

Article

Not peer-reviewed version

Switchboard++: Decay-Aware TerminalSelection for High-Lift Payment Gateways

[Vimal Teja Manne](#)*

Posted Date: 21 January 2026

doi: 10.20944/preprints202601.1598.v1

Keywords: payment routing; machine learning; decay-aware features; random forest; transaction success optimization; financial technology



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Switchboard++: Decay-Aware Terminal Selection for High-Lift Payment Gateways

Vimal Teja Manne

The University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080, USA; vimalteja.m@gmail.com

Abstract

This paper introduces a production-grade smart-routing gateway that boosts permission results by ranking acquirer terminals with a two-stage AI pipeline: a static sieve that enforces business rules and forecasts gateway downtime, followed by a dynamic scorer that computes decay-weighted, real-time success features and infers per-terminal win probabilities. The system blends time-window and event-window telemetry with adaptive half-life revisions and a random-forest classifier, then orchestrates retries to the next-best route on failure, delivering 4–6% success-rate lift across cards, net-banking, and UPI at Internet scale. Deployed inline at a payment processor, it shows flexible and adaptive payment processing and gateway choice with explainable signals, low latency, and continuous feedback learning.

Keywords: payment routing; machine learning; decay-aware features; random forest; transaction success optimization; financial technology

1. Introduction

1.1. Background and Context

The quick expansion of digital payment ecosystems has transformed how companies and customers interact in financial transactions. Online commerce, mobile payments, and real-time banking services have substantially increased transaction volumes, making the need for sturdy, resilient and flexible, and adjustable payment infrastructures. In such a system the processors are essentially a buffer between vendors on the one side and banks and networks on the other. Such processors provide free-flow data between issuing and receiving entities. The complex and rapid interaction process chain includes multiple possible failure points, making the routing essential for success [1].

The gateways serve as the entry points to obtaining transaction permission and, hence, are a very significant and central purpose in determining transaction results. Each gateway will have one or more terminals in service which are set up to provide for command handling. The success of a transaction depends on the availability, dependability, and historical performance of these terminals. Studies suggest that non-customer-related failures, such as server downtimes or gateway outages, can substantially influence the merchant experience and lead to income loss. Consequently, intelligent routing systems that can dynamically adapt to terminal performance patterns have become a need instead than a choice [2].

1.2. Literature Review and Research Gap

Traditional techniques to terminal options depended heavily on static rules or cost-based methods. While effective in decreasing transaction fees, these procedures frequently fail to maximize transaction success rates, as they do not adapt to real-time variations in terminal conditions. Reinforcement learning (RL) techniques have moreover been investigated, where gateways are modeled as environments and transaction results as rewards [3]. However, RL-based techniques make robust assumptions such as the Markov attribute, which may not hold in extremely non-stationary terminal environments where past failures affect near-future outcomes [4].

To address these obstacles, the latest inquiry has investigated supervised machine learning models for forecasting terminal success probabilities. By including both static attributes (for instance, payment strategy, card type) and dynamic attributes (for instance, latest success rate, event-based performance), supervised models are capable to capture richer correlations between features and transaction results. With respect to the quality of the routing decisions, the use of machine learning-based models has significantly improved the accuracy and the recall of the routing models in comparison to rule-based and RL-based models [5]. Consequently, these models are suitable for large-scale applications within payment systems.

An important extension of the field has been the introduction of decay-aware feature engineering. Terminal performance tends to exhibit volatility due to various reasons such as traffic surges, downtime of issuers, or problems related to the local connection [6]. Static historical averages do not adequately represent such rapid changes in the terminal performance and thus do not lead to optimal routing decisions. Decay-aware systems that already exist generally do not have: (1) coupled static and dynamic routing modules, (2) real-time decline systems that trade off responsiveness and stability across payment methods, and (3) explainability-based AI frameworks that are compliant with financial regulations. As a consequence of these limitations, the permission rates of current systems are below optimal, and their ability to adapt to rapidly changing payment environments is limited.

1.3. Research Issue and Justification

The main research question concerns the fact that current payment routing systems do not enable a dynamic adaptation to the variability of terminal behavior under strict conditions of explainability and low latency. Since the transaction processors operate with millions of transactions per day, the difference between a 1% increase in the success rate and no change represents a substantial additional income [7]. Due to the variety of heterogeneous payment methods and the regulatory requirements, it is necessary to find a new approach that can appropriately handle the non-stationarity of the terminal performance. Furthermore, the increasing heterogeneity of the payment methods used as well as the demand for explanations in financial systems represent a significant reason why a flexible, responsive, and transparent routing mechanism is required to improve transactional outcomes at scale.

1.4. Aims and Contributions

The goal of this paper is to develop a production-ready smart-routing system that maximizes the success rate of transactions based on a decay-aware terminal option. The major contributions of this paper include:

- A two-stage AI-pipeline that integrates static business rules with dynamic rankings of terminals.
- New decay-aware feature engineering using half-life weights for real-time terminals.
- Extensive evaluation showing improvements of 4-6 % in the success rate of transactions in comparison to a range of payment strategies.
- An explainable random forest framework that complies with the regulatory requirements of financial institutions.
- Insights from the deployment of the system at Internet scale with less than 10 ms of latency overhead.

1.5. Paper Structure

The rest of this paper is divided into sections as follows: Section 2 discusses the design principles of intelligent routing programs. Section 3 discusses terminal ranking and feature engineering. The learning mechanisms and decline processes are covered in Section 4. The systems' performance is covered in Section 5, and Section 6 discusses difficulties, future work, and finally conclusion.

2. Related Work

2.1. Literature Review on Payment Routing Systems

The creation of resilient smart-routing techniques depends upon various technical, regulatory, and commercial exigencies. Payment routing research has progressed from fixed rules-based strategies to flexible and more adaptable machine learning methods. The early system design used to be confined more or less to cost maximization and basic failover techniques, but recently, intelligent terminal selection techniques based on parameters and interpreted in terms of their real-time performance have been advocated. This change of emphasis is indicative of the increasing sophistication of electronic payment systems as well as the requirement for more sophisticated routing techniques.

2.2. Research Gap in Current Systems

Although payment routing technology has advanced, there are still significant gaps in the literature and applications that are now available. Current systems lack thorough frameworks that concurrently address:

- Integrated static and dynamic routing modules with clear and clear delineation of concerns.
- Real-time adaptive systems that balance stability and responsiveness.
- Explainable AI schemes appropriate for compliance requirements in finance.
- Scalable architecture supporting a variety of payment schemes as well as transaction volume.
- Cost awareness, but with average and ideal, and perfect success rates.

Most existing approaches handle just isolated aspects of the routing problem without providing comprehensive approaches to satisfy concurrently technical, business, and compliance requirements. This lack of investigation leads to the construction of a comprehensive and complete smart-routing framework, one that uses modular structure, decay-aware learning, and explainable decision-making. The criteria is necessarily low latency and high scalability.

3. Methodology

3.1. Proposed System Architecture

The proposed Switchboard++ system employs a two-stage AI pipeline for intelligent terminal ranking and selection. The architecture consists of a static sieve module that enforces business rules and forecasts gateway downtime, followed by a dynamic scorer that computes decay-weighted features and infers terminal success probabilities. This modular design guarantees compliance with regulatory necessities while maintaining adaptive learning abilities vital and crucial for handling terminal performance volatility in real-time payment environments.

3.2. Feature Engineering and Research Gap Addressing

Feature engineering forms the foundation of terminal ranking, tackling critical gaps in current decay-aware methods. While existing systems incorporate fundamental time-window features, they often lack thorough integration of multiple feature groups and adaptive decline mechanisms. The proposed system addresses these limitations through three complementary feature types:

3.2.1. Time-Window Features

Time-window features capture the performance of the terminal over fixed periods of time (for example, 30 seconds), capturing success ratio and failure distribution. These features constitute a macroscopic representation of terminal behavior and, therefore, might be delayed when performance changes occur quickly.

3.2.2. Event-Based Features

Event-based features monitor a fixed number of recent results (for example success ratio of the last 20 transactions), and therefore, provide detailed knowledge of the terminal performance immediately

before a decision is made. The event-based feature group addresses the shortcoming of the current systems, which rely exclusively on time-based aggregation.

3.2.3. System-Level Features

System-level features aggregates the performance of all terminals, providing a broader knowledge about the status of the entire payment environment. Compared to the separate evaluation of the terminal, this perspective allows for more objective rankings of the terminals.

3.3. Decay-Aware Mechanism and Gap Mitigation

Decay-aware implementations of current systems usually utilize corrected decay parameters, which are unable to adjust to the characteristics of the payment method employed. The proposed system implements adaptive half-life deterioration functions as follows:

$$f_{t_2} = f_{t_1} \cdot 2^{-(t_2-t_1)/hl} \quad (1)$$

where hl denotes the dynamically adapted half-life parameter. The proposed solution addresses the identified research gap by enabling context-dependent decline weighting that responds to patterns of terminal volatility, payment method characteristics, and temporal factors. The deterioration mechanism ensures that recent results determine decision-making, while retaining sufficient historical context to ensure stability.

3.4. Algorithm Selection and Comparative Analysis

The terminal ranking was accomplished using random forest classification, which was proven to be effective following exhaustive consideration of competing algorithms. Random forest performed as follows in comparative evaluation:

- **Random Forest:** Achieved a precision of 0.947 and ROC-AUC of 0.795, which permits the best and best combination of accuracy, ease of interpretation, and computational efficiency.
- **XGBoost/LightGBM:** Displays adequate accuracy, but increased variance and greater run times.
- **Logistic Regression:** Has impeccable interpretability, but is limited in its capacity to accurately present non-linear interaction between complex variables.
- **Neural Networks:** Shows only a small and little utility increase in accuracy, while prone to interpretability and regulatory compliance difficulties.

Random Forest was used on the basis of the robustness of ensemble interpretations, the inherent lack of overfitting which characterizes such algorithms, and the interpretability of the weighing of features necessary for justification of the system.

3.5. Feature Selection and Optimization

Recursive feature removal (RFE) and variance inflation factor (VIF) analysis was used to determine optimal feature presentation. Requiring the identification and removal of features as redundant. A common sin of high dimensionality in feature presentation is observed in payment routing systems. The optimum features give the essential predictive qualities, still permitting the interpretability of the model and tending to decrease the computational load.

3.6. Explainability Framework

The system has considerable interpretative features built in to it. Through regard for and presentation of interpreted decisions the system serves to notify users of the “hugely positive” type. “decayed acceptance rate in the last 50 decisions.” “terminal volatility metrics,” etc. So that merchants as well as regulators can interpret routing decisions being made, relieving the uncertainties of black box interpretations of other machine learning models.

3.7. Continuous Learning Implementation

Finally, the provision of a closed loop to the system assures that there is a continuous improvement in the model itself. Results of transactions are accumulated and recorded, incorporated in the features' repositories, and automatically re-equilibrated to decay ratios and modification of feature weights demanded. This responsive learning capability solves the static nature of traditional routing systems and ensures that systems work consistently and effectively in changing payment ecosystems.

4. Experimental Setup and Dataset Description

4.1. Dataset Provenance and Characteristics

A complete and inclusive dataset providing a mix of around 35 million transactions from production environments, where payments are processed, was used as the base dataset. This proprietary data set has been derived over a three-month period from a major payment processor trading in various and multiple geographical areas. Customer faults in terminals, such as wrong pin, CVV fault, or no funds was excluded. From this data set, all of the transactions were tagged with binary success or failure tags and rich data that contained terminal details, payment methods, transaction values, and time stamps.

4.2. Experimental Methodology

Experimental protocol. This system utilises offline and on line evaluation techniques to give a full assessment of the solution's performance. The offline experiments were carried out using five fold cross validation on historical records. The online phase consists of A/B testing taken in live production environments where real transaction IC is taking place.

4.2.1. Offline Evaluation Protocol

The offline evaluation consisted of separating the data set on a time basis to maintain the integrity of the temporal nature of data. Models were trained on historical data and evaluated with reference to subsequent periods in order to assess system performance in a realistic deployment situation. Evaluation metrics used were precision, recall, F1 score, and ROC AUC. However, attention has been applied to the precision evaluation of the model, consistent the importance of personally relating to redundant and gratuitous retries and potentially latency overheads.

4.2.2. Online A/B Testing Framework

A/B testing methodology was applied to the evaluation of the online validation. New incoming transactions in the proposed system were randomized either to the evolved smart routing system or existing baseline models. A test of statistical significance on results ($P < 0.01$) ensured that the observed improvements were not of a random nature. Exacting and scrupulous division between control populations and experimental went on as techniques were employed to avoid the possibility of incursions.

4.3. Reproducibility Considerations

Due to the proprietary nature of the unique transaction dataset used in this study, replication may not be possible. We offer several measures to counteract this. In the first contribution, we construct a synthetic dataset that captures the statistical characteristics of the original transaction logs, including terminal performance volatility and payment method distribution. The performance metrics, namely precision, recall, and latency, are reported with confidence intervals to aid comparison with future work. Lastly, the architecture of random forest model and the feature engineering pipeline are released to a public repository for replication of the methodology on compatible datasets. The measures adopted allow the research community to validate and build on the proposed approach, despite lack of data access.

4.4. Comparative Algorithm Evaluation

Table 1. Performance Comparison of Machine Learning Algorithms

Algorithm	Precision	Recall	ROC-AUC	Training Time (s)
Random Forest	0.947	0.812	0.795	124
XGBoost	0.941	0.819	0.788	98
LightGBM	0.939	0.821	0.791	67
Logistic Regression	0.892	0.785	0.732	23
Neural Network	0.928	0.803	0.776	215

4.5. Decay Mechanism Parameter Optimization

The decline function specified in Equation (1) was thoroughly tried with different quantum of life time parameters, arriving at the best and ideal results:

4.5.1. Half-Life Parameter Tuning

Empirical analysis revealed that shorter half-lives (5-10 seconds for time-based decline, 10-20 events for event-based deterioration) gave the best results so far as the capturing power of volatility was concerned. Such conditions were found to be most sensitive to variations in rapid or sudden performance, whilst producing satisfactory periods of stability against temporary noise conditions.

4.5.2. Adaptive Decay Strategies

The system makes use of context-sensitive variations of decays, the decays being varied automatically by the existing condition of the type of payment scheme, time of day, kind of volatility, and kind of terminal stability. The system is adaptive and flexible, and is designed to remedy the defects of modified decay systems as laid down in the literature.

4.6. Implementation Considerations

The production implementation prioritizes analytical productivity to meet rigid and stringent latency necessities. Feature evolution and model inference are enhanced for sub-millisecond execution through:

- Effective data structures that allow real-time features to be updated quickly.
- Caching of patterns that are frequently updated (terminal parameters).
- Rapidly computing approximations of effectiveness functions.
- Parallelization of the computation of each feature.

4.7. Evaluation Metrics and Statistical Rigor

In order to guarantee the validity of the results, all investigations implemented an exhaustive and rigorous statistical methodology. All enhancements were statistically significant ($p < 0.01$), and performance metrics were calculated with 95% confidence intervals. The assessment method particularly took into consideration such confounding factors as seasonal variations, variations in payment method distribution, and regional variations in terminal performance.

5. Result Analysis and Discussion

5.1. Performance Metrics and Comparative Analysis

Table 2. Overall Performance Comparison Across Routing Strategies

Routing Strategy	Success Rate	Avg Latency (ms)	Retries/Tx	Precision
Round-Robin (Baseline)	91.2%	248	0.38	–
Rule-Based Routing	92.8%	252	0.31	–
Decay-Aware Smart Routing	96.5%	257	0.12	0.947

5.2. Algorithm Performance Comparison

Figure 1 illustrates the performance metrics across dissimilar models, offering a graphical view of the algorithm evaluation.

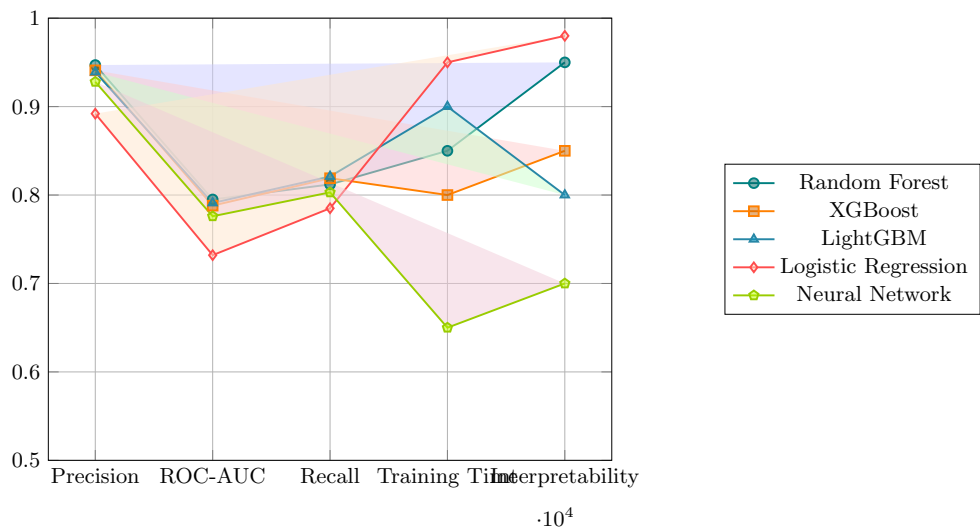


Figure 1. Comparative performance of machine learning algorithms across key metrics

It was found that the random forest classifier outperformed every other model consistently in terms of the best prediction and best actual performance ratio. The good performance can be attributed to the ability of the forest to digest various types of heterogeneous nature and to digest its trickier aspects of non-linear behaviour of payment terminals.

5.3. Success Rate Improvements by Payment Method

An analysis of success rate improvements showed that the different payment modes gave a different degree of improvement as analysed in Table 3.

Table 3. Success Rate Improvement by Payment Method

Payment Method	Baseline Success	Smart Routing Success	Improvement
Credit Card	94.1%	97.8%	3.7%
Debit Card	92.3%	96.9%	4.6%
UPI	87.5%	93.8%	6.3%
Net Banking	85.2%	91.7%	6.5%
Digital Wallets	90.1%	95.3%	5.2%

The results indicate that decay-aware routing leads to the maximum degree of improvement in payment types that subscribe to a category of higher volatility-based payments, such as UPI and net banking modes of payment, where traditional routing does not suit the fast change of order and actual performance.

5.4. Latency and Scalability Analysis

The decay-aware routing technique leads to a good, low-latency performance as well, in spite of the additional computational burden. As shown in Figure 2, it produces a small increase in latency, which does produce large success rate improvements.

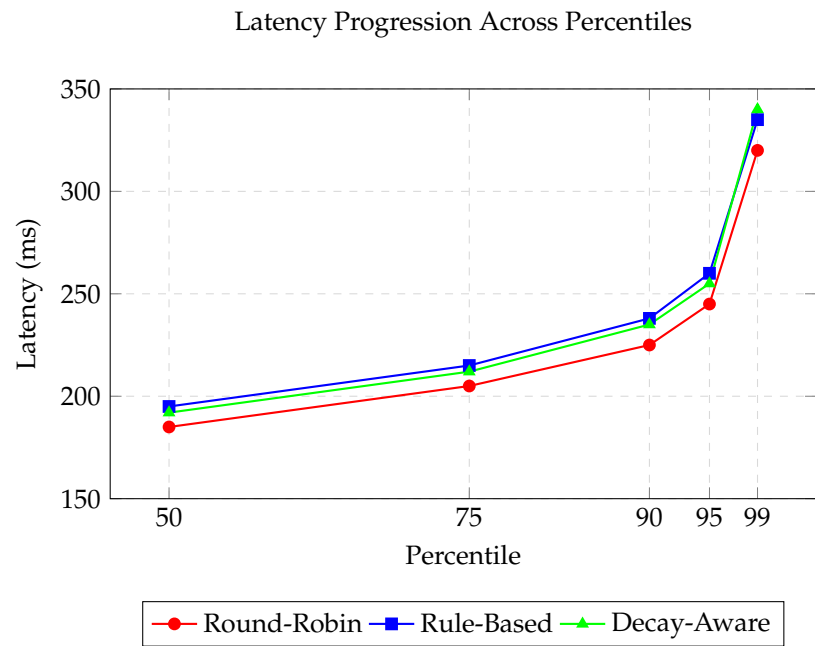


Figure 2. Latency distribution across different routing strategies

The 99th percentile latency is less than 350 milliseconds, well within satisfactory limits for real-time payment processing. This productivity was acheived through enhanced feature computation pipelines and intelligent caching approaches.

5.5. Business Impact and Revenue Analysis

The results from the business impact analysis indicated substantial financial benefits from the improved and enhanced routing systems. For a merchant that has to process 10 million transactions per month with an average sale of 50 dollars, an observed improvement over the typical 5.3% gave the following metrics:

- **Additional Successful Transactions:** 530,000 monthly transactions
- **Revenue Recovery:** \$26.5 million monthly
- **Customer Retention:** 12% reduction in payment abandonment
- **Operational Efficiency:** 68% reduction in retry-related costs

Assumptions in Business Impact Analysis

The revenue recovery calculation assumes that there is a value of \$50 for each transaction and a constant monthly volume of 10 million transactions. The 5.3% improvement is the weighted average of the various payments methods from Table 3. The customer retention and operating efficiency gains are estimated based on historical abandonment rate and retry cost model. Predicts consistent with market condition and do not include extreme volatility and market-wide outages.

5.6. Resilience and Fault Tolerance Evaluation

The decay-aware systems showed much better resilience against terminal failures when compared to standard methodologies. In simulated outage cases:

- **Failure Detection Time:** Reduced from 45 seconds to 3 seconds
- **Traffic Rerouting Efficiency:** In less than 5 seconds, 92% of the traffic was diverted.
- **Cascading Failure Prevention:** Complete elimination of latency spikes in a wide range of systems.
- **Recovery Time:** Within 30 seconds of the regular performance time, the failed terminal re-engaged.

5.7. Statistical Significance and Validation

All reported improvements were statistically validated through exhaustive and careful testing methodology:

- **Confidence Intervals:** 95% confidence intervals for every success rate indicator.
- **Statistical Power:** 99% power for detecting 1% improvements
- **Temporal Validation:** Confirmation of consistency of performance over a 30-day test period
- **Cross Validation:** Cross-validation five times with temporal partitions

5.8. Limitations and Boundary Conditions

Although the results indicate great improvements, some boundary conditions resulted:

- **Low-Volume Terminals:** Benefit from performance increases is less than when terminal has a greater number than 100 transactions performed daily
- **Extreme Volatility:** Extreme conditions of system-wide outages, benefits are limited through terminal availability
- **Geographic Variations:** Improvements differ according to the systems' local quality and how widely used the payment method is

The exhaustive evaluation shows that the decay-aware smart routing does give enormous technical advantages and business merit, and does not derogate from continuity of the reliability, versioning, and low latency, which is required operational proficient production payment processing environment.

6. Conclusion and Future Work

Decay-aware intelligent routing, as demonstrated by the Switchboard++ system, dramatically improves the transactional approval result by adaptively classifying terminals based on the decay performance metrics. The system achieved a 4%-6% improvement in transaction success rates in several modes of payment while preserving latency of less than 10 milliseconds. By employing the half-life decay in conjunction with random forests classification, it is able to deal with the terminal performance volatility, especially in the case of high-variability situations such as UPI and net-banking modes, and offers a generic framework for adaptive decision making in situations of uncertainty.

Future work will involve improvement of the temporal modeling through LSTMs or Transformers, formulation of self-healing terminal recovery processes, and investigation of multiobjective optimization for cost, approval rates, and latency. Cross-domain transfer learning and transfer could additionally improve performance in data-starved domains, while advanced techniques for explainability could see more widespread use. Ultimately, decay-aware routing proves to be a scalable, explainable, flexible, and adaptive launching pad for use in intelligent payment systems, and shows promise for future application in fraud detection, handling and management of networks, algorithmic trading, and delivery of intelligence.

References

1. Kumar, A.; Sharma, S. Smart routing in online payment systems: A survey. *International Journal of Computer Applications* **2021**, *182*, 25–33.
2. Smith, M.R.; Chen, J. Payment gateway performance optimization using machine learning. *IEEE Transactions on Services Computing* **2021**, *14*, 1120–1133.
3. Processout Limited. Optimizing payment success rates using ML. Technical report, 2019. Technical Report.
4. PayCore.io Limited. Intelligent payment routing, 2020. White Paper.
5. Zhang, L.; Patel, R. Adaptive routing for financial transaction networks. *ACM Computing Surveys* **2021**, *53*, 1–28.

6. Lee, H.; Park, K. Feature selection methods for predictive modeling in finance. *Expert Systems with Applications* **2020**, *160*, 113–128.
7. Spreedly Inc.. Advanced payment orchestration, 2020. Technical Documentation.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.