JULY 2023

NuData Security
mastercard

NUDATA SECURITY - MASTERCARD

DATA STRATEGY CASE STUDY

PRESENTED BY: GAGANDEEP KAUR
VANCOUVER BC

NuData Security
mastercard

CASE STUDY PROMPT

Scenario Description

Existing fraud detection methods rely on signals e.g. typing patterns, device and network information that may indicate when account activity is originating from a different user that is not the actual owner of the account. However online fraud happens in many ways including but not limited to social engineering and scams. In a scam scenario for example, the good user may be tricked into performing a transaction (e.g. a money transfer) to benefit the fraudster. In such a scenario, existing fraud detection methods are not effective because the account activity is originating from the actual owner of the account. As a result, existing methods are limited to only being able to catch third-party fraud i.e., where a fraudster takes over a good user's account to perform fraud.

Task

The objective is to develop a data-driven strategy for improving the existing detection methods. The proposed strategy should at least address the following aspects:

Data Needs: What kind of data would be needed to better identify instances of first-party and elaborate scam fraud?

Data Analysis: What methodologies would you employ to analyze the collected data effectively?

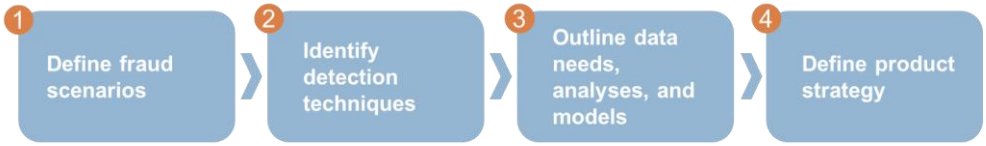
Predictive Modelling: How can the analyzed data be utilized to predict potential fraud cases accurately?

Product Strategy: How could your findings inform product strategy?

DATA STRATEGY FRAMEWORK

We first lay down a structured strategy that uses first principles to identify and address the common fraud scenarios (Exhibit 1). We subsequently expand upon each of the 4 elements of the strategy.

Exhibit 1: Data Strategy Framework



1. To ensure the relevant business scenarios are defined, we'll first define the common fraud scenarios
2. We'll then identify and define detection techniques for each scenario
3. We expand on the data needs, analyses and modelling required for each of these techniques
4. The techniques will then be mapped to current and potentially new NuData products. The products should then be prioritized, and a roadmap laid out. We'll propose a few criteria to prioritize the products but will treat the actual prioritization and roadmap as out of scope for the purposes of the report.

1 | DEFINE FRAUD SCENARIOS

We identify 5 common first-party fraud scenarios (Exhibit 2) to develop effective countermeasures and enhancing fraud detection methods, particularly for first-party fraud.

Exhibit 2: Fraud scenarios

Scenario	Examples	Definition
1. Social Engineering	Phishing	Fraudsters deceive good users into unwittingly assisting them in committing fraud.
2. Elaborate Scam	Bust-out fraud, sleeper fraud, multiple accounts [1]	Fraudsters create an appearance of legitimacy by engaging in genuine activities, then suddenly commit fraud and vanish.
3. False Claims and Disputes	Chargeback fraud, refund fraud, de-shopping, goods lost in transit [2]	Fraudsters exploit dispute mechanisms to fraudulently claim refunds or benefits.
4. Identity Purchasing	Synthetic identity, stolen identity	Fraudsters develop synthetic identities or acquire identities, often from vulnerable populations (e.g., immigrants, senior citizens), for illicit purposes.
5. Application Fraud	Fronting, multiple applications [3]	Fraudsters use false documentation and stolen identities to apply for loans, credits, or accounts deceitfully

02 | IDENTIFY DETECTION TECHNIQUES

For each fraud scenario, multiple detection techniques are identified (Exhibit 3).

Exhibit 3: Detection techniques

Scenario	Examples	Detection Techniques	Technique Type
1. Social Engineering	Phishing	**Checkout Fraud Detection**: Monitor account and transaction activity for abnormal behavior using NuData's behaviour and device intelligence solution.	*Existing NuData technology*
2. Elaborate Scam	Bust-out fraud, sleeper fraud, multiple accounts	**Social Network Analysis:** Monitor links to known fraud accounts through social media and partner intelligence sharing.	AI-based
3. False Claims and Disputes	Chargeback fraud, refund fraud, de-shopping, goods lost in transit	**Social Media Activity Risk**: Identify signals of known false claims from social media behaviour and develop risk scoring based on claim archetypes.	AI-based
		Historical Claim Analysis: Evaluate frequency and patterns of claims to create archetypes.	Research-based
4. Identity Purchasing	Synthetic identity, stolen identity	**Digital Footprint Verification:** Validate account creation against social identity and location information.	AI-based
		Dark Web Intelligence: Monitor dark web for stolen identities.	Research-based
5. Application Fraud	Fronting, multiple applications	**Digital Footprint Verification:** In addition to rule-based checks on submitted documents and information, use applicant's digital footprint to identify inconsistencies or mismatches in applicant data.	AI-based

These technologies focus on improving profiling and risk scoring to augment traditional rules-based detection systems. We expand upon these techniques below:

1. **Checkout Fraud Detection:** Monitor account and transaction activity for abnormal behavior using NuData's behaviour and device intelligence solution.
2. **Social Network Analysis**: This technique can be used to identify relationships among accounts and applications in addition to identifying direct connections. Utilizing NLP techniques to analyze textual data such as social media posts or comments, we can identify mentions or references to other users, indicating potential connections. Data from relationships between applicants can be used to detect potential fraud rings or collusion. Such an analysis can be used to combat elaborate bust-out frauds, such as the $200M fraud committed by a group of 18 individuals in New York using 7000 false identities [Link]).
3. **Social Media Activity Risk Scoring**: Digital identity verification tools use data points such as IP addresses, locations, cookies, and email checks to stop fraudsters from using someone else's identity or accounts. However, it doesn't stop people from using valid identification documents to commit first party fraud. Therefore, we need behavioural profiling to detect risk in such cases. Two specific use cases are listed below:
   - Social media activity can indicate spending behaviour that can be used to model behaviour-based archetypes. An example of such an archetype could be "low-income youth with high spending" known to

be associated with chargeback fraud. This can be used to flag potential fraud where a new customer fitting the archetype makes large purchases.

- Behaviour profiling can be used to flag identity purchasing fraud, by matching account activity with behaviour archetype to identify irregularities (for example, if an 18-year-old student's identity is used to purchase mortgage).

4. **Digital Footprint Verification**: In addition to identity verification using digital data points (e.g., IP addresses, locations) we can link social media profiles to account application to validate the applicant's digital presence. We can analyze social media activity to determine a score for identity risk - based on variety of parameters, including number of social accounts linked to the email, age of accounts, categorizing account activity into normal vs suspicious (based on posts, comments, likes), and sentiment analysis of activity (including presence of extreme negative sentiments such as racism, sexism etc.).

5. **Historical Claim Analysis**: We can build a collaborative network across financial institutions and merchants to share fraud intelligence. This will help build a 360-degree profile of an individual linking historical data and activities across accounts and applications to identify suspicious behaviour. For example, previous chargebacks is a great indicator of false claims fraud. According to Chargebacks911, 40% of first party fraudsters commit further fraud within 60 days [Link]. This technique can also be used to address multiple accounts, elaborate scams and stolen identity fraud.

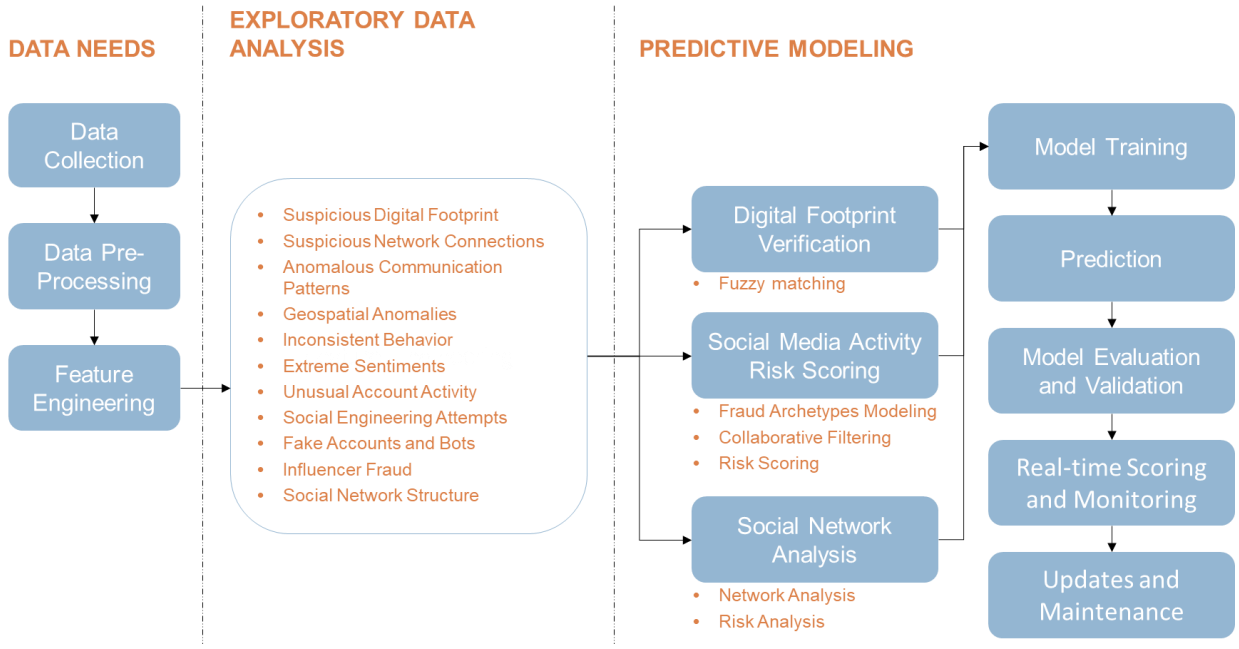
6. **Dark Web Intelligence**: Dark web intelligence can be used to provide early warnings of data breaches, monitor stolen credentials, verify identities, detect fraudulent services and tools, raise awareness of social engineering attempts, identify fraud rings, and uncover money laundering schemes. Integrating dark web intelligence with existing threat intelligence for can provide a comprehensive view of the threats landscape.

For the rest of our report, we shall focus on AI-based techniques (Social Network Analysis, Social Media Activity Risk, Digital Footprint Verification).

03 | OUTLINE DATA NEEDS, ANALYSES, AND MODELS

We now define a pipeline to outline the data needs, analysis, and modeling required across the detection techniques (Exhibit 4).

Exhibit 4: Machine Learning Pipeline



DATA NEEDS: WHAT KIND OF DATA WOULD BE NEEDED TO BETTER IDENTIFY INSTANCES OF FIRST-PARTY AND ELABORATE SCAM FRAUD?

All suggested AI-based techniques work on publicly available social media data from websites such as Facebook. We can extract the following data from public social media profiles for existing and new accounts:

- **Profile Information**: Basic profile data such as the user's name, username, profile picture, bio, and any publicly shared personal details.
- **Contact Information**: Contact information such as email addresses or website links, if publicly shared.
- **Employment and Education**: Publicly shared information about the user's employment history, educational background, and professional interests.
- **Location Data**: Some social media platforms allow users to share their current location or tag locations in their posts, providing information about their whereabouts.
- **Connections and Friends**: Information about the user's connections, friends, followers, and those they follow can indicate their social network and potential affiliations.
- **Posts and Updates**: Publicly shared posts, status updates, photos, and videos can provide insights into the user's activities, interests, and preferences.
- **Public Comments**: In some cases, public conversations and comments with other users on the platform may be accessible.
- **Interests and Likes**: Information about the user's likes, interests, and pages they follow can provide insights into their hobbies and preferences.
- **Publicly Shared Events**: Users may share information about events they plan to attend or have attended, providing clues about their activities and social engagements.

This data would require several cleaning and processing steps to prepare for analysis. We can expect to have a lot of missing data, and handling missing data would be crucial to manage the performance of our models. Data processing and modelling should incorporate privacy-preserving mechanisms (like differential privacy ML) to make sure the models are not biased. Given below are the key steps we need to complete before any analysis:

Data Collection

- Gather relevant data from various sources, including social media profiles, account activity logs, historical fraud data, financial records, application data, and identity information.
- Extract attributes such as user demographics, social media activity, transaction history, account creation details, relationships, and connections.

Data Preprocessing

- Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.
- Convert categorical variables into numerical representations for modelling.
- Apply text preprocessing techniques to prepare textual data for fuzzy matching and NLP-based analysis.

Feature Engineering

- Extract meaningful features from social media activity, transaction history, and account details.
- Create behavioural features like spending patterns, transaction frequencies, post frequency, follower counts, etc.
- Quantify network relationships using metrics like degree centrality, clustering coefficients, and link strength.

EXPLORATORY DATA ANALYSIS: WHAT METHODOLOGIES WOULD YOU EMPLOY TO ANALYZE THE COLLECTED DATA EFFECTIVELY?

Social media data can provide valuable insights for fraud detection, especially when combined with other data sources. Listed below are some of the key signals that need to be assessed as part of exploratory analysis:

- **Suspicious Digital Footprint**: Does the individual have fake social media accounts or operate multiple accounts?
- **Suspicious Network Connections**: Are there individuals who have a higher-than-expected number of connections to known fraudsters or suspicious accounts? This can help identify potential accomplices or members of a fraud ring.

- **Anomalous Communication Patterns**: Are there unusual communication patterns between certain individuals or accounts, such as a sudden increase in interactions or use of specific keywords or phrases? This could indicate coordinated fraudulent activities.
- **Geospatial Anomalies**: Does the social media data reveal geospatial patterns that do not align with the claimed locations of individuals or transactions? Geotagged posts or location information can be cross-referenced with other data to identify potential discrepancies.
- **Inconsistent Behavior**: Does the social media activity of certain individuals contradict the information they provided in applications or transactions? This can help identify instances of identity fraud or first-party fraud.
- **Extreme Sentiments**: Can sentiment analysis be performed on social media posts related to certain individuals or accounts? Unusually positive or negative sentiments might indicate attempts to manipulate public perception or conduct fraudulent activities.
- **Unusual Account Activity**: Are there sudden spikes in followers, likes, or other social media metrics for specific accounts that do not align with normal behavior? This could indicate use of fake accounts to inflate an account's popularity.
- **Social Engineering Attempts**: Are there messages or posts that attempt to deceive individuals into revealing sensitive information or participating in fraudulent schemes?
- **Fake Accounts and Bots:** Can the data be used to detect fake accounts or bots that might be involved in spreading misinformation or engaging in fraudulent activities?
- **Influencer Fraud**: Are there instances of influencers or high-profile individuals engaging in fraudulent behavior or promoting fraudulent products or services?
- **Social Network Structure**: Analyze the network structure to identify key individuals with high centrality or influence within a given social network that has known fraud presence. These individuals could be crucial players in fraudulent activities.

PREDICTIVE MODELING: HOW CAN THE ANALYZED DATA BE UTILIZED TO PREDICT POTENTIAL FRAUD CASES ACCURATELY?

To uncover these metrics, various analytics and machine learning techniques can be employed. These include natural language processing for sentiment analysis and text mining, graph analysis for studying social connections, machine learning for detecting anomalies, and data visualization for understanding patterns and relationships within the data. We can use the output from above analysis to design risk scoring and predictive machine learning model to power our selected fraud detection techniques, as below:

Digital Footprint Verification

Fuzzy Matching

- To link account applications with the respective individuals' social media profiles to validate digital footprint.
- Utilize Levenshtein Distance, Jaccard Similarity, Cosine Similarity, and Jaro-Winkler Distance to measure textual similarity and differences between strings. Set a similarity threshold to identify when two textual entities are considered a match.
- Assign similarity scores and establish links between applicants and their social media profiles.

Social Media Activity Risk Scoring

Fraud Archetype Modeling

- Use unsupervised learning algorithms like K-Means, Hierarchical Clustering, or Gaussian Mixture Models to create fraud archetypes based on behavioral features.
- Train the model to cluster similar profiles into distinct archetypes.

Collaborative Filtering

- Implement collaborative filtering techniques to find similar profiles based on behavioral features and network connections.
- Use methods like user-based or item-based collaborative filtering to identify profiles with similar spending patterns and social media behavior.

Fraud Risk Scoring

- Create a risk scoring model using supervised learning techniques such as Logistic Regression, Random Forest, or Gradient Boosting.
- Train the model using labeled fraud data and behavioral features to predict the likelihood of fraud for each account.

Social Network Analysis

Network Analysis for Fraud Ring Identification

- Build a graphical representation of the network using the collected and preprocessed data.
- Analyze the network structure and identify potential fraud rings or clusters using graph-based algorithms and centrality measures.

Link Analysis for First-Party Fraud and Identity Fraud Detection

- Identify and extract links or connections between accounts based on shared attributes, transactional patterns, or co-occurrences.
- Create features based on the identified connections, such as the number of links each account has and the frequency of transactions between linked accounts.

Predictive Model for detecting Potential Fraud

Training a Predictive Model

- Utilize the features from fraud archetypes, collaborative filtering, risk scoring, network analysis, and link analysis to train a predictive machine learning model.
- This model can be a classification model (e.g., logistic regression, decision tree, random forest) that predicts the probability of an entity being associated with fraud.

Making predictions

- Fraud Ring Identification: Once the model is trained, apply it to the network data to identify potential fraud rings or networks. Entities with high probabilities of being associated with fraud can be flagged for further investigation. Suspicious clusters of connected entities may indicate fraud rings.
- First-Party Fraud Detection: By analyzing individual entities and their connections within the network, the model can flag entities with a high likelihood of engaging in first-party fraud.
- Identity Fraud Detection: By analyzing patterns of connections between accounts and looking for anomalies, the model can help identify potential cases of identity fraud.
- Identifying Influential Fraudsters: The network analysis can also be used to identify the most influential fraudsters within the network. These individuals may play a crucial role in orchestrating fraudulent activities, and their detection can aid in breaking down fraud networks.

Model Evaluation, Deployment and Maintenance

Model Evaluation and Validation

- Split the data into training and testing sets to evaluate the model's performance.
- Use appropriate metrics like precision, recall, F1-score, and ROC-AUC to assess the model's effectiveness in detecting different types of fraud.

Real-time Scoring and Monitoring

- Deploy the predictive model to continuously evaluate new accounts and transactions in real-time.
- Set up alerts for accounts with high fraud probabilities for further investigation.

Updates and Maintenance

- Set up dedicated teams to investigate instances of recent fraud, track dark web for stolen identities and develop techniques that incorporate the information into production model regularly.
- Track latest research in first party detection techniques.
- Continuously update the model with new data to adapt to emerging fraud patterns.
- Periodically retrain the model to improve accuracy and reduce false positives and false negatives.

04 | DEFINE PRODUCT STRATEGY

The identified techniques will now be aligned to the most relevant products as new features (Exhibit 5).

Exhibit 5: Defining new product features

Product	Product status	New product feature
Account creation	Existing product	Digital Footprint Verification
Account access & update	Existing product	Social Media Activity Risk
Digital payments & transactions	Existing product	*Not applicable*
Network Intelligence	New product	Social Network Analysis

Digital Footprint Verification would add identity verification capability in conjunction with the existing behavioral and device intelligence to the Account Creation product. Social media behaviour can be used to compliment clickstream-based behaviour intelligence for enriching the indicators for first party fraud. Network Intelligence, driven by Social Network Analysis is a potential new product domain for NuData focused on first party fraud. Pursuing this new product would be dependent on the relative prioritization against other product features.

Now that we have identified the product features, we need to prioritize development based on the following proposed criteria. Prioritization is followed by outlining a roadmap.

1. Alignment to NuData / Mastercard strategy
2. Market size, profitability and competition
3. Ease and duration of implementation
4. Technological risk in terms of model performance
5. Data privacy concerns

To take an example, alignment to Mastercard strategy can be assessed by outlining core and adjacent markets, and aligning them to existing or new customers. The product strategy should also incorporate continuous innovation and explore potential generative AI use-cases such synthetic fraud scenario generation to improve model performance.

For the purposes of this report, we consider the actual prioritization and roadmap as out of scope.

CHALLENGES AND LIMITATIONS

Using Personally Identifiable Information (PII) and social media data for machine learning presents certain challenges, listed below with potential remediations:

- **Data Privacy and Ethics:** Prioritize user privacy, comply with data protection regulations, and handle data responsibly.
- **Lawful Access:** Obtain social media data legally, following platform terms of service and usage policies.
- **Data Bias:** Identify and mitigate biases in the data to ensure fair and unbiased machine learning results.
- **Data Use Policy:** Review and adhere to the data use policy of social media platforms.
- **Consent and Privacy:** Obtain user consent and follow best practices for data privacy.
- **Compliance:** Follow relevant data protection laws and regulations.
- **Anonymization and Aggregation:** Consider anonymizing or aggregating data to protect user identities.
- **Transparency and Explainability:** Ensure transparency and explainability of machine learning models for user trust and accountability.

REFERENCES

1. https://www.fico.com/blogs/what-first-party-fraud
   - Sleeper fraud occurs when a fraudster acquires a form of credit and, over time, builds up what appears to be normal customer behaviour. As the customer builds trust with the service provider over months or even years, they eventually ask for more credit and then cash in, taking the maximum amount of cash and any goods with them, never to be seen again.
   - Bust-out fraud, also known as hit and run, can happen on many types of financial services. It's quick and sometimes easy, with credit cards and loans being the easiest prey today. In some countries where cheques are in use or have slower clearing cycles, fraudsters can exploit these inefficiencies to inflate a credit balance by up to 10 times the limit and cash out before these transactions are caught.
2. https://linkurious.com/first-party-fraud/
   - Chargeback fraud: A customer buys something with a credit card, and then claims to their credit card company that the payment was fraudulent in an attempt to fraudulently get back the funds.
   - Goods lost in transit fraud (GLIT): After ordering something online, a consumer may claim those goods were not delivered in order to gain a fraudulent refund.
   - De-shopping: When a consumer buys something, uses those items, and then returns them for a refund.
3. https://plaid.com/resources/fraud/first-party-fraud/
   - Mortgage application fraud: A person submits false documentation in an effort to gain access to better interest rates or a higher mortgage amount. For example, renting an asset and claiming you own it.