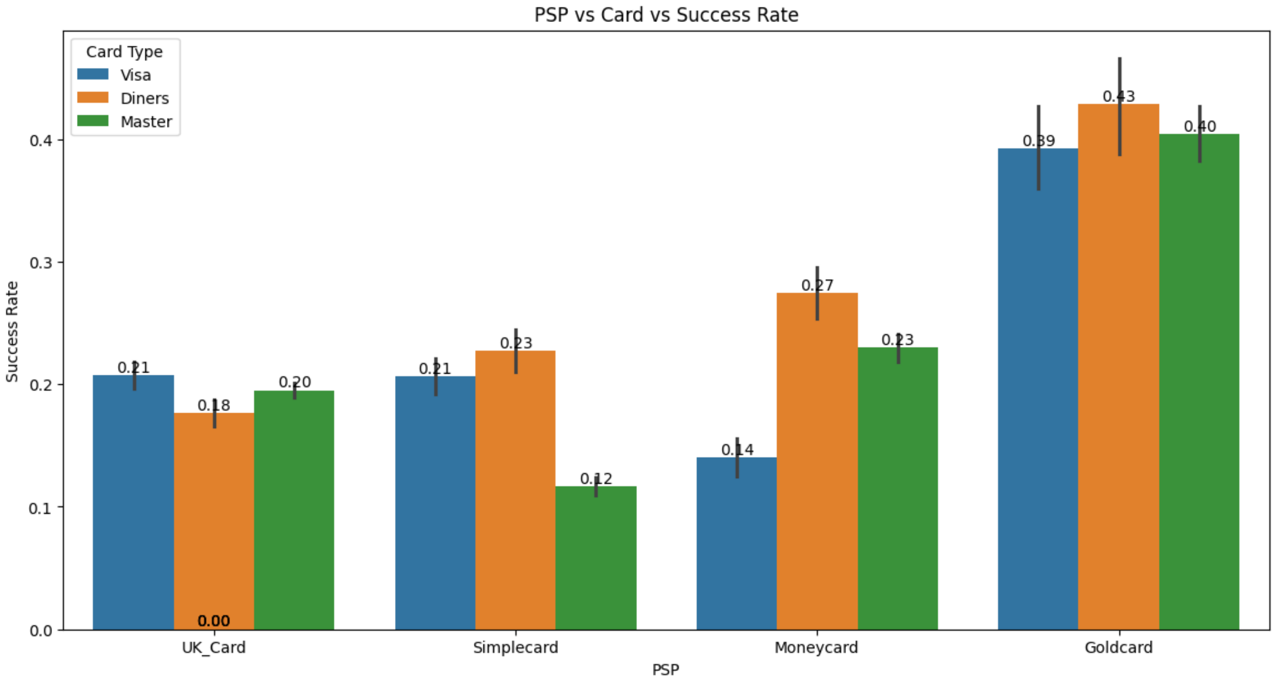


Figure 15: PSP vs Card vs Success Rate

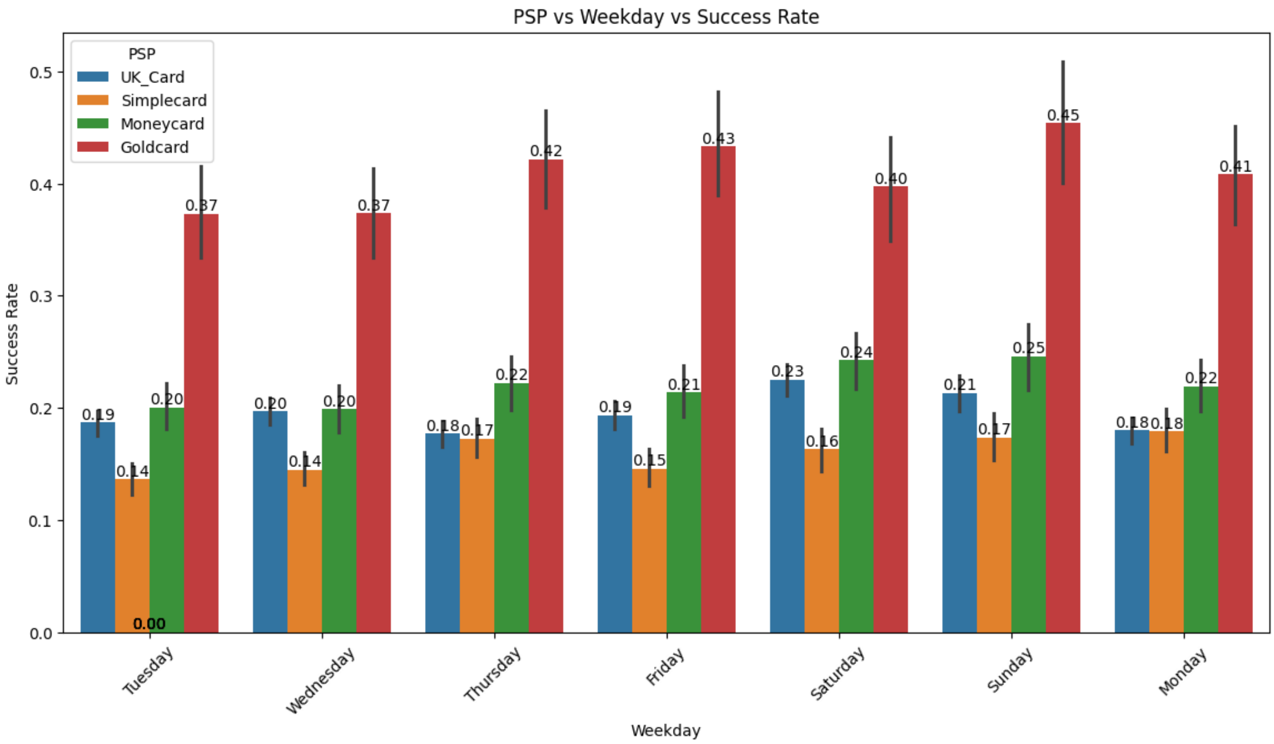


Figure 16: PSP vs Weekday vs Success Rate

creditworthiness.

Transaction Hour feature provides categorization of the hour of transaction into broad time blocs (Midnight, Morning, Afternoon, Evening) so that one gets a finer perspective of the behavior of transaction occurrences between these different parts of the day. One could also examine a success rate comparison of ”Evening” hour of the transactions against ”Morning” hour to detect various customer preference shifts or system performance fluctuations.

Amount Category feature classifies the amounts of transactions into ”Low,” ”Medium,” and ”High” ranges, which enables how the value within transactions influences success to be investigated. Amounts that are higher may be more heavily checked or have different risk profiles. One might find that transactions of ”High” amounts are dramatically more successful than transactions of ”Low” amounts, shedding light on whether transaction value is correlated with creditworthiness.

The success rate for specific card types in different countries would be captured using the Country-Wise Card Success Rate feature. This could then be used to understand regional variations in payment patterns and card usage patterns. Comparing the success rate of transactions in ”UK Card” between Germany and Austria would therefore help give insights into differences in successrate in different regions or card acceptance policies.

Oversampling with SMOTE is used for the highly skewed distribution of PSP’s. It generates synthetic samples for the classes having fewer numbers of samples in it and balances the class distribution. This allows for a balanced relationship between the classes and decreases the chance of a model being overly biased during training. All categorical features are converted into numerical using LabelEncoder so that models can suitably process these features during training. This data preprocessing step ensured that the numerical and categorical data were fed to the models so that they could learn meaningful patterns from this that could improve the prediction accuracy.

3.3 Modelling and Results

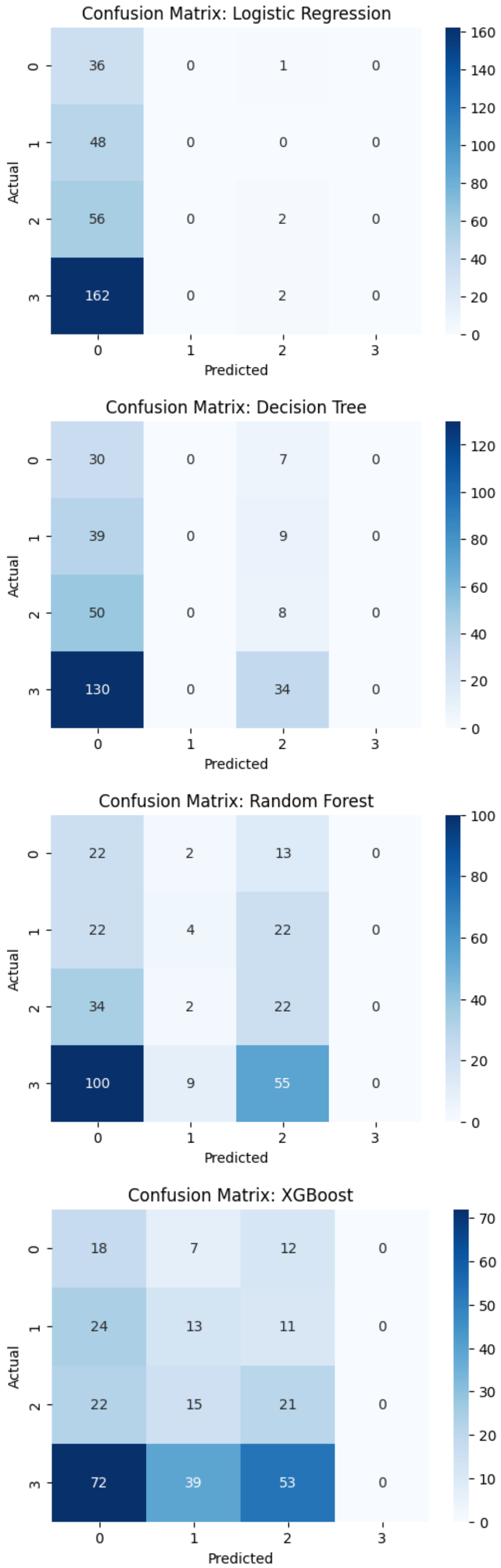
This section explains the modeling procedure that was used to predict changes in credit scores based on the engineered features. Four most widely used machine learning algorithms that are Logistic Regression, Decision Tree, Random Forest, and XGBoost are used in the experimentation.

Each model was trained on oversampled training data, and then they were all tested on the unseen test set and Performance metrics were calculated such as accuracy, F1 score, precision, recall, and ROC AUC on the metrics to estimate the performance of the model. Fig 17 displays the baseline model’s confusion matrices and evaluation metrics.

Initial indicated that the XGBoost model had surpassed all others by a good margin due to its high accuracy and F1 score. This can be evidenced by the better classification based on the confusion matrices, where the XGBoost achieved greater accuracy for all classes. Cross-validation and hyperparameter tuning techniques are used to fine-tune the performance of the XGBoost model. The former is estimating how well a model would generalize to new, unseen data, while the latter seeks the optimal combination of model parameters with a view of getting more accurate results.

The hyperparameter tuning process yielded the hyperparameters mentioned in the table 1:

These parameters resulted in a significant improvement in the model’s performance, achieving a test accuracy of 53.42%. This indicates that the hyperparameter-tuned XGBoost model effectively captured the complex relationships between the features and the credit score adjustment prediction. The XGBoost model, after hyperparameter tuning, emerged as the most effective model for predicting credit score adjustments. Its performance is evidenced by the evaluation metrics and confusion matrix analysis as mentioned in Fig 18



	Model	Accuracy	F1 Score	Precision	Recall	ROC AUC
0	Logistic Regression	0.123779	0.037593	0.089937	0.123779	None
1	Decision Tree	0.123779	0.051343	0.040579	0.123779	None
2	Random Forest	0.156352	0.092806	0.088795	0.156352	None
3	XGBoost	0.169381	0.109593	0.084320	0.169381	None

Figure 17: Confusion Matrix and Evaluation metrics of the base models

Hyperparameter	Value
colsample_bytree	0.8
learning_rate	0.01
max_depth	3
n_estimators	50
subsample	0.8

Table 1: XGBoost Tuned Hyperparameters

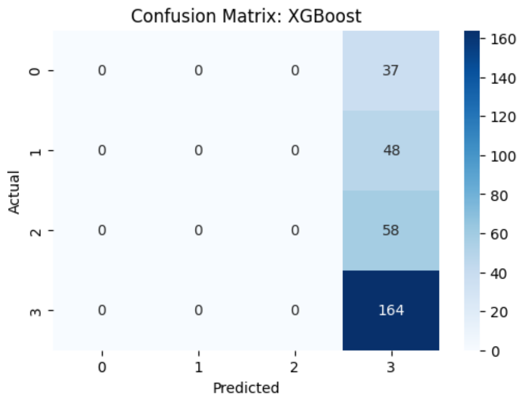


Figure 18: Confusion Matrix of the Hyperparameter tuned XGBoost

3.4 Model Evaluation metrics

This section accounts for the performance criteria evaluation metrics of the trained models in predicting adjustment of credit scores.

Accuracy metric measures correctly classified instances, the number of correctly predicted credit score adjustments from the total number of instances in the test set. It generally gives an overview of the model's overall performance. A balanced harmony of precision and recall, giving an average measure of the model's performance, constitutes the F1 score. It accounts for both the model's ability to correctly identify positive instances which correspond to the determination of adjustments (precision) and its ability to capture all positive instances in the data set. Precision measures the percentage of the correctly predicted credit score adjustments among all instances predicted as adjustment and explains the ability to avoid false positives on the part of the model.

Recall metric measures the percentage of the PSP's for which the actual PSP's are correctly predicted. This essentially simply says that out of all of the actual modifications, this model has captured all of the true positives.

Based on the metrics of evaluation, the best model in terms of all the models with regard to tuning hyperparameters was XGBoost. The highest accuracy, F1 score, precision, and recall for the exact payment service provider prediction were shown by XGBoost.

3.5 Deployment and Model Monitoring

It is designed to deploy in real-time for predicting the payment service provider for each individual transactions. It can be deployed as a web service that accepts user inputs and responds immediately with recommendations. The user interface takes Country, Card, Transaction Date, Amount, 3D_Secured features as input and The "Country-wise Card Success Rate" features, which were computed from the training data, are stored as static lookup tables. These tables are used to map user-provided country and card information against the success rate.

A strong monitoring system must be in place tracking at regular intervals the main performance metrics -

Accuracy, F1 Score, Precision, and Recall - so any deterioration in performance is picked up in due time. When major deviations are seen, the model will be retrained with updated data to maintain the accuracy of the model.

4 Conclusion

The study leverages machine learning model, particularly XGBoost, to analyze payments behaviour and providing valuable insights for choosing right payment service provider for each country with minimal transaction cost. The study highlights the importance of addressing data imbalances, which are prevalent in real-world scenarios. The implementation of SMOTE, a synthetic oversampling technique, effectively balanced the dataset, mitigating the risk of biased model training and enhancing the model's generalizability.

The comprehensive evaluation process, utilizing metrics such as accuracy, F1 score, precision, recall, and ROC AUC, confirmed the superior performance of the hyperparameter-tuned XGBoost model.

The deployment architecture, designed for real-time payment service provider prediction by leveraging static feature mapping to ensure consistent and accurate feature extraction for new data points. The robust monitoring system, tracking key performance metrics and implementing automated retraining mechanisms, ensures the model's continued effectiveness and adaptability to evolving market conditions.

5 Appendix

The code for the above mentioned modelling approach is available in the following github link:
Github: https://github.com/GSaiDheeraj/Credit-Card-Routing-for-Online-Purchase-via-Predictive-Modelling/