INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

MSc. Data Science (120 ECTS)

Model Engineering (DLMDSME01)

CASE STUDY

Credit Card Routing for Online Purchase: ***A Predictive Modelling Approach***

Joël Kazadi

Matriculation Number: 9213934

Tutor: Prof. Dr. Sahar Qaadan

Delivery date: November 9, 2024

Contents

List of Figures

List of Tables

Abbreviations

GB	Gradient Boosting
KNN	K-Nearest Neighbors
KPI	Key Performance Indicator
ML	Machine Learning
PSP	Payment Service Provider
RF	Random Forest
ROC AUC	Receiver Operating Characteristic – Area Under the (ROC) Curve
SMOTEENN	Synthetic Minority Oversampling TEchnique and Edited Nearest Neighbours

1 Problem definition

In today's competitive retail landscape, the company "PaySmart" faces significant challenges in managing online credit card payments. Business stakeholders have raised concerns regarding the high failure rate of these transactions, which has resulted in substantial financial losses and diminished customer satisfaction. The current routing logic for selecting payment service providers (PSPs) is static and rule-based, leading to inconsistencies and inefficiencies. This approach not only increases the likelihood of transaction failures but also incurs higher transaction fees, as the routing decisions do not consider the specific characteristics of each transaction.

To address these challenges, PaySmart aims to replace the manual system with a predictive model. The model should be capable of dynamically routing each transaction to the most suitable PSP, based on a variety of factors, such as transaction amount, credit card type (secure or not), and past success rates. This data-driven technique allows for real-time decision-making, enabling better adaptation to patterns and trends in transaction data that a static, rule-based system would miss.

The potential benefits of implementing such a predictive model are significant, both in terms of value proposition for the customer and success definition for the company. By selecting the most cost-effective PSP for each transaction, the company can lower transaction fees to enhance customer satisfaction (value proposition). Additionally, by improving the success rate of payments, the company can minimize the the risk of losing customers to competitors in order to increase the overall profitability (success definition).

2 Data collection and Preprocessing

The data required for this project is provided by the online payment department of the company. Table 1 describes the variables contained in the dataset in its raw state, i.e. before the pre-processing phase. There are a total of 50,410 instances and 7 attributes (not including the target variable).

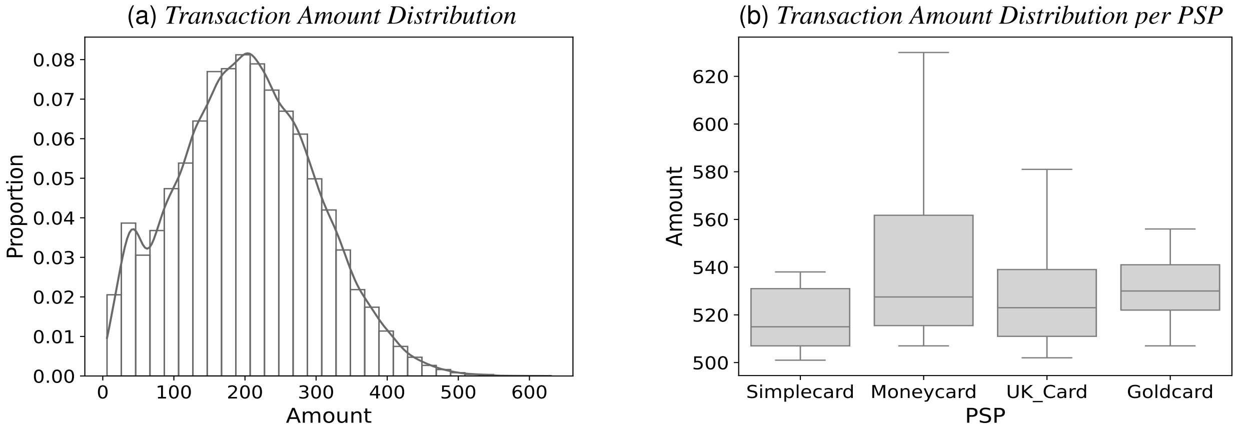
Tab. 1: Variables Description

Role	Variable Name	Type	Definition
Features	*ID*	String	Identification number of the transaction.
	tmsp	Date Time	Timestamp during which the transaction was completed.
	country	Categorical	Country of transaction (Germany, Switzerland, Austria).
	amount	Integer	Transaction amount (in euros).
	success	Binary	1 if payment is successful, 0 otherwise.
	3D_secured	Binary	1 if customer uses a more secure credit card (3D identification).
	card	Categorical	Credit card provider (Master, Visa, Diners).
Target	*PSP*	Categorical	Name of one of the four payment service providers.

2.1 Data Quality Assessment

The data quality assessment shows that the dataset does not contain any missing values. Furthermore, all the data conforms to the predefined variable format. Therefore, there is no inconsistency in the structure of the data.

Fig. 1: Outliers checking



However, an examination of the '`transaction_amount`' variable indicates that it does not follow a normal distribution, which raises concerns about the presence of outliers, particularly for transaction amounts exceeding €500 (*see* panel (a) in Figure 1). Notably, these outliers are unlikely to be fake ones, as they exhibit a correlation with the PSP utilized. This implies that certain PSPs accommodate significantly higher transaction amounts (*see* panel (b) in Figure 1). Since we are dealing with true outliers, we will keep them in the dataset to preserve the integrity of the analysis.

2.2 Feature Engineering

The online payment department has informed the data scientists team that many transactions fail on the first attempt. Consequently, customers often try several times to transfer money, in hopes of successfully completing their payments. This pattern is reflected in the dataset where two transactions occurring within a one-minute window, with the same amount and from the same country, are payment attempts for the same purchase. To account for this reasonable hypothesis in the modeling efforts, a new feature '`attempt_count`' is created to capture the total number of attempts made for the same payment under the specified conditions.

Tab. 2: Distribution of the number of attempts

Number of attempts	Count	Proportion
1	27,337	54.23%
2	12,323	24.45%
3	5,712	11.33%
4	2,656	5.27%
5	1,253	2.49%
> 5	1,129	2.24%
Total	**50,410**	**100.00%**

The newly engineered feature reveals that customers may make up to ten attempts to complete an online transaction. This is an indication of a very low success rate, as discussed later. To mitigate the high occurrence of duplicate transactions within the dataset, a limit of five attempts per transaction is set. Any records reflecting more than five attempts are removed from the dataset. This exclusion has minimal impact on the overall dataset in terms of loss of useful information, especially as the proportion of transactions with more than five attempts constitute only about 2.24% of the total entries (*see* Table 2).

As a final operation in this part of the project, two critical variables will be extracted from the 'timestamp': (i) the 'hour' of the transaction, and (ii) the 'day' of the transaction. These features are expected to significantly influence the routing decisions to a specific PSP and, consequently, the likelihood of transaction success. For instance, transaction patterns may vary by hour due to factors such as the availability of certain PSPs during peak shopping times. Similarly, the day of the week may exhibit distinct trends related to consumer spending habits, which can affect transaction outcomes. The incorporation of these new variables aims to enhance the model's predictive capability by capturing temporal dynamics. Thus, their contribution will be properly evaluated during the modeling phase, helping to determine how effectively they improve the overall performance of the ML model.

3 Exploratory Data Analysis

The primary objective of this project is to enhance the online payment process for the company "PaySmart" by automating the credit card routing to payment service providers through predictive modeling. To achieve this, the online payment department has provided the data scientists team with a comprehensive dataset, which has undergone several preprocessing operations to ensure its readiness for analysis. Tables 3 and 4 present a detailed statistical summary of the cleaned data, offering insights into both qualitative and quantitative variables.

Tab. 3: Statistical summary of qualitative variables

PSP	**freq**	**country**	**freq**	**card**	**freq**
UK_Card	52.31%	*Germany*	59.90%	*Master*	57.56%
Simplecard	24.73%	*Switzerland*	20.51%	*Visa*	23.10%
Moneycard	16.55%	*Austria*	19.59%	*Diners*	19.34%
Goldcard	6.41%				
day	**freq**	**attempt_count**	**freq**		
2	16.99%	*1*	55.47%		
3	16.24%	*2*	25.01%		
4	15.48%	*3*	11.59%		
1	14.42%	*4*	5.39%		
5	13.70%	*5*	2.54%		
6	12.46%				
7	10.71%				
3D_secured	**freq**	**success**	**freq**		
0	76.09%	*0*	79.68%		
1	23.91%	*1*	20.32%		

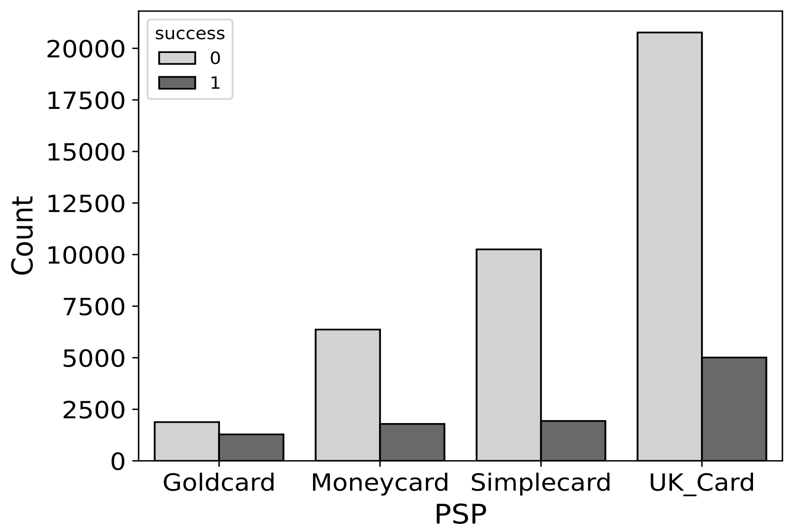
The vast majority of PaySmart customer transactions are completed using either Visa or Master card, collectively accounting for over 80% of total transactions. Additionally, nearly 60% of these transactions originate from Germany, indicating a strong regional concentration. Notably, less than a quarter of customers utilize 3D-identified cards, which are generally associated with enhanced security features. The average transaction amount stands at €202.3, with a standard deviation of €±96.2, suggesting a diverse range of transaction sizes. Transactions are predominantly completed around 12:30 p.m., with a standard deviation of ±6.9 hours, highlighting peak transaction times. Alarmingly, the overall success rate of online payments is only 20.3%, a level that signals significant challenges for both customer satisfaction and company profitability.

Tab. 4: Statistical summary of quantitative variables

	amount	**hour**
mean	202.29	12.50
std	96.23	6.92
min	6.00	1.00
25%	133.00	6.00
50%	201.00	13.00
75%	269.00	18.00
max	630.00	24.00

Focusing on the transaction success rate per PSP, Figure 2 illustrates a troubling trend: failed transactions vastly outnumber successful ones across all PSPs. This disparity is most pronounced for the "UK Card," where failed transactions significantly exceed successful ones. In contrast, the "Gold Card" demonstrates a somewhat better performance, with a higher proportion of successful transactions with respect to failed ones. This raises the question: what factors contribute to these varying success rates? Preliminary analysis suggests that the likelihood of a successful transaction increases when customers use 3D-secured cards. Table 5 confirms this trend, showing that success rates improve across all PSPs when a secure card is used, with the "Gold Card" experiencing a notable jump from 34.85% to 58.37% for 3D-identified transactions.

Fig. 2: Transaction outcome per PSP



Tab. 5: Comparison of Success Rates across PSPs by Card Security

	non-secure card	**secure card**
Goldcard	34.85%	58.37%
Moneycard	21.09%	24.64%
Simplecard	14.72%	19.48%
UK_Card	18.35%	22.91%

To assess the statistical reliability of the relationship between 3D card identification and transaction success rates, a chi-square test is conducted. This test evaluates the independence of the '`success`' variable in relation to the remaining qualitative variables one after the other. Next, the same test is applied to analyze the relationship between the '`PSP`' variable and these same qualitative variables. The results, reported in Table 6, reveal that the null hypothesis of independence is rejected between '`success`' and '`3D_secured`'. This evidence supports the intuition that utilizing a secure card maximizes the chances of successful online payments. Furthermore, the variables '`card`' and '`day`' also exhibit statistically significant relationships with the success rate. With regard to the '`PSP`', the only significant relationship identified is with the variable '`attempt_count`'. In other words, the number of attempts on an operation varies remarkably based on the chosen PSP. These insights will be helpful in interpreting the outputs of the Machine Learning (ML) models to be developed in section 4.

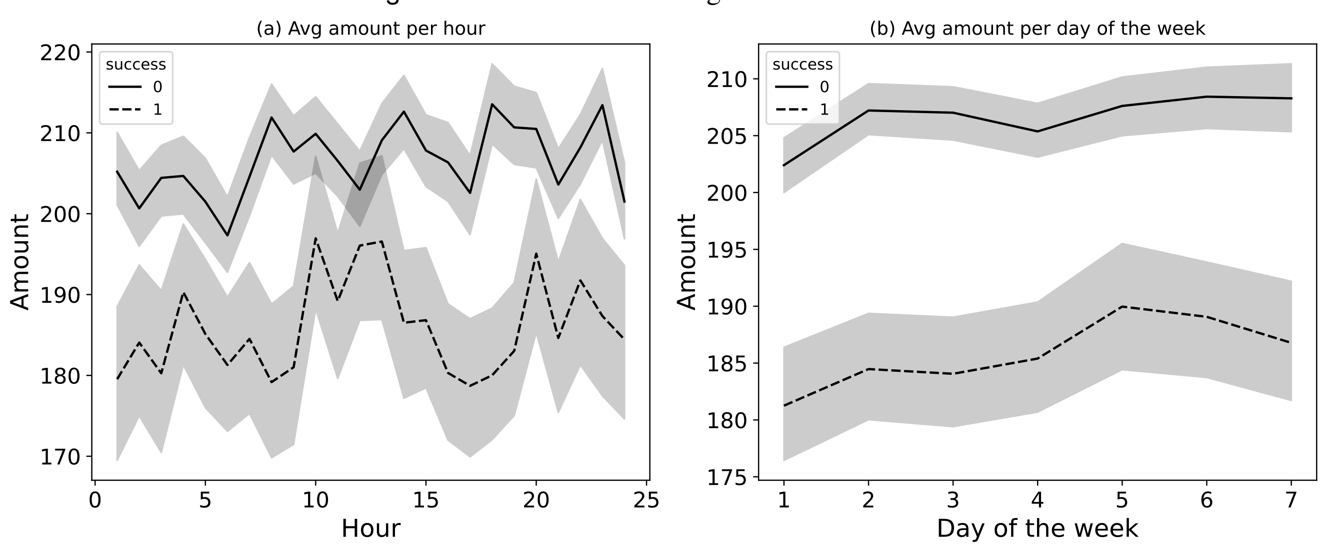
Tab. 6: Chi-square test results

	Chi-sq stat for '`success`'	**Chi-sq stat for '`PSP`'**
country	0.904008	7.183437
3D_secured	180.18628***	0.733151
card	42.18376***	15.784259**
attempt_count	9.107492*	426.800346***
day	50.864021***	33.339523**

Note: *, **, and *** denote significance at 10%, 5% and 1% respectively.

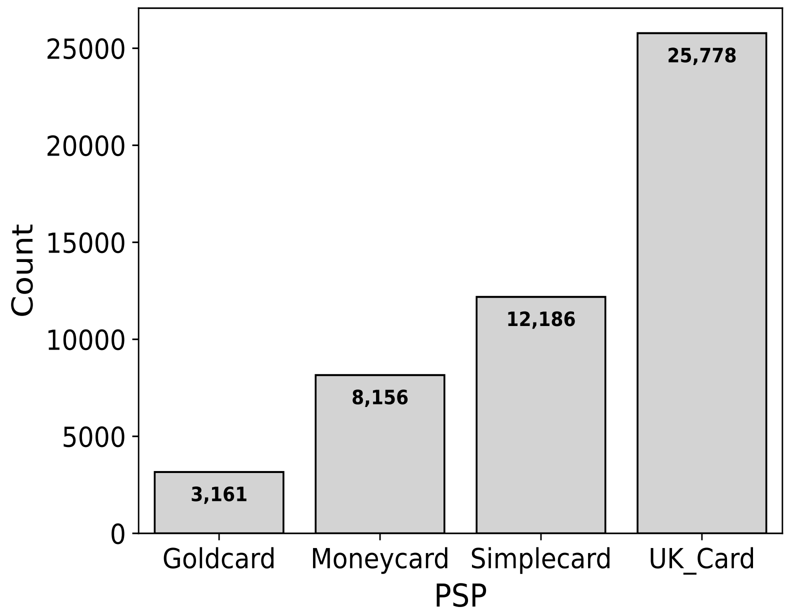
We now turn our attention to the two time-based features created in subsection 2.2. It has been assumed that transaction volume is likely influenced by the availability of certain PSPs during peak hours, as well as customer spending habits on specific days of the week. This hypothesis is visually represented in Figure 3, which illustrates that the average volume of failed transactions declines between 10 a.m. and 3 p.m., while the volume of successful transactions increases during this time frame (*see* panel (a)). This suggests a higher propensity for transactions to succeed in this time slot. Conversely, during weekends, the average volume of successful transactions decreases, while the average volume of failed transactions continues to rise (*see* panel (b)). So, consumers are less likely to engage in online payments during this period due to lower success probabilities.

Fig. 3: Evolution of the average transaction amount



As we prepare for modeling, it is crucial to analyze the balance of classes within the target variable ‘PSP’. In Machine Learning, ensuring class balance is essential to avoid any concerns that could undermine the model’s generalization ability (Kuhn and Johnson 2013). Figure 4 reveals that more than half of the instances in the dataset are concentrated in the “UK Card” class, indicating a significant class imbalance. To tackle this issue, a random resampling technique will be implemented to reduce the bias towards the majority class, thereby fostering a more equitable representation of all PSPs in the dataset.

Fig. 4: Classes balance of the target variable



4 Machine Learning modeling

The dataset provided to the data scientists team includes several qualitative features, which necessitate careful handling since ML models are generally unable to process non-numeric variables directly. To address this, a numerical assignment technique is employed for encoding these categorical variables. This method is preferred over one-hot encoding due to its lower memory usage and reduced computational power, making it particularly suitable for larger datasets (Cerda et al. 2018).

To ensure robust model training and evaluation, the dataset is split into two distinct partitions: (i) the training set (80%), and (ii) the test set (20%). To mitigate class imbalance issues prevalent in the target variable, the Synthetic Minority Oversampling TEchnique and Edited Nearest Neighbours (SMOTEENN) resampling method is implemented on the training set. This method allows to create a more balanced distribution of classes (Lemaître et al. 2017), thereby enhancing the model’s ability to learn effectively from the data.

As a preliminary approach, we begin modeling with a multi-nominal logistic regression, which serves as the benchmark model. The results of this initial model are then compared against those of more sophisticated algorithms, such as Random Forest and Gradient Boosting, to determine if they can achieve better predictive accuracy. Ultimately, the final model is selected based on its performance metrics evaluated on the test set.

4.1 Baseline model

Multi-nominal logistic regression is established as the benchmark model. This model is used to assess the performance improvement of more advanced models compared with this reference base. Before training the model, standard scaling is applied to the feature '`amount`', which exhibits a wide range of values (as detailed in Table 4). The regression analysis results are reported in Table 7, and offer useful insights into how various features affect the likelihood of routing transactions to different payment service providers.

Tab. 7: Logistic regression results

	Class = Moneycard		Class = Simplecard		Class = UK_Card		Class = Goldcard	
Feature	**Coefficient**	**Odds Ratio**	**Coefficient**	**Odds Ratio**	**Coefficient**	**Odds Ratio**	**Coefficient**	**Odds Ratio**
success	1.381707	3.981692	-0.347873	0.706189	-1.041127	0.353056	-0.454034	0.635061
scaled_amount	0.000000	1.000000	0.000000	1.000000	-0.068618	0.933684	0.099194	1.104281
hour	-0.003000	0.997005	0.002753	1.002757	0.001977	1.001979	0.004917	1.004929
day	-0.030369	0.970088	-0.012328	0.987748	-0.012515	0.987563	0.062787	1.064800
country	-0.052385	0.948964	-0.099315	0.905457	0.000000	1.000000	0.226790	1.254567
card	-0.215810	0.805888	-0.156310	0.855294	0.000000	1.000000	0.433720	1.542987
attempt_count	-0.419576	0.657325	-0.343506	0.709279	-0.339184	0.712351	0.969567	2.636804
3D_secured	-0.299405	0.741259	-0.000677	0.999323	0.000000	1.000000	0.345793	1.413110

Overall, the regression results reveal significant differences in how different factors influence the routing of transactions to various PSPs, indicating that customer preferences and transaction dynamics are complex and multifaceted. Notably, the usage of secure cards emerges as a critical predictor; while the negative coefficient for Moneycard indicates a reduced likelihood of its use when a secure card is employed, Goldcard shows a positive association, suggesting that 3D-secured cards increase its selection probability. The transaction amount has limited impact on most PSPs; however, it does favor Goldcard. The count of attempts is particularly influential, indicating that Goldcard is more likely to be selected after multiple attempts, possibly reflecting its perceived reliability. Conversely, the hour and day of the week have odds ratios close to one across all classes, indicating negligible effect on PSP choice. Moreover, the historical success of a transaction is a strong predictor for Moneycard, while UK card exhibits the opposite trend. Understanding these relationships can help PaySmart to refine its routing strategies, ultimately enhancing customer satisfaction and improving transaction success rates.

To assess the performance of this baseline model, the ROC AUC score is used. This metric is particularly suited for multi-class classification tasks involving imbalanced data (Bradley 1997). When confronted with the test set, the logistic regression achieved a performance of **54.36%**. While this score falls short of ideal performance, it serves as a foundation for exploring more sophisticated modeling techniques capable of achieving a much better performance, which are discussed in subsection 4.2.

4.2 Improved models

This subsection delves into more advanced ML models that are better equipped to capture the complex relationships inherent in the data and accurately predict the appropriate PSP for online operations. While enhancing predictive performance, these complex models may sacrifice some interpretability (Hastie et al. 2009), which remains a key strength of the benchmark model. The three models selected for training at this stage are: (i) K-Nearest Neighbors (KNN), (ii) Random Forest (RF), and (iii) Gradient Boosting (GB). As argued by Chakravorti (2003) and Hemkiran et al. (2022), these models have demonstrated strong predictive power in financial applications.

To optimize the learning process, hyper-parameters fine-tuning is essential for each model. For KNN, the optimized hyper-parameters are the number of neighbors, the power of the Minkowski distance, and the weighting function. Regarding the RF model, the hyper-parameters are the number of trees and the criterion for splitting. In the case of GB, the number of estimators and the learning rate are prioritized for tuning. This optimization process employs 5-fold cross-validation to ensure model robustness. After training on the learning dataset, the models are evaluated on the test dataset. The performance metrics for each model are summarized in Table 8.

Tab. 8: Performance scores of advanced models

	ROC AUC	**F1-score**	**Accuracy**
K-Nearest Neighbors	50.66%	27.87%	25.01%
Random Forest	53.89%	35.89%	33.45%
Gradient Boosting	54.88%	38.71%	37.71%

As shown in Table 8, the Gradient Boosting model achieves the highest performance metrics among the tested models. Consequently, it is selected as the final predictive model for routing transactions to the PSP that maximizes the chances of success. However, it is important to note that not all sophisticated models yield exceptional scores during the test phase. Indeed, in the dataset provided, the 'PSP' column indicates the payment service providers selected based on the pre-existing static routing rule. The aim of the modeling effort is to rectify inconsistencies in this manual routing method to enhance the overall success rate. It is therefore obvious that the classes predicted by the model may differ from the classes observed in the test set. These discrepancies explain the low performance scores of the models, and should not be viewed as a major concern.

4.3 Final model

Gradient Boosting is the definitive model for this project. As an ensemble method, it combines multiple weak learners so that the overall performance of the aggregated estimator is maximized. However, the complexity inherent in this model makes its outcomes difficult to interpret, i.e. the decision rule is like a black box. Since the GB model is not intrinsically interpretable, the importance of each feature will be examined to understand how individual predictions are made by the model (Breiman 2001).

Fig. 5: Features importance plot

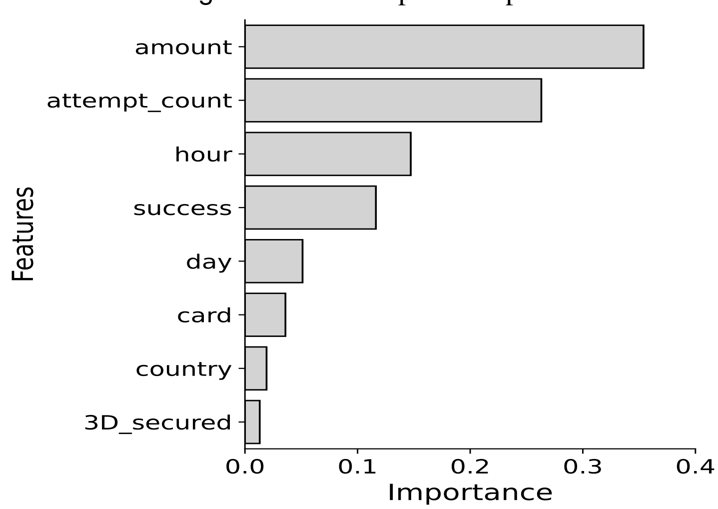


Figure 5 illustrates the ranked importance of features contributing to the model's predictive power. The most influential features, in descending order of importance, include: (i) the transaction amount, (ii) the number of attempts, (iii) the hour when the payment is made, and (iv) the history of past successes. Unlike the baseline model, the use of a secure card has a marginal effect on the final model's predictions. Additionally, the card category and the country of origin are found to have little impact on the choice of PSP.

5 Business implications and Deployment

This section explores the benefits of implementing a Machine Learning model to automate the routing of online transactions to the relevant PSP. The discussion focuses on two primary dimensions: (i) customer satisfaction, reflecting the value proposition, and (ii) company profitability for success definition. On one hand, the model is expected to generate considerable value for customers by reducing the average cost of transactions. A decrease in transaction fees directly translates to more affordable services, enhancing customer satisfaction and loyalty. On the other hand, the company stands to benefit from increased transaction success rates following the deployment of the predictive model, which is crucial for driving profitability. Throughout this section, we closely monitor these two key performance indicators (KPIs) to assess the model's impact on both customer experience and business outcomes.

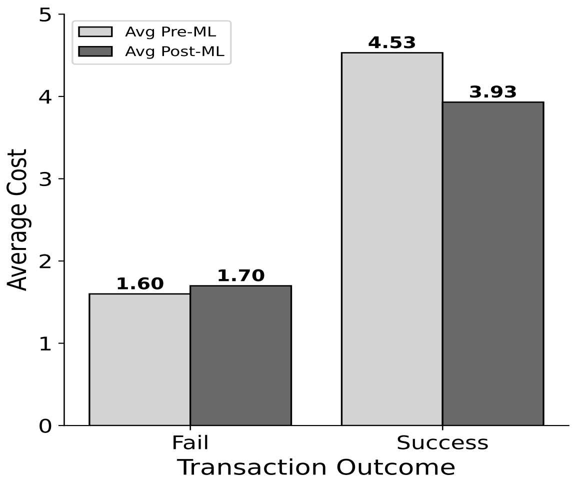
Given the unit fee structure for transaction services at PaySmart (as outlined in the task description), it is possible to compute the global costs associated with successful and failed transactions prior to the model's implementation, by analyzing the number of successful/failed transactions per PSP. After deploying the model, the same process can be performed by using the predicted classes for counting the number of transactions. This will facilitate a comprehensive evaluation of the average cost per transaction pre- and post-ML implementation, differentiated by transaction outcomes (is the payment successful or not?).

Most of the time, the predicted classes do not align with the observed classes, a result that was expected as the model is designed to rectify the routing errors of the manual system that prevailed in the company. The analysis reveals that out of 9,857 instances in the test dataset, approximately 62.3% correspond to cases where the model's predicted class diverges from the observed class. This indicates that nearly two-thirds of the PSP selections made under the static routing rule were incorrect. We therefore postulate that for all instances where the predicted class differs from the observed class, the transaction is more likely to succeed. In other words, since the model has been adequately trained, the predicted PSP should reliably guarantee the completion of any online payment. This corrective functionality for assignment errors is reflected in the "success rate" KPI, as shown in Figure 6.

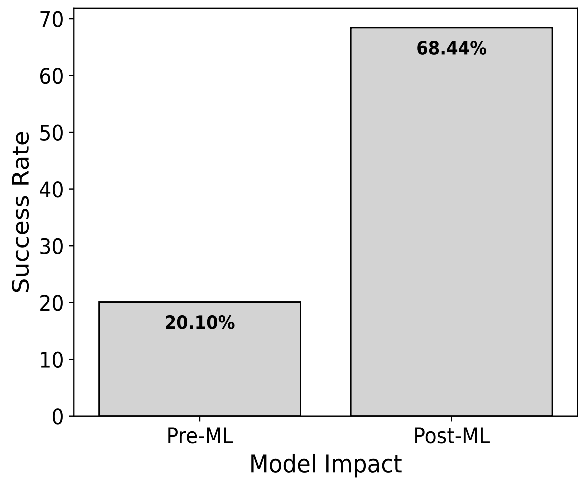
Panel (a) demonstrates a significant reduction in the average cost of successful transactions, which has decreased from €4.53 to €3.93 following the model's implementation, i.e. an impressive drop of 13.25%. This reduction in service fees leads to substantial savings for customers, and fosters increased satisfaction and loyalty towards PaySmart's services offerings. Panel (b) illustrates a dramatic increase in the success rate, which rises from 20.10% to 68.44% after the deployment of the predictive model. This significant jump is likely to reduce the risk of losing clients to competitors and ultimately contributing to higher profitability for the company. Summing up, the deployment of the predictive model developed by the data scientists team presents a dual advantage: it enhances the customer experience by lowering transaction costs while simultaneously boosting the company's success rate and profitability.

Fig. 6: Benefits of the predictive model

(a) *Average cost per transaction*



(b) *Succes rate dynamics*



6 Conclusion

The development of a predictive model for dynamic routing of online credit card transactions at PaySmart represents a significant advancement in addressing the challenges posed by high transaction failure rates and associated costs. By transitioning from a static, rule-based routing system to a data-driven approach, PaySmart can enhance transaction success rates, thereby improving customer satisfaction and fostering loyalty.

The comprehensive analysis conducted throughout this project has illuminated critical factors influencing payment outcomes, such as transaction amounts, the number of attempts, and the 3D-identification of cards. These insights have been harnessed to develop a robust machine learning model that effectively predicts the most suitable payment service provider for each transaction. The results demonstrate a remarkable increase in the success rate of payments — from 20.10% to 68.44% — alongside a significant reduction in transaction costs, marking a 13.25% decrease in average service fees.

The implications of this predictive model extend beyond mere cost savings; they encompass a strategic enhancement of PaySmart's competitive positioning within the retail landscape. By leveraging advanced analytics, PaySmart not only mitigates the risk of customer attrition to competitors but also capitalizes on new opportunities for profitability. This dual advantage underscores the importance of integrating machine learning into financial processes, ultimately reinforcing the company's commitment to delivering exceptional value to both customers and stakeholders.

As PaySmart prepares for the full-scale deployment of the predictive model, it will implement a phased rollout strategy, ensuring thorough testing in live environments while monitoring system performance and user feedback. This approach will allow for real-time adjustments and optimizations based on evolving transaction patterns and customer behaviors. Additionally, ongoing training and support for staff will be crucial to facilitate a smooth transition and maximize the model's effectiveness. By maintaining this proactive stance and continuously refining the model, PaySmart can ensure sustained success and innovation in its online payment processing capabilities.

References

Bradley, A. P. (1997). The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning algorithms. *Pattern Recognition*, 30(7):1145–1159.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

Cerda, P., Varoquaux, G., and Kégl, B. (2018). Similarity encoding for learning with dirty categorical variables. *Machine Learning*, 107:1477–1494.

Chakravorti, S. (2003). Theory of Credit Card Networks: A Survey of the Literature. *Review of Network Economics*, 2(2).

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd edition.

Hemkiran, S., Sadasivam, G. S., Rahavendra, A. P., and Anjhanna, A. K. (2022). Design of Automatic Credit Card Approval System using Machine Learning. In *Artificial Intelligence and Technologies*, Lecture Notes in Electrical Engineering, pages 1–9. Springer.

Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling*. Springer.

Lemaître, G., Nogueira, F., and Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*, 18:1–5.