Leveraging Financial Mobile App Interaction Data for Credit Risk and First-Party Fraud Detection

Jyoti Prakash Bal
Revolut Group Holdings Ltd, Madrid, Spain
Global Credit Management - Data Science
Data Scientist
jyoti.prakash@revolut.com

August 29, 2025

Abstract

Traditional credit risk evaluation often relies on historical credit data and financial records, which may not fully capture real-time user behavior and emerging risk patterns. This study proposes a novel approach that leverages user interaction data from mobile and web applications to enhance both credit risk assessment and the detection of first-party fraud, specifically No Intent to Pay (NITP) fraud. By analyzing user interactions up to 45 days before credit applications, we engineer a comprehensive set of behavioral features that reveal patterns linking user app behavior to financial risk.

Key behavioral indicators include session dynamics, device attributes, interaction frequency, sudden activity spikes, user engagement velocity trends and application abandonment trends. Spike metrics capture abrupt increase in user engagement, which may indicate impulsive behavior or heightened interest. Velocity metrics analyze the speed and clustering of user actions, helping to distinguish between organic engagement and suspicious activity. Cancellation metrics track abandoned applications, which can signal indecisiveness or potential risk tendencies.

To validate this approach, we trained both Logistic Regression and LightGBM models, demonstrating improved predictive accuracy in identifying high-risk borrowers and NITP fraud cases. Our methodology is adaptable, supporting various machine learning models, including interpretable approaches that ensure regulatory compliance.

By integrating these behavioral insights, lenders can refine risk assessment processes, reduce fraud losses, and enhance decision-making efficiency. This study underscores the value of digital interaction data in complementing traditional financial data for a more holistic evaluation of creditworthiness and fraud detection.

1 Introduction

In recent years, the financial services industry has experienced a transformative shift driven by the rapid adoption of digital banking platforms and mobile financial applications. With increasing digital engagement, traditional credit risk assessment and fraud

detection methodologies—which largely rely on static data such as demographic information, credit bureau scores, and historical transaction records—face notable limitations. These conventional approaches often lack the granularity and timeliness required to capture evolving customer behaviors and emerging patterns, particularly in fast-paced digital environments.[4]

Furthermore, they may underperform when assessing thin-file or new-to-credit segments, where limited financial history hinders effective risk evaluation. Mobile banking applications generate vast amounts of granular, time-stamped event data that chronicle detailed user interactions: from app navigation flows and session durations to specific user actions such as product views, form submissions, and cancellations. These digital footprints provide a rich behavioral substrate that reflects users' intent, engagement, and decision-making processes in ways static data cannot.[5]

Harnessing such data offers the potential to unlock previously untapped predictive signals crucial for nuanced credit risk and first-party fraud detection. This paper investigates a novel, event-driven machine learning framework that leverages anonymized mobile app usage data collected up to 45 days prior to credit applications. By transforming raw event streams into a comprehensive set of behavioral features, the proposed approach captures subtleties of user activity patterns, enabling more dynamic and context-aware risk scoring.

The models are designed to address multiple complementary risk targets: predicting early repayment risk and detecting first-party fraud attempts—areas where timely intervention can significantly reduce financial losses. In addition, this research recognizes and addresses several practical challenges inherent to event data modeling in financial contexts: data quality variability across geographies; evolving app usage patterns over time; and the need to balance model complexity with interpretability and operational deployment constraints.[8]

By sharing detailed methodologies for data engineering, feature extraction, model training, and validation, this paper aims not just to present predictive improvements but also to foster transparency and reproducibility. Ultimately, the integration of behavioral event data into credit and fraud risk models represents a critical step forward in financial risk science, aligning with broader fintech trends toward real-time, customer-centric decision-making, and inclusive credit access.[9]

2 Background and Motivation

2.1 The Shift to Alternative Data

The proliferation of alternative data—defined as information outside traditional credit bureau or financial databases—has transformed lending and fraud risk analytics. Mobile app interaction data, geolocation traces, and digital footprints provide context-rich, real-time behavioral indicators that enhance model accuracy, extend reach to new customer segments, and accelerate onboarding and decisioning processes.[7]

2.2 Behavioral Event Streams in Finance

Mobile applications log an array of user-generated events: screen navigation, action timestamps, device changes, cancellations, session durations, and more. When anonymized and aggregated, these signals can act as proxies for intent, engagement, and even deception.

Research in behavioral finance and computational fraud detection affirms the value of such high-frequency event streams for risk segmentation and anomaly detection. [6]

2.3 Motivation for Event-Based Modelling

The central motivation lies in addressing the limitations of static, transaction-based models by incorporating dynamic user behavioral signals:[7]

- **Improved risk detection:** Event models uncover subtle behavioral anomalies that static models miss.
- **Inclusivity:** Extends credit and fraud assessment to users with sparse financial histories but rich app engagement.
- **Speed and responsiveness:** Supports real-time or near real-time scoring aligned with digital onboarding workflows.

3 Literature Review

3.1 Alternative Data in Credit Risk and Fraud Modeling

The proliferation of alternative data sources in the financial sector has transformed how lenders assess creditworthiness and detect fraudulent activity. Traditional models historically relied on demographic data, credit bureau scores, and account-level transaction history. However, researchers have found that incorporating alternative signals—such as mobile phone usage, social networks, and e-commerce behaviors—improves predictive performance, particularly for thin-file customers or those with limited formal financial histories.[2]

Mobile behavioral data, including app event streams, offer granular, real-time insights into user intent and risk signals. It is demonstrated that real-time event streams processed from digital banking applications can be effectively used for fraud detection, delivering results in operational timeframes and enhancing traditional rule-based systems[10]. This echoes a growing consensus that behavioral and engagement signals supplement standard risk metrics and help close inclusion gaps in digital lending.

3.2 Real-Time and Complex Event Processing

The deployment of real-time event processing platforms has been a topic of both academic inquiry and practical innovation. For example, complex event processing frameworks enable timely identification of suspicious patterns in streaming app and transactional data—supporting immediate fraud intervention and adaptive risk scoring. Early-stage evidence suggests these approaches can increase detection rates for both credit risk and fraud beyond what static-variable models achieve.[3]

4 Data and Feature Engineering

4.1 Data Sources and Event Selection

4.1.1 Raw App Events

The raw data is comprised of anonymized event logs—each event marking a discrete user action or navigation within the mobile app. For the scope of risk and fraud modeling, the following restrictions and refinements were enforced:

- **Event types:** Only events from user journeys relevant to onboarding, product selection, and credit-related decisions were included. Excluded were casual browsing, customer support, and external integrations.
- **Observation window:** User event sequences were considered if they occurred within a fixed window (e.g., the 45 days leading up to, but not including, the application date), ensuring behaviors preceding the application influence the model.
- **Anonymization and privacy:** All events are fully anonymized—no direct identifiers, device IMEIs, IP addresses, or transactional metadata that could compromise user privacy are retained or used.

4.1.2 Data Quality Controls

Robust quality controls were instituted to filter corrupt, incomplete, or out-of-sequence event records. Cross-validation against known onboarding flows helped ensure the integrity and relevance of captured event types.

4.2 Revolut App Events Raw—How Data is Stored (Simulator)

The raw app event data generated by user interactions is ingested and stored within a dedicated event data store, designed for high scalability, reliability, and privacy compliance. Data ingestion pipelines process, validate, and persist event streams in an optimized format to facilitate model development and analytics.

Figure 1 below provides an overview of the event simulator, illustrating how test events are generated, structured, and routed into the storage system.

4.3 Revolut App Events—Filtering Screens for Model

To maximize both the predictive value and manageability of the dataset, only user events originating from the **Accounts**, **Product**, and **Credit** screens were included in the modeling pipeline. This targeted approach focused on those event flows most relevant for onboarding and credit risk analysis, while:

- Ensuring signal relevance to user intent, application journey, and risk assessment.
- Excluding extraneous events (such as generic browsing, customer service, or marketing interactions) that add noise and inflate data size.
- Maintaining a manageable, efficient dataset, reducing storage and processing costs without compromising model quality.

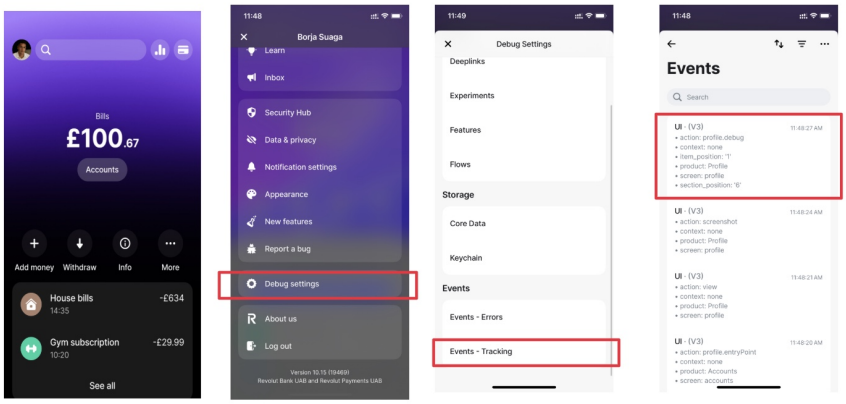

Figure 1: Overview of the app events simulator used for generating and testing event flows.

Figure 2 illustrates the event filtering logic, showing how relevant screens are identified and retained for downstream analytics.

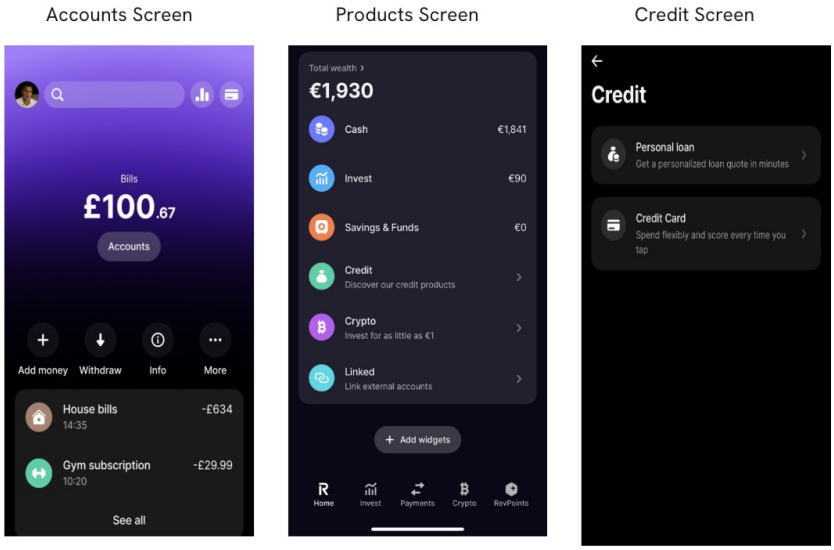

Figure 2: Event selection and filtering: Only interactions from Accounts, Product, and Credit screens were retained for modeling.

4.4 Event Aggregation: Sessions and Actions

Session: A session is defined as a contiguous sequence of user actions bounded by explicit logouts, application closures, or inactivity thresholds.
Interaction: An interaction constitutes a user-performed event (such as a screen view, tap, or selection) associated with a particular product or workflow phase. [1] Aggregation techniques (mean, median, min/max, standard deviation) transform granular event sequences into robust features capturing individual session properties, inter-session patterns, and cross-device and locale consistency.

4.5 Feature Engineering—In-Depth Overview

Over 250 behavioral features are engineered

- **Session Metrics:** Session frequency, durations, session timing variability.
- **Device and Locale Patterns:** Number/type/stability of devices, device changes, locale-country match.
- **Engagement & Navigation:** Unique screens visited, action type ratios, consistency within/across sessions.
- **Credit-Interaction Proxies:** Sequencing/timing of credit actions, loan navigation frequency, application patterns.
- **Spike & Responsiveness:** Magnitude/frequency/timing of spikes, inactivity spells, responsiveness at peak times.
- **Cancellation & Abandonment:** Cancel/abandon counts, timing, time-to-cancellation, form incompletion.
- **Action/Velocity:** Rapid-action clusters, action velocity.
- **Questionnaire Engagement:** Time on forms, interaction density, engagement ratio.

Examples:

- High ratio of evening/night sessions as risk proxy.
- Sudden device changes or mismatches as possible fraud signals.
- Multiple rapid application attempts with inactivity spells indicating uncertainty or fraudulent intent.

4.6 Visual Overview: Data and Model Pipeline

5 Model Development

5.1 Label Definition and Target Construction

Credit Risk: Flags users detected as 14+ or 30+ days past due at 3 months on book.
First-Party Fraud: NO Intent to pay proxies users showing unrecoverable delinquency (90+ days past due within 4 months).

5.2 Feature Selection and Importance

Feature selection is a critical step when working with high-dimensional behavioral data, as it reduces model complexity, improves generalization, and highlights the most predictive signals relevant to credit risk and first-party fraud detection. In this study, a combination of SHapley Additive exPlanations (SHAP) values, forward feature elimination, and correlation-based filtering methods was employed to select and prioritize features.

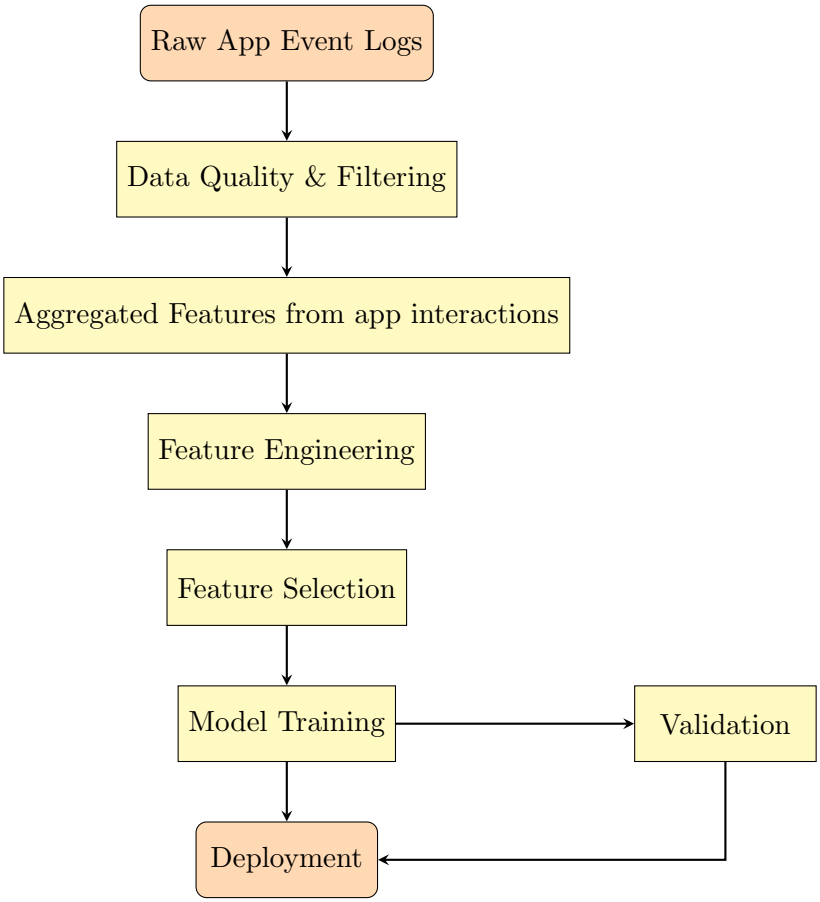

Figure 3: Data pipeline example.

5.2.1 SHAP Values for Feature Importance

SHAP values provide a unified, model-agnostic approach to quantify the contribution of each feature to model predictions. They measure the impact of a feature on the prediction while considering interactions with other features.

The procedure was as follows:

- A LightGBM model[13] was trained using the full feature set.
- SHAP values were computed for each feature across a representative holdout validation set.
- Mean absolute SHAP values were aggregated per feature to quantify overall importance.

Features exhibiting higher mean absolute SHAP values indicate greater influence on predictions. For example, behavioral features such as session regularity, timing and intensity of interaction spikes, and engagement concentration emerged as top predictors. SHAP also provides interpretability by revealing feature effect directionality, with irregular session timing associated with increased fraud risk and consistent engagement tending to lower risk scores.

5.2.2 Forward Feature Elimination

Forward feature elimination is a wrapper-based method used to identify a minimal subset of features that achieve near-optimal predictive performance. This approach iteratively adds features based on incremental gains in model accuracy.

The procedure implemented was:

1. Begin with an empty feature set.
2. At each iteration, train candidate models by adding one remaining feature at a time.
3. Evaluate model performance (e.g., cross-validated AUC or GINI coefficient).
4. Select the feature providing the largest improvement.
5. Repeat until performance gain plateaus or a predefined number of features is reached.

This method often resulted in selecting a small subset (approximately 20–50) of behavioral features capturing most of the predictive signal. The features selected through forward elimination largely aligned with the top SHAP-ranked features, indicating robustness and helping to avoid model overfitting caused by redundant variables.

5.2.3 Correlation-Based Feature Filtering

Given that behavioral features extracted from event streams can exhibit high multicollinearity (e.g., similar session timing metrics), correlation filtering was applied as a preprocessing step to reduce redundancy.[12]

The steps included:

- Calculation of pairwise Pearson correlation coefficients among features using the training data.

- Identification of feature pairs or groups exceeding a high correlation threshold (e.g., $|r| > 0.5$).
- Retention of the feature within each correlated group showing higher SHAP importance or better individual predictive performance, while discarding others.
- Removal of features with low variance or negligible predictive importance.

This filtering improves model stability, interpretability, and computational efficiency by eliminating redundant features and mitigating issues arising from multicollinearity.

5.2.4 Feature Selection Workflow

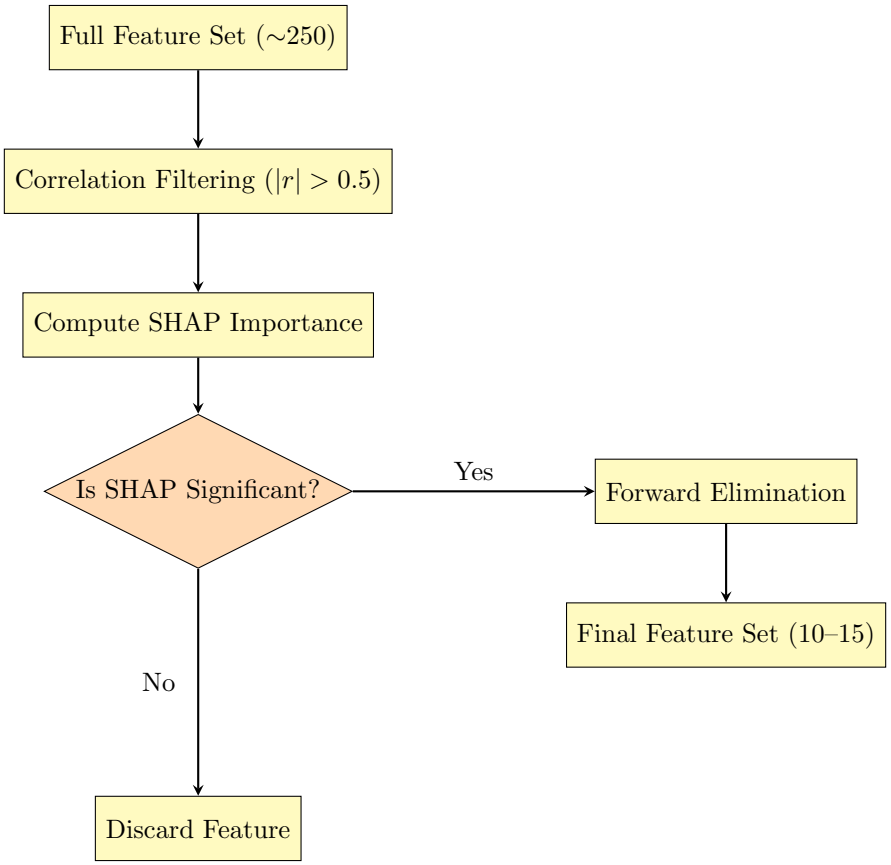

Figure 4: Feature selection workflow: combining correlation filtering, SHAP-based ranking, and forward elimination.

5.2.5 Summary

The final feature selection pipeline integrates these methods sequentially:

1. **Correlation Filtering** to prune highly collinear features.
2. **SHAP Importance Ranking** to prioritize features by predictive contribution.
3. **Forward Feature Elimination** to iteratively construct a compact, high-performing feature subset.

This combined approach produces a parsimonious set of behavioral features balancing interpretability, robustness, and predictive power. It enables domain experts to identify key behavioral drivers such as session irregularity, rapid action bursts, and application abandonment, facilitating actionable insights and supporting regulatory transparency in credit risk and fraud modeling.

5.3 Model Architecture & Algorithms

The modeling framework leverages Light Gradient Boosting Machines (LightGBM)[13] due to their efficiency and robustness in handling large-scale, high-dimensional, and correlated feature spaces typical of behavioral event data. The overall architecture is designed to balance predictive performance with interpretability and operational feasibility.

5.3.1 Light Gradient Boosting Machines (LightGBM)

LightGBM is a state-of-the-art gradient boosting framework that builds decision trees sequentially to minimize a differentiable loss function. It offers several advantages suitable for modeling complex behavioral data:

- **Gradient-based One-Side Sampling (GOSS):** LightGBM efficiently handles large datasets by focusing on instances with large gradients, which improves training speed without degrading accuracy.
- **Exclusive Feature Bundling (EFB):** It reduces dimensionality by bundling mutually exclusive features, effectively managing the high-dimensional and sparse feature space encountered in event-based behavioral modeling.
- **Handling Correlated Features:** LightGBM's tree-based learners naturally accommodate multicollinearity by selecting splits that maximize information gain, reducing the need for explicit feature decorrelation.
- **Regularization:** Supports L1 and L2 regularization to control overfitting given the large number of behavioral features.
- **Support for Missing Values:** Built-in mechanism to handle missing or incomplete data points common in event logs.

5.3.2 Rationale for Algorithm Choice

Compared to linear or simpler models such as Logistic Regression, LightGBM provides superior flexibility and non-linear modeling capability, key for capturing complex interaction patterns in app usage behavior. Yet, it remains computationally efficient and interpretable through tools like SHAP for feature explanations, thus aligning well with regulatory and operational requirements.[11]

5.3.3 Validation Strategy

To rigorously assess model generalization and mitigate risks from temporal distribution shifts and overfitting, a comprehensive validation strategy was employed incorporating:

- **Hold-out Sets:** A dedicated portion of the dataset, unseen during training, was reserved as a hold-out test set to provide an unbiased evaluation of final model performance.
- **K-Fold Cross-Validation:** The training data was partitioned into K folds (typically $K = 5$ or 10). In each iteration, one fold serves as validation while the rest form the training set. This methodology provides robust estimates of predictive performance and variability by averaging results across folds.
- **Temporal Validation Splits:** Recognizing temporal dynamics in user behavior and app updates, training and test sets were split chronologically to mimic production scenarios. This approach validates model stability over time and ensures that the model does not rely on data from the future relative to prediction points, preventing data leakage.

Hyperparameter tuning of LightGBM (e.g., number of trees, learning rate, maximum tree depth, minimum data in leaves) was performed within the cross-validation folds using grid or Bayesian optimization to optimize key performance metrics such as Area Under the Receiver Operating Characteristic curve (AUC-ROC) and GINI coefficient. Early stopping mechanisms based on validation loss were employed to prevent overfitting.

5.3.4 Model Interpretability

Though LightGBM is a powerful ensemble method, model interpretability is crucial for trust and compliance. Post-training, feature importance was analyzed through SHapley Additive exPlanations (SHAP), which provide consistent, local and global interpretability of how behavioral features influence individual predictions and overall model behavior.

5.3.5 Model Development Workflow

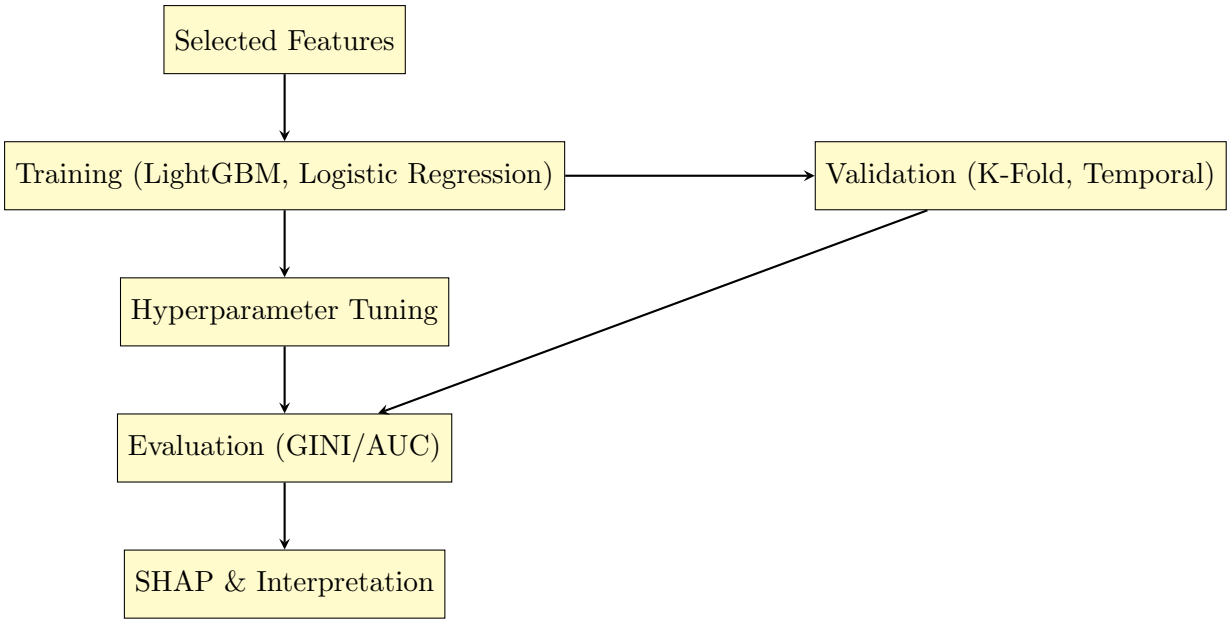

Figure 5: Model development and validation workflow.

5.4 Evaluation and Results

5.4.1 Performance Metrics

Table 1 summarizes the GINI coefficient ranges achieved by the event-based models across key risk targets. These metrics demonstrate consistent improvement in predictive discrimination, particularly for first-party fraud (No Intent to Pay) and early repayment risk.

Risk Target	**GINI (%)**
First-Party Fraud	50–60
Early Repayment Risk	36–45

Table 1: Summary of GINI coefficients for key risk targets using event-based models

5.4.2 Detailed Model Performance Across Samples

The following tables provide granular insights into LightGBM model performance, including approximate sample sizes, counts of bad outcomes, and bad rates for various delinquency definitions and dataset partitions. Exact GINI values are presented as ranges to maintain summary clarity. The consistency across training and test splits supports model robustness and generalizability.

First Party Fraud – 90+ Days Past Due at 4 Months on Book (4MOB)

Sample	**LightGBM GINI (%)**
Overall	55–60
Train	56–58
Test	53–55

Early Repayment Risk – 30+ Days Past Due at 3 Months on Book (3MOB)

Sample	**LightGBM GINI (%)**
Overall	40–45
Train	42–44
Test	36–39

5.4.3 Discussion

These results demonstrate that:

- LightGBM models exhibit stable and consistent performance across training and test samples, indicating strong generalization capability.
- Bad rates remain consistent within respective delinquency cohorts, confirming reliable segmentation of risk levels.
- Event-driven behavioral features contribute significantly to prediction quality across both early delinquency and first-party fraud scenarios, underscoring their operational value in credit risk management.

Overall, the evaluation supports the practical effectiveness of leveraging granular app event data for enhanced credit risk and fraud detection modeling.

GINI improvements for event-based models are typically 4–6 points over key segments.

5.5 Benchmarking Against Traditional Models

A rigorous benchmarking framework was established to compare the effectiveness of the event-based machine learning models against conventional approaches commonly used in credit risk and fraud assessment. Specifically, the following strategies were applied:

5.5.1 Comparison with Logistic Regression Baselines

The initial benchmark involved traditional Logistic Regression models trained on the similar set of features. This comparison served as a baseline to assess the incremental value brought by the lightGBM Model

Empirically, the stand-alone event models—particularly those based on LightGBM—consistently exceeded the predictive accuracy of Logistic Regression baselines.

5.5.2 Integration with Local Acquisition Models and Bureau Scores

Beyond stand-alone application, the event-based models were also deployed in conjunction with existing local acquisition risk models. These local models typically leverage bureau credit scores and other context-specific variables. In production settings, the event model scores were combined as additional features in the local models

Evaluation results show that integrating event-driven predictions with bureau-based models delivers significant incremental uplift. The ensemble approaches achieved the strongest results in terms of GINI, AUC, and early fraud detection. This uplift was particularly meaningful in pilot geographies, driving improvements in approval rates and portfolio quality.

5.5.3 Summary of Uplift and Best Practices

- Stand-alone event models perform on par or better than legacy, bureau-only models, highlighting the predictive power of behavioral signals. - Combined/ensemble models, leveraging both event-based and traditional bureau features, provide a robust uplift in risk segmentation and fraud detection, supporting more inclusive and accurate credit decisioning. - The approach is production-ready, adaptable, and enhances existing risk assessment frameworks without replacing established best practices.

This benchmarking strategy supports the conclusion that behavioral event data, when used thoughtfully alongside traditional risk scores, materially improves credit and fraud model performance while maintaining operational compatibility and regulatory transparency.

5.6 Feature Contribution Analysis

The analysis of feature contributions provides key insights into which behavioral variables drive model predictions for credit risk and first-party fraud detection. Understanding these top predictors enables both improved model interpretability and actionable business

insights. Below is a detailed discussion of the leading features identified through feature importance rankings such as SHAP values and their relevance:

- **Median Interval Between Sessions:** This metric captures the typical time gap between user app sessions. Longer or highly irregular intervals may indicate inconsistent engagement or reduced intent to repay, correlating with elevated risk. Conversely, steady and frequent sessions often signal committed users.

- **Timing and Intensity of Interaction Spikes:** Sudden bursts of activity—such as rapid navigation through credit-related screens or rapid completion of forms—may reflect impulsive behavior, heightened interest, or attempts to conceal fraudulent intent. Both the magnitude and temporal positioning of these spikes provide critical signals.

- **Session Regularity and Variability:** Consistency in session frequency and duration, as well as variability patterns, serve as proxies for behavioral stability. High variability or erratic usage patterns often align with increased risk or fraud likelihood.

- **Number and Timing of Cancellations or Abandonments:** Frequent application cancellations or form abandonments, especially near critical decision points, can signal uncertainty, hesitancy, or intentional risk evasion. Timing features, such as rapid abandonment soon after starting an application, strengthen this signal.

- **Action Velocity:** This measures the speed and clustering of user actions within sessions, capturing how quickly users navigate or provide inputs. Exceptionally fast or bursty velocities may indicate bots, automation, or fraudulent behavior, whereas natural pacing suggests legitimate intent.

Additional important predictors include:

- **Device and Locale Consistency:** Sudden changes in device type or discrepancies between device locale settings and declared user country can indicate account takeover or deceptive behavior.

- **Unique Screens Visited:** The diversity of app screens navigated reflects user familiarity and intent focus. Limited or overly broad navigation may both carry risk implications.

- **Engagement During Verification or Application Forms:** Metrics measuring time spent, interaction density, and focus during key verification steps help distinguish genuine applicants from fraudulent or negligent users.

- **Evening or Night Session Ratios:** A higher proportion of sessions occurring during unusual hours may correlate with increased risk, reflecting atypical user behavior often seen in fraud.

- **Frequency of Device Changes:** Multiple device swaps in a short time frame may signal attempts to mask identity or evade detection.

Together, these features form a robust behavioral signature set, enabling the models to detect subtle patterns linked with elevated credit risk and No Intent to Pay fraud. Their combined importance verifies the value of leveraging rich app interaction data beyond traditional credit bureau metrics for enhanced predictive accuracy and operational risk management.

6 Deployment Considerations

6.1 Model Integration

Models are integrated as adjusters or supplements, allowing real-time or batch scoring. Scores can be tailored for specific geographies or product offerings.

6.2 Impact Measurement

The deployment of the event-based credit risk and fraud detection models in controlled pilot environments has yielded significant positive impacts across multiple key performance indicators. Comprehensive measurement of these pilots underscores the practical value and operational benefits of integrating behavioral event data into existing risk frameworks.

- **Up to 8–9% Improvement in Approval and Sales Rates:** By incorporating rich behavioral features derived from mobile app usage, the models better differentiate between low- and high-risk applicants, particularly in segments with thin or sparse credit histories. This enhanced risk segmentation facilitates more informed and confident credit decisions, leading to increased approval rates without compromising portfolio quality. The uplift of 8–9% in approvals observed during pilot testing translates directly into greater customer acquisition and revenue growth, demonstrating the commercial viability of the approach.
- **Enhanced Early Fraud Detection with Minimal False Positives:** The incorporation of event-driven behavioral signals significantly improves the early identification of first-party fraud attempts, such as No Intent to Pay cases. Leveraging metrics like session spikes, rapid application attempts, and device inconsistency allows the models to flag suspicious patterns promptly. Critically, the models maintain a low false-positive rate, minimizing unnecessary application declines or verification interventions that can degrade customer experience. This balance enhances fraud prevention efficacy while preserving user trust and operational efficiency.
- **Operational Efficiency and Scalability:** Beyond predictive gains, the event-based modeling pipeline integrates smoothly with existing credit and fraud analytics infrastructure, supporting both batch and near real-time scoring. This facilitates automated decision workflows aligned with digital onboarding processes, reducing manual review workloads and accelerating application turnaround times.

In summary, the pilot implementations demonstrate that leveraging app event behavioral data materially advances approval efficiency, fraud prevention, and portfolio performance, delivering tangible business and operational benefits within production environments.

7 Limitations and Future Work

7.1 Model Limitations

While the proposed event-driven modeling framework demonstrates considerable advances in credit risk and first-party fraud detection, there remain several noteworthy limitations:

- **Coverage Gaps:** The effectiveness of behavioral event models critically depends on the granularity and recency of user event histories. Applicants who exhibit minimal engagement with the mobile app, or those new to digital banking, may not generate sufficient event data to facilitate reliable scoring. As a result, these users may fall outside the scope of event-based risk models, necessitating fallback or hybrid approaches leveraging traditional credit/bureau data.

- **Cross-Geography Variability:** Mobile app usage patterns, logging practices, and feature availability can vary substantially across geographies, user cohorts, and product versions. Differences in local regulations, user onboarding flows, and technical implementations may affect both the quality and distribution of event features. To maintain predictive validity, models may require geo-specific calibration or localization, introducing complexity in model management and deployment at scale.
- **Event Data Quality and Consistency:** Event logs may be incomplete, erroneously timestamped, or inconsistently structured due to factors such as application updates, device firmware differences, or backend service interruptions. Ensuring rigorous monitoring, validation, and robust data engineering pipelines is essential to maintain model reliability and prevent unforeseen performance degradation.

7.2 Future Directions

Future research and development will focus on augmenting the modeling framework to address current limitations and extend predictive capabilities:

- **Automated Feature Engineering via Deep Learning:** By leveraging deep learning architectures (e.g., recurrent neural networks, attention mechanisms), sequential modeling of raw event logs can be automated to capture complex temporal dependencies and emergent behavioral motifs. This could reduce manual feature engineering, uncover subtler risk patterns, and continuously adapt to evolving user behaviors, yielding better generalization across dynamic digital environments.
- **Adaptive Models for Auto-Retraining and Active Learning:** To ensure sustained model accuracy amid shifting user populations and digital trends, future implementations should embed mechanisms for periodic or automated retraining based on fresh data ingestion. Active learning strategies, where the model selectively seeks expert feedback on ambiguous or novel cases, will accelerate adaptation to fast-emerging risk types or fraud vectors.
- **Integration of Hybrid and Transfer Learning Paradigms:** Hybrid approaches combining behavioral event models with traditional credit scoring and transfer learning methods may enable robust performance for thin-file users or underrepresented geographies, further democratizing credit access and risk detection in diverse markets.

Collectively, these future directions aim to reinforce the adaptability, inclusivity, and resilience of event-based credit risk and fraud modeling as digital banking ecosystems continue to evolve. “‘

8 Conclusion

Harnessing mobile app event data marks a substantial advance for credit and fraud risk, enabling new levels of predictive granularity, inclusivity, and speed. Our event-based machine learning framework delivers meaningful performance gains in risk segmentation and fraud detection and is an essential complement to traditional data in modern financial services.

References

[1] Adelakun Matthew Adebowale and Olayiwola Blessing Akinnagbe. Cross-platform financial data unification to strengthen compliance, fraud detection and risk controls. *World Journal of Advanced Research and Reviews*, 20(3):2326–2343, 2023.

[2] Jonathan Kwaku Afriyie, Kassim Tawiah, Wilhemina Adoma Pels, Sandra Addai-Henne, Harriet Achiaa Dwamena, Emmanuel Odame Owiredu, Samuel Amening Ayeh, and John Eshun. A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. *Decision Analytics Journal*, 6:100163, 2023.

[3] Fereshteh Baratzadeh and Seyed M. H. Hasheminejad. Customer behavior analysis to improve detection of fraudulent transactions using deep learning. *Journal of AI and Data Mining*, 10(1):87–101, 2022.

[4] Aparna Krishna Bhat. Improve real-time fraud detection with dataops on resilient elastic platforms. *International Journal of Science and Research (IJSR)*, 13(10):193–198, 2024.

[5] Kishore Reddy Gade. Event-driven data modeling in fintech: A real-time approach. *Journal of Computational Innovation*, 3(1), 2023.

[6] Abdellah Hanae, Saida, and Youssef. End-to-end real-time architecture for fraud detection in online digital transactions. *International Journal of Advanced Computer Science and Applications*, 14(6), 2023.

[7] Jordan Harris, Emily Wong, and Samuel Gold. A cloud-native architecture for real-time transaction analysis. *Journal of Event-Driven Fraud Detection*, 3(1):27–35, 2025.

[8] Amarnath Immadisetty. Real-time fraud detection using streaming data in financial transactions. *Journal of Recent Trends in Computer Science and Engineering (JRTCSE)*, 13(1):66–76, 2025.

[9] Michael Levi. Frauds and their controls: Some scholarly reflections. *Crime and Justice*, 54(1), 2025.

[10] Ramchander Malkoochi. Event-driven fraud detection system: A cloud-native architecture for real-time transaction analysis. *World Journal of Advanced Engineering Technology and Sciences*, 15(2):1684–1693, 2025.

[11] NVIDIA Corporation. Financial fraud detection blueprint. NVIDIA AI, June 2025.

[12] Arjun Sirangi. Retail fraud detection via log analysis and stream processing. *Computer Fraud & Security Bulletin*, pages 21–32, 2018.

[13] Di-ni Wang, Lang Li, and Da Zhao. Corporate finance risk prediction based on lightgbm. *Information Sciences*, 602:259–268, 2022.