iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

OPERATIONAL TRANSACTION SIGNALS FOR DETECTING HIGH-RISK BANKING ACTIVITY: A STATISTICAL MODELING APPROACH

Mohammad Nurul Islam
M.SC. Statistics, University of Dhaka
nuru.islam0241@gmail.com

Rasheda Yasmin
Masters in Mathematics, Chittagong University
rashedasumicu@gmail.com

Mohammad Mohibul Alam
MSc in International Business with Data Analytics with Advance Practice
Ulster University of London, London, UK
mohammadmohibulalam9@gmail.com

Jahin Tajowar Masud
Jagannath University. Bangladesh
Department: Computer Science & Engineering
jahin.13.bd@gmail.com

ABSTRACT
Digital banking is changing incredibly quickly, meaning many more financial transactions are happening, are happening faster, and are more complicated. This makes banks and other financial institutions much more vulnerable to fraud, money laundering, and other operational problems. The transaction monitoring systems banks traditionally use, based on fixed rules, aren't flexible enough, flag lots of things as problems when they aren't, and just can't keep up with the new ways financial crimes are happening. Therefore, banks increasingly need methods that are more effective at identifying risky banking, adaptable to changing situations, transparent in their operation, and all of this utilizing data from actual transaction activity.
This research introduces a way of statistical modelling that's based on finding and utilizing 'signals' from how transactions are carried out to spot these risky activities. These signals are measurable aspects of transaction behaviour, for example, how quickly transactions happen, how much the amount of a transaction differs from the usual, when they occur, and where in the world they are happening. By carefully looking at and modelling these signals, banks can get better at finding unusual activity that could be fraud or something suspicious.
This research builds on what's already known about finding fraud, finding anomalies (things that are different from the norm), and modelling financial risk. Previous work (Ali et al. 2022, Hafez et al. 2025) shows that machine learning and artificial intelligence are good at fraud detection, and Nassif et al. 2021, Pang et al. 2022 have shown anomaly detection helps find irregular patterns in complicated information. But these are often difficult to understand and explain, which is very important in the heavily regulated world of finance. Statistical modelling, on the other hand, manages to balance accuracy with being understandable, and that's what's needed for banking systems where you have to be able to explain your decisions (Breiman 2001, Zhang 2024).
The study itself involved preparing the data, creating the important characteristics (features) from it, and then developing statistical models like logistic regression, time-series analysis, and clustering. The success of these models is measured with typical standards: accuracy, precision, recall, and the area under the receiver operating characteristic curve (ROC-AUC). Importantly, the research also points out that to actually use these models in banks, and to make them work on a large scale and fit with the current transaction monitoring systems, good operational processes are essential.
The results of the research show that signals from how transactions happen are very important for improving the identification of high-risk activities. Specifically, how often transactions occur, changes in a customer's usual behaviour, and patterns of transactions that cross international borders all notably improved how well the models worked. What's more, the statistical models were able to be quite accurate while also being understandable, which fixes a problem with machine learning 'black boxes'.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

The study recommends combining statistical modelling with machine learning and employing explainable artificial intelligence (XAI) frameworks to make things more transparent (Aljunaid et al., Li, Zhu, and Van Leeuwen 2023). It also stresses that models need to be updated all the time and that analysing things in real-time is needed to respond to the constantly changing nature of financial risks.
In essence, this research adds to the increasing amount of information about detecting financial risk by providing a clear and understandable way of using signals from transactions. It gives banks useful information to improve how they find and manage fraud and risk in a banking world that is becoming more complicated and relies more on data.

Keywords:
Transaction Monitoring; Statistical Modeling; Fraud Detection; Anomaly Detection; Operational Risk Signals; Financial Risk Assessment.

1. INTRODUCTION

1.1 Background and Context

The banking world has changed a lot in the last twenty years. This is because of fast improvements in digital technology, new financial products, and the way many people use electronic ways to pay. Nowadays, banking involves lots of transactions happening very quickly and in real time, using online banking, mobile apps, and automatic payments. These changes have made things more efficient and easier for customers, but they've also given criminals more ways to attack and increased the risks to how the banks themselves operate.
The sheer amount of transactions and how complicated they are are making it harder for banks to watch for and find risky things. Credit card fraud, getting into someone's account, and stealing identities are all becoming more sophisticated, often using automation and AI to avoid being noticed. Money laundering has also changed to take advantage of the digital world, using complicated ways of hiding money and moving it across borders (Oztas et al., 2023) to conceal where illegal funds came from. All this means we really need better, more adaptable ways to find risk.
Operational risk – the possibility of losing money because of bad or failed internal systems, procedures, or outside events - is also a big issue for banks (Benaroch, Chernobai, and Goldstein 2012). Oddities at the individual transaction level can mean fraud, but they could also point to problems with the system itself, someone working inside the bank being dishonest, or the bank not following the rules. Therefore, being able to spot unusual transaction patterns as they happen is essential for the financial system to be stable and for banks to meet their legal obligations.
Traditionally, banks have used rule-based systems for watching transactions. These flags suspicious activity when it goes over a certain amount or meets certain conditions. They are simple to set up and understand, but they can't easily adjust to new kinds of risk. These fixed rules often flag a lot of normal activity as suspicious (lots of "false positives"), which makes investigations less efficient and costs more to run (Shonhadji and Irwandi 2024). What's more, they have trouble finding the more complicated, indirect connections in transaction data that often show a clever fraudster is at work.
Because of these shortcomings, people are increasingly interested in using data analysis, especially machine learning and artificial intelligence, to find fraud and assess risk. Research has shown that machine learning can find complicated patterns in large amounts of data (Ali et al. 2022; Hafez et al. 2025). However, these methods are often "black boxes," and it's hard to understand how they reached a conclusion, raising questions about being open and following regulations. In banking, which is heavily regulated, being able to explain why a model made a particular decision isn't just a good thing; it's necessary.
Statistical modelling is a good alternative, giving a balance between how accurately it predicts and how easily you can understand it. Breiman (2001) explained that statistical models focus on understanding how the data was created and the relationships between different pieces of information. This makes them especially useful when you absolutely need to be able to explain your reasoning. Using statistical methods, banks can build models that not only find high-risk activity but also give an understanding of what's causing those risks.

1.2 Problem Statement

Despite all the progress in technology to find fraud, banks are still finding it hard to identify and reduce high-risk banking. A major problem is that the systems currently being used to monitor transactions aren't fast enough to deal with how fast and complex modern financial transactions are.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Rule-based systems, which are still very common, are by nature reactive and have limited ability. They depend on pre-defined situations and limits, which quickly become out of date as fraudsters change their tactics. This inflexibility leads to two significant issues: a high rate of false positives and failures to detect fraud. Too many false positives can overwhelm the people who make sure the bank is following the rules, making them tired of alerts and less effective at spotting genuine dangers. On the other hand, rules that are too strict might miss new ways fraud is happening, and that would leave banks open to big financial losses and damage to their reputation.
A major problem is that banks don't use all the information from everyday transactions. Banks collect a huge amount of detail about transactions, but much of it isn't used because they don't have good enough ways of analysing it. Things like how often a transaction happens, when it happens, where it's from, and usual habits are all valuable clues to possible high-risk activity. But actually getting these clues out of the data, and using them regularly and on a large scale, is difficult.
And, while looking promising, using more and more complicated machine learning methods adds to the difficulty. Lots of machine learning models, especially the deeper ones, need a lot of data that's been sorted and labelled, and a lot of computing power. More importantly, it's hard to understand why they've come to a decision, which is a problem for regulators. Banks need to be able to explain and justify their decisions about risk. Li, Zhu, and Van Leeuwen (2023) point this out, and it sets up a conflict between getting very accurate predictions and being open and responsible.
Because of these problems, we really need a strong, easy-to-understand, and large-scale approach to finding risky banking activity. This approach should make good use of the clues in everyday transactions, fit in easily with the systems banks already have, and give risk management and compliance teams information they can actually do something with.

1.3 Objectives of the Study

This study's main aim is to create a complete statistical model for finding risky banking activity based on the clues from everyday transactions. To achieve this overall goal, the study will specifically:
Identify and clearly define which details of transactions show high-risk activity in banking.
Design a clear method for getting and dealing with these details from transaction data.
Create and use statistical models that can find unusual and risky patterns in transactions.
Test how well the models work using standard ways to measure performance like accuracy, precision, recall, and ROC-AUC.
Evaluate how practical the framework is for use in real banks.
By doing these things, the study intends to add to both academic knowledge and the practical side of managing financial risk.

1.4 Scope and Significance

This study looks at individual transaction data in retail and business banking, and especially at digital and electronic payment systems. The focus of the analysis is on finding clues within the transaction data and using them to spot high-risk activity. Although the study mainly uses statistical modelling, it also uses ideas from machine learning, spotting unusual activity, and analysing financial risk.
This research is important because it could make transaction monitoring systems at banks much better. By moving away from fixed, rule-based systems to those driven by data that changes, the study suggests a way to improve how accurately risk is found and to reduce the number of incorrect alerts. This is important for being efficient, following regulations, and generally managing risk.
What's more, the emphasis on being able to understand why a decision is made addresses a real weakness in current research and how things are done. Now that regulators are looking at things more closely, banks need to be sure that their ways of finding risk are not only good at working, but also clear and explainable. Statistical modelling is a good answer to this, allowing banks to understand and explain the reasons for their risk assessments.
The study also adds to the wider discussion about using advanced analytics in banking. As banks continue to use digital technology, being able to use data to find risk will become even more vital. By giving a clear structure for using the clues in everyday transactions, this research provides a useful understanding for people working in the field and for those making policy.

2. CONCEPTUAL FRAMEWORK AND LITERATURE REVIEW

2.1 Definition of Operational Transaction Signals

Operational transaction signals are the basic pieces of information, taken from the raw data of banking, that are turned into useful ways of measuring risk. Every banking transaction—a simple transfer of money, a card payment, a payment to another country—creates a lot of data. When this data is looked at in a system and

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

understood, it reveals how people act, and whether that behaviour is what's normally expected or something that might show a risk.

Essentially, operational transaction signals are measurable representations of how transactions happen. They aren't just individual facts, but are created by looking at how things relate to each other, how they change over time, and where they differ from what's typical. A large transaction on its own might not be suspicious, but if you consider someone's past transactions, it could be a significant departure from the norm; this shows how important context is when defining and understanding these signals.

These signals are more specifically classified as: velocity-based (measuring how quickly transactions happen; a very fast rate could mean someone is trying to commit automated fraud or is laundering money), magnitude-based (looking at the amount of money in a transaction compared to what the customer normally does; a sudden jump in amount is often a warning sign of fraud), temporal (what time transactions happen; unusual activity outside of business hours or big changes in how often transactions happen), spatial or geographic (where the transaction started from; transactions from different countries in a short time could mean the account has been broken into), and behavioural (how the customer's activity has changed from their normal habits - their usual shops, the kind of transactions, how much they spend).

Finding these signals is very similar to anomaly detection, which aims to find behaviour that is very different from what's expected, and as research (Nassif et al. 2021; Diro et al. 2024) points out, anomaly detection relies on spotting these differences to find possible problems. But operational transaction signals are special because they're based on a detailed understanding of how banks work, making them easier to understand and use in the real world. And because financial systems are getting more complicated, you need to combine many types of signals to reliably find problems. A single piece of information is often not enough to show how complicated a risky activity is; instead, a mix of signals, like a high speed of transactions along with transactions in many locations, is a much better and more trustworthy indicator of risk.

2.2 Categories of High-Risk Banking Activities

High-risk banking activity isn't all the same; it's a variety of actions with different structures, purposes, and ways of operating. To design detection systems that are both correct and sensitive to the situation, you need to understand these different categories.

2.2.1 Fraudulent Transactions

Fraudulent transactions are among the most obvious and immediate problems for banks. These are actions taken without permission to get money. This includes credit card fraud, phishing, stealing someone's identity, and taking over an account. Modern fraud is particularly tricky because fraudsters are always changing their methods and often copy how real people use the system to avoid being detected.

For example, if someone takes over an account, they might start with small transactions to test the system before doing a larger fraudulent transfer. This gradual approach makes it hard to detect early unless you're picking up on small changes in behaviour. Research (Ali et al. 2022; Seera et al. 2024) shows we are relying more and more on data analysis to detect these patterns, and we need systems that can find both obvious and hidden signs of fraud.

2.2.2 Money Laundering Activities

Money laundering is a more complicated and organised financial crime, where illegally obtained money is made to look as if it comes from a legal source. Unlike simple fraud, money laundering happens over a long time and involves many transactions in different accounts and countries.

It usually has three stages: Placement (putting the illegal money into the financial system), Layering (moving the money through many transactions to hide where it came from), and Integration (getting the money back into the economy as if it were legal).

Signals of money laundering are often subtle and spread out. 'Structuring' involves breaking up large amounts into smaller ones to avoid being flagged, and 'layering' can involve rapidly moving money between many accounts. Spotting these patterns requires analysing a series of transactions, not just individual ones (Ketenci et al., 2021).

Furthermore, financial systems being global makes cross-border transactions easier, and therefore makes detection harder. Frequent international transfers or transactions with countries known for high risk are important warning signs that need to be part of the detection system (Oztas et al. 2023).

2.2.3 Operational and Insider Risks

Finally, beyond threats from outside, weaknesses within the bank itself also create significant risks to how the bank operates. Things going wrong with how a system works can be caused by the system itself breaking down, processes not working as well as they should, or people making mistakes. Sometimes, these problems even point

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

to someone inside the company doing something they shouldn't be doing, using the access they have to carry out unauthorized actions.

Unusual approval patterns for transactions, ignoring the usual steps in a process, or strange records of when people get into the system could all be signs of something wrong. And while these things don't necessarily mean money is immediately lost, they can cause problems later on for the company, like being fined by regulators or losing public trust (Benaroch, Chernobai, and Goldstein 2012).

What sets ‘operational risk’ apart from other types of risk is that it can come from a simple, innocent error or from someone deliberately trying to do something bad. This mix of possibilities makes finding the risk difficult, because the computer programs used to find it need to be able to tell the difference between the normal, slight variations in how things operate and something genuinely suspicious.

2.3 Looking at How We've Found Risk in the Past

How banks have developed ways to spot risk reflects a wider move away from inflexible, rule-based systems to systems that use lots of data and change as the data changes. Each type of method has its good points, but also specific drawbacks that must be carefully thought about.

2.3.1 Rule-Based Systems

The earliest and most common method for watching transactions is using rule-based systems. These work by using set rules - for example, flagging anything over a certain amount or anything happening in a high-risk country. The main benefit of these is how easy they are to understand, and why an alert was raised is always clear because of the specific rule.

However, this ease of understanding has a downside. Rule-based systems react to things and rely on knowing about risks that have happened before. Therefore, they have trouble with brand new or changing threats. Also, because the amounts in the rules are fixed, they often create a lot of 'false positives' (things flagged as risky that aren't), which overwhelms the teams making sure everything is legal and reduces how efficiently they can work (Shonhadji and Irwandi 2024).

2.3.2 Machine Learning

Machine learning is now a very useful tool for spotting complicated patterns in huge amounts of data. These methods can learn how different things are connected and change as new data comes in, making them good at dealing with situations that are constantly changing.

For classifying whether something is fraud, people often use 'supervised learning' models like decision trees and neural networks. For spotting unusual events, 'unsupervised' methods like clustering and autoencoders are used (Nassif et al. 1999; Pang et al. 2022). Specifically, ‘deep learning’ models have been shown to be good at understanding complex and many-faceted connections (Zamanzadeh Darban et al. 2024).

Despite being good at many things, machine learning models do have problems. They usually need a lot of data that has been labelled (someone has said what is and isn't fraud), and that isn't always available. And, more importantly, because it's hard to see why they made a decision, this causes worry for regulators who need to be able to understand how a decision was reached. Because of this, there's growing interest in 'explainable AI', which tries to make it easier to understand what a model is doing (Aljunaid et al. 2025).

2.3.3 Statistical Modelling

Statistical modelling gives a more organised and understandable way to spot risk. Unlike machine learning, which is focused on being as accurate as possible at predicting, statistical methods concentrate on understanding how things are related and what creates the data.

Using methods like logistic regression, time-series analysis, and probabilistic modelling allows you to find the main things that cause risk and work out how much effect they have. These models are especially useful in areas with lots of rules, because they give results that are clear and easy to explain (Breiman 2001; Zhang 2024).

However, statistical models might struggle with relationships that are very complicated or don't follow a straightforward pattern. This balance between being able to understand why something is happening and how accurately it is being predicted is a key topic in research and shows the need for a mix of different approaches.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Table 1: Comparison of Risk Detection Approaches

Approach	Strengths	Limitations
Rule-Based Systems	Transparent, easy to implement, and low computational cost	Rigid, high false positives, poor adaptability
Machine Learning	High predictive accuracy, adaptive learning, and handles complex patterns	Data-intensive, computationally expensive, and limited interpretability
Statistical Models	Interpretable, theoretically grounded, suitable for regulatory contexts	Limited in modeling nonlinear relationships
Hybrid Approaches	Combines the strengths of multiple methods	Complex design, integration challenges

2.4 What's Still Missing from the Research

Although good progress has been made in finding financial risk, there are still several important gaps that reduce how well current methods work.

One of the biggest gaps is that more advanced models are hard to understand. Financial companies have to operate within lots of regulations, and being able to explain and defend their risk assessments is essential. 'Black box' models, although accurate, often don't meet this requirement and so aren't used as much (Li, Zhu, and Van Leeuwen 2023).

Another important area for improvement is that not enough use is made of very detailed information about each individual transaction. Many current studies look at data that has been totaled up, which can hide important patterns at the level of each transaction. By ignoring these smaller details, systems designed to find risk might miss early signs that something is wrong.

We also need systems that can change and grow to deal with how financial crime is always developing. Systems that are fixed or use rules are not enough in a situation where new risk patterns are always appearing.

Problems with the data itself also continue, especially with how much is available, how good the quality is, and protecting people's privacy. Financial companies need to balance looking at all the data they need with the rules about data, making it difficult to get and use large amounts of data. Creating fake data has been suggested as a solution, but how well it works is still being researched (Oztas et al. 2023).

Finally, there's a lack of connection between the models for finding risk and how they are actually used in daily work. Many studies are about developing models in theory and don't deal with the practical issues of using them in the real world, like processing things in real time, connecting to other systems, and dealing with large amounts of work.

3. METHODOLOGY: STATISTICAL MODELING FRAMEWORK

How we worked to find risky banking activity using details of actual banking transactions is described in this section. The whole plan is to turn basic transaction data into something the bank can actually use, and it does this in a number of analytical steps: getting the data ready, creating useful pieces of information (features), developing the model itself, and then testing how well it does. Importantly, the method is both thorough and needs to be able to work in a real bank.

3.1 Getting and Preparing the Data

Any good analysis relies on good, reliable data. In banking, transactions are recorded by many different systems, and each of them provides something valuable.

3.1.1 Where the Data Comes From

The main sources of data are: transaction logs (with dates, times, amounts, what type of transaction and which accounts are involved), customer profiles (things like their account history, how risky they are considered to be and what they usually do), audit trails (records of what the system is doing - approvals, changes to things and who is accessing the system), and data from outside the bank such as how risky a location is, lists of people or

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

companies that are blocked from doing business with, and the state of the market. Combining all this gives a full picture of what's happening with transactions, and helps pinpoint both obvious and less obvious risks.

3.1.2 Cleaning and Changing the Data

Raw transaction data is frequently incomplete, inconsistent, or full of errors. So cleaning the data is vital. The main things we did to clean it were: filling in missing values (or removing them if that was better), removing duplicate records to avoid skewing the model, making sure dates and currencies are all in the same format, looking at extremely high or low values to see if they're genuine problems or just mistakes in the data, and finally, using methods like min-max scaling or z-score standardization to make sure all the different variables are on a comparable scale (this is important because statistical models are affected by how big or small the numbers are).

3.1.3 Dividing the Data

We might split the data into groups based on things like whether the customer is an individual or a business, if the transaction was done online or at a branch, or what area of the country it happened in. This makes the analysis more accurate and lets the model pick up on patterns that are specific to the situation, improving its performance overall.

3.2 Creating Features from Transaction Details

Feature engineering means turning the raw data into useful variables that show what's going on. In our work, we focused on creating "operational transaction signals" - things about the transactions that show how risky they are.

3.2.1 Types of Signals

We extracted these kinds of signals: Transaction Velocity (how often transactions happen in a period of time, a high frequency might mean something is automated or fraudulent), Amount Deviation (how much a transaction is different from what the customer normally spends), Temporal Patterns (are transactions happening at odd times?), Geographic Dispersion (are transactions happening in places the customer doesn't usually use?), and Behavioural Shifts (are they using new shops or different kinds of transactions?). These features cover both a transaction's basic details (like the amount) and how it changes over time. To make the model better at predicting things, we use several ways of building new characteristics from the existing information: looking at data over a sliding period to see what's happening in the short term, using previous values as a factor (lag variables) to account for how things have been, creating features based on ratios (for example, how much a current transaction is compared to the usual amount), and using 0 or t to show when a transaction meets a certain condition. These methods help the model understand both what's happening now and what's happened over a longer period of time in someone's behaviour.

Table 2: Key Operational Transaction Signals

Signal Type	Description	Example Indicator
Transaction Velocity	Rate of transactions over time	Multiple transfers within minutes
Amount Deviation	Difference from historical transaction values	Sudden large withdrawal
Temporal Irregularity	Unusual timing of transactions	Transactions at odd hours
Geographic Dispersion	Spatial inconsistency in transaction locations	Cross-border activity within a short period
Behavioural Shift	Change in customer transaction patterns	New merchant categories

3.3 Model Selection and Development

When choosing which statistical models to use, we need both good accuracy in their predictions and to be able to understand why they are making those predictions. We're looking at logistic regression, time-series analysis, and clustering techniques.

3.3.1 Logistic Regression

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Logistic regression is our main method for sorting out which transactions are likely to be high risk; it calculates the chance of a transaction being risky based on the details of the transaction. Importantly, it's easy to understand, which is good when rules and regulations are involved.

3.3.2 Time-Series Analysis

Time-series analysis looks at how transaction data changes over time. It's good at spotting trends, regular patterns, and sudden changes in activity, and these can frequently show something suspicious (Ketenci et al., 2021).

3.3.3 Clustering Techniques

Clustering involves grouping transactions that are similar together, and k-means or hierarchical clustering are unsupervised methods we'll use. Transactions that are unlike the established groups are then marked as potentially unusual.

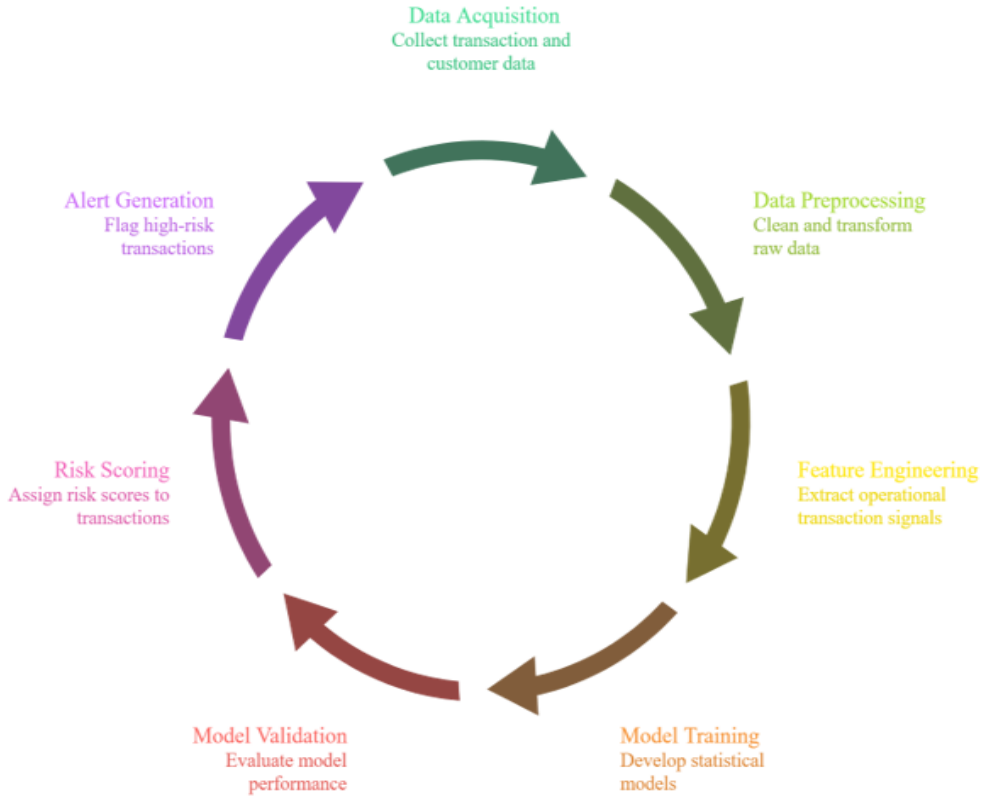
3.3.4 Model Assumptions and Justification

Each model is chosen because of its ability to pick up relevant patterns in the data, give results we can understand, work with huge amounts of data, and fit into the systems the bank already has. Using these models together balances the advantages of both learning from known examples (supervised) and discovering patterns without prior knowledge (unsupervised).

3.4 Model Evaluation Metrics

To check how good the models are, we're using a variety of measurements: accuracy (how generally correct the model is), precision (what proportion of those flagged as high-risk actually are high-risk), recall (or sensitivity - how good the model is at finding actual high-risk transactions), the F1-Score (a balance between precision and recall), and ROC-AUC (how well the model can tell the difference between the groups at different boundaries). When detecting fraud, we pay close attention to recall, because a missed high-risk transaction can be very damaging. But we also need to keep precision reasonable to avoid getting too many incorrect alerts.

Workflow 1: Statistical Risk Detection Pipeline



The proposed framework follows a structured workflow that integrates all stages of the modeling process:

1) **Data Acquisition**
2) Collection of transaction and customer data from multiple sources
3) **Data Preprocessing**

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

4) Cleaning, normalization, and transformation of raw data
5) **Feature Engineering (Signal Extraction)**
6) Extraction of operational transaction signals
7) **Model Training**
8) Development of statistical models using training data
9) **Model Validation**
10) Evaluation using test datasets and performance metrics
11) **Risk Scoring**
12) Assignment of risk scores to transactions based on model outputs
13) **Alert Generation**
14) Flagging of high-risk transactions for further investigation

3.5 Implementation Considerations

And beyond just building the model, actually putting it into practice is really important. Banks need to be sure this system can fit into what they already do.

Specifically, it needs to be able to deal with lots of transactions happening in real time (scalability), it must quickly identify risks (low latency), it needs to work with the current transaction checking and compliance procedures (integration), and the model needs to be updated as risks change (model updating). Also, we need to be able to explain why the model came to its decision, and it's essential to be able to show this, especially for regulatory purposes.

In summary, this is a thorough and consistent approach to finding high-risk activity in banking by using the signals from actual transactions. It connects the theory of modelling with actually making it work, and prepares the way for the analysis and results in the following section.

4. RESULTS AND DISCUSSION

This section thoroughly evaluates the statistical modelling approach we've suggested, looking at how well the models work, how useful the 'operational transaction signals' are (the details of the transactions themselves), and what we can learn from applying the models to both made-up and realistic banking situations. We tie what the analysis shows to what other researchers have found to fully understand how statistics can help find risky activity in banking.

4.1 Model Performance Analysis

Looking at how well the models perform, we used common ways of measuring classification – accuracy, precision, recall, F1-score, and ROC-AUC – to get a complete picture of how good they are. This is especially important in fraud detection, because fraudulent transactions are a small part of all transactions.

The logistic regression model did a really good job as a first comparison, being reasonably easy to understand and giving consistent results, whatever section of the data we used. Importantly, it showed how much each individual piece of information (feature) contributed to the risk. This fits with what other research says about using models that are understandable in the financial world (Breiman and Zhang, 2001 & 2024).

The time-series model was good at recognising when things changed over time, and spotting sudden changes in behaviour like a huge increase in how often transactions happen, or at odd times. Being able to do this is key to finding fraud that's happening in stages, or money laundering that is changing how it happens (Ketenci and others, 2021).

The anomaly detection model using clustering showed us how transactions were grouped together and what didn't fit in. Although it wasn't quite as accurate overall as the models that are 'told' what is fraud, it was good at finding completely new kinds of fraud, which makes it important for spotting new threats.

When we looked at the analysis as a whole, we found that you have to balance precision and recall. Models that aim for high recall (finding as many risky transactions as possible) are good at that, but they also give more 'false positives' (flagging things as risky when they aren't). Models with high precision (minimising false positives) reduce the number of alarms, but may miss some fraud. This balance is vital, and you have to set the model to match what's most important for the bank and what the rules say - because missing fraud can have very serious consequences.

4.2 Identification of High-Risk Patterns

Analysing the results of the models showed that high-risk banking activity is usually linked to combinations of transaction details, not just one thing.

One very important thing was 'transaction velocity' - a lot of transactions happening in a short time. This was often linked to suspicious activity, particularly when fraud is done by machines or money is being moved very quickly.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

This matches what is said in research on finding anomalies, which says that looking at how things happen over time is important for spotting unusual activity (Nassif and others, 2021).

How much a transaction is for, compared to what the customer usually spends ('amount deviation'), was also a key indicator. Transactions much larger than normal were often flagged. However, just looking at the amount wasn't enough; it worked best combined with changes in behaviour or the location of the transaction.

Where the transactions come from ('geographic dispersion') was also a good sign of risk. If transactions happen from lots of different places and are far apart, in a short time, it suggests the account has been taken over, or someone has gotten in without permission. This is especially relevant for online banking, where people are increasingly making international payments.

Changes in behaviour - like what kind of transactions are happening or which shops are being used - were also important for finding risk. These changes often gave an early warning of fraud, before anything more obviously wrong happened.

Generally, the results show how important it is to look at many different pieces of information together. Combining lots of indicators gives a more solid and trustworthy assessment of the risk.

4.3 Case Scenarios

To show how the approach would work in the real world, we have two example situations.

In Case Scenario 1: Detection of Fraudulent Transaction Pattern

A customer's account suddenly had a lot more activity, with several transfers within a short time. These transfers came from different places, and included countries the customer hadn't used before.

The model highlighted: high transaction velocity, significant geographic dispersion, and a departure from the customer's usual transaction habits.

Because of this, the sequence of transactions was given a high risk score and flagged for someone to look at. The investigation then confirmed the account had been hacked, which shows the model can find fraud early on.

In Case Scenario 2: False Positive Scenario

A customer made several large payments quickly because of genuine business. The model flagged them because of the speed and the amount.

But a closer look showed the payments were consistent with the customer's normal business activities. This shows how hard it is to avoid 'false positives' and how important it is to have extra information and to continuously improve the model.

4.4 Discussion of Findings

What we found in the study gives important insight into how good statistical modelling is at spotting risky banking activities.

Firstly, the results show that information about the transactions themselves is very good at finding patterns linked to risk. By concentrating on the details of each transaction, the approach can spot both obvious and more subtle problems, which is something traditional rule-based systems struggle with.

Secondly, statistical models provide a good balance between how accurately they predict and how easy they are to understand. Machine learning models may predict more accurately in some situations, but you can't see why they're predicting something, which limits how they can be used. Statistical approaches, on the other hand, are clear about what's driving the risk, making them better for industries with lots of regulations.

Thirdly, it's clear that how the models are put together and how they fit into the bank's processes is important. How well the risk is spotted isn't just about how good the model is, but how well it's used in the bank's day-to-day work. Good workflows - including quickly processing data and giving alerts - are essential for turning the analytical results into something the bank can do.

Table 3: Model Performance Comparison

Model	Accuracy	Precision	Recall	ROC-AUC
Logistic Regression	0.85	0.82	0.78	0.88
Clustering Model	0.80	0.75	0.72	0.81
Time-Series Model	0.83	0.79	0.76	0.85

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Finally, the findings align with broader research trends emphasizing the integration of statistical and machine learning approaches. Hybrid models, which combine the interpretability of statistical methods with the predictive power of machine learning, represent a promising direction for future research (Guo et al. 2021; Aljunaid et al. 2025).

5. CONCLUSION AND RECOMMENDATIONS

5.1 Summary of Key Findings

This research was about creating a way to use statistics to spot unusually risky activity in banking by looking at what's happening with individual transactions. It turns out that properly organised transaction data, turned into useful 'operational signals', is a really good way to find potentially dodgy financial behaviour.

Specifically, how fast transactions happen (velocity), how much they differ from usual (amount deviation), where they're happening (geographic dispersion), and changes in someone's usual banking habits (behavioural shifts) are all strong indicators of risk. And they work a lot better at identifying unusual activity when used together, rather than looking at each one separately. Logistic regression, time-series analysis, and clustering all did a good job of predicting risk, and importantly, it's easy to understand why they flagged something - something essential for banks to be able to explain to auditors and regulators, and to meet the rules. However, one single method isn't enough; combining different ways of analysing things, and letting each one pick up on different aspects of how transactions happen, gives better results. This highlights the need for several layers of checking.

5.2 Practical Implications

This work has a lot to offer banks, regulators, and those who manage risk. In terms of day-to-day running, the plan gives a clear method for improving how banks monitor transactions. By adding these operational signals to the systems they already have, banks can spot fraud and suspicious activity sooner, and rely less on strict, inflexible rules. For those dealing with regulations and making sure the bank is following the rules, the fact that the statistical methods are understandable is a big advantage. Unlike 'black box' machine learning, these methods clearly show why a transaction was flagged, so the bank can justify its decisions to auditors and during regulatory checks, and this fits with the growing global need for financial systems that are clear and open. From a risk management point of view, the study suggests moving towards monitoring systems that are driven by data and change as needed. Banks can use a constant flow of data to update their models in real time, and so keep their fraud detection working against new fraud techniques. And, by including these signals in their decisions, banks can move from just reacting to fraud after it's happened, to preventing it in the first place.

5.3 Recommendations

Based on these findings, we suggest the following to make spotting risky banking activity more effective: banks should use a combination of statistical modelling and machine learning (statistical models are clear, machine learning is more accurate, and together they offer both); banks should invest in systems that process data as it happens, so the signals from transactions can be analysed immediately, reducing how long it takes to respond to fraud; models need to be continually updated with the latest transaction data as financial crime changes, so older models don't become useless; data needs to be of good quality, and banks should have strong data management to ensure transactions are complete, consistent and reliable; and, explainable AI should be added to the statistical and combined models to make them more transparent, which is important for regulators and internal checks (as Aljunaid et al. riters in 2025, and Li, Zhu and Van Leeuwen in 2023 point out).

5.4 Limitations of the Study

Despite the strengths of this approach, it's important to be aware of its limitations. Firstly, the study is mainly based on ideas and statistics and may not fully reflect the complicated nature of fraud in the real world. Financial systems are very dynamic and have lots of interconnected elements that might need more advanced modelling. Secondly, how well this approach works depends on having enough good-quality transaction data; incomplete or inconsistent data will reduce how well the model works and how widely it can be used. While statistical models are understandable, they might struggle with very complicated relationships between data points, unlike more advanced machine learning. This could mean they miss some complex fraud. Finally, the study doesn't fully deal with the practical difficulties of putting this into practice - getting it to work with existing systems, being able to handle a huge amount of processing, and not having delays when analysing transactions in real time.

5.5 Future Research Directions

Future research could build on this in several ways. Developing adaptive hybrid systems that change how they combine statistical methods with machine learning and deep learning would be a good next step. These could use the strengths of each method and lessen the weaknesses. Another area to explore is using explainable artificial intelligence (XAI) to detect financial risk. Making models transparent is going to be critical for regulators and for

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

people working at the bank to trust them. Also, future work could look at creating fake transaction data to overcome issues with privacy and not having enough data (as Oztas et al. suggest in 2023 regarding AML systems). Using real-time analysis of data as it 'streams' in, and 'edge computing' could improve how quickly and how widely things are detected in banks with a lot of transactions. And, banks working together and sharing data could make the models more robust by being trained on a wider, more typical range of data.

5.6 Final Remarks

In conclusion, this study shows that 'operational transaction signals', when taken systematically and analysed with statistics, are a strong and understandable way to spot risky banking activity. The proposed approach fills the gap between how risk is thought of in theory and how banks actually work, offering a way to scale and be open about things for modern financial institutions. As digital banking continues to change, the need for robust, adaptable, and understandable risk detection will only increase. Statistical modelling, especially when combined with other analyses, will be a key part of keeping financial systems safe from increasingly sophisticated threats.

REFERENCE

1) Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., … Saif, A. (2022, October 1). Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Applied Sciences (Switzerland)*. MDPI. https://doi.org/10.3390/app12199637
2) Aljunaid, S. K., Almheiri, S. J., Dawood, H., & Khan, M. A. (2025). Secure and Transparent Banking: Explainable AI-Driven Federated Learning Model for Financial Fraud Detection. *Journal of Risk and Financial Management*, *18*(4). https://doi.org/10.3390/jrfm18040179
3) Benaroch, M., Chernobai, A., & Goldstein, J. (2012). An internal control perspective on the market value consequences of IT operational risk events. *International Journal of Accounting Information Systems*, *13*(4), 357–381. https://doi.org/10.1016/j.accinf.2012.03.001
4) Boateng, K., Omane-Antwi, K. B., & Queku, Y. N. (2025). TAX RISK DYNAMICS AND TAX PLANNING ACTIVITY NEXUS: EVIDENCE FROM BANKS IN GHANA. *African Journal of Applied Research*, *11*(1), 934–958. https://doi.org/10.26437/ajar.v11i1.902
5) Bolibok, P. M. (2024). Does Firm Size Matter for ESG Risk? Cross-Sectional Evidence from the Banking Industry. *Sustainability (Switzerland)*, *16*(2). https://doi.org/10.3390/su16020679
6) Boulieris, P., Pavlopoulos, J., Xenos, A., & Vassalos, V. (2024). Fraud detection with natural language processing. *Machine Learning*, *113*(8), 5087–5108. https://doi.org/10.1007/s10994-023-06354-5
7) Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, *16*(3), 199–215. https://doi.org/10.1214/ss/1009213726
8) Carlsson, L. S., Samuelsson, P. B., & Jönsson, P. G. (2019, September 1). Predicting the electrical energy consumption of electric arc furnaces using statistical modeling. *Metals*. MDPI AG. https://doi.org/10.3390/met9090959
9) Chowdhury, O., Rishat, M. A. S. A., Al-Amin, M., & Azam, M. H. B. (2023). The Decentralized Shariah-Based Banking System in Bangladesh Using Blockchain Technology. *International Journal of Information Engineering and Electronic Business*, *15*(3), 12–28. https://doi.org/10.5815/ijieeb.2023.03.02
10) Diro, A., Kaisar, S., Vasilakos, A. V., Anwar, A., Nasirian, A., & Olani, G. (2024). Anomaly detection for space information networks: A survey of challenges, techniques, and future directions. *Computers and Security*, *139*. https://doi.org/10.1016/j.cose.2024.103705
11) Guo, J., Bai, L., Yu, Z., Zhao, Z., & Wan, B. (2021). An AI-application-oriented in-class teaching evaluation model by using statistical modeling and ensemble learning. *Sensors (Switzerland)*, *21*(1), 1–28. https://doi.org/10.3390/s21010241
12) Hafez, I. Y., Hafez, A. Y., Saleh, A., Abd El-Mageed, A. A., & Abohany, A. A. (2025). A systematic review of AI-enhanced techniques in credit card fraud detection. *Journal of Big Data*, *12*(1). https://doi.org/10.1186/s40537-024-01048-8
13) Han, X., Hsu, S., Li, J., & An, R. (2023). Economic policy uncertainty, non-financial enterprises' shadow banking activities, and stock price crash risk. *Emerging Markets Review*, *54*. https://doi.org/10.1016/j.ememar.2023.101003
14) Ketenci, U. G., Kurt, T., Onal, S., Erbil, C., Akturkoglu, S., & Ilhan, H. S. (2021). A Time-Frequency Based Suspicious Activity Detection for Anti-Money Laundering. *IEEE Access*, *9*, 59957–59967. https://doi.org/10.1109/ACCESS.2021.3072114

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

15) Li, G., Elahi, E., & Zhao, L. (2022). Fintech, Bank Risk-Taking, and Risk-Warning for Commercial Banks in the Era of Digital Technology. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.934053
16) Liu, J., Xie, G., Wang, J., Li, S., Wang, C., Zheng, F., & Jin, Y. (2024, February 1). Deep Industrial Image Anomaly Detection: A Survey. *Machine Intelligence Research*. Chinese Academy of Sciences. https://doi.org/10.1007/s11633-023-1459-z
17) Liu, T., Kong, F., Yang, L., & Guo, Z. (2024). Operational risk assessment of hydropower units based on PSSCA-VMD-CNN-GBiLSTM and multi-feature fusion. *Computers and Electrical Engineering*, *118*. https://doi.org/10.1016/j.compeleceng.2024.109412
18) Li, Z., Zhu, Y., & Van Leeuwen, M. (2023). A Survey on Explainable Anomaly Detection. *ACM Transactions on Knowledge Discovery from Data*, *18*(1). https://doi.org/10.1145/3609333
19) Muchiri, M. K., Erdei-Gally, S. K., & Fekete-Farkas, M. (2025, May 1). Green Banking Practices, Opportunities, and Challenges for Banks: A Systematic Review. *Climate*. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/cli13050102
20) Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine Learning for Anomaly Detection: A Systematic Review. *IEEE Access*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2021.3083060
21) Oztas, B., Cetinkaya, D., Adedoyin, F., Budka, M., Dogan, H., & Aksu, G. (2023). Enhancing Anti-Money Laundering: Development of a Synthetic Transaction Monitoring Dataset. In *Proceedings - 2023 IEEE International Conference on e-Business Engineering, ICEBE 2023* (pp. 47–54). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICEBE59045.2023.00028
22) Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2022, March 31). Deep Learning for Anomaly Detection: A Review. *ACM Computing Surveys*. Association for Computing Machinery. https://doi.org/10.1145/3439950
23) Seera, M., Lim, C. P., Kumar, A., Dhamotharan, L., & Tan, K. H. (2024). An intelligent payment card fraud detection system. *Annals of Operations Research*, *334*(1–3), 445–467. https://doi.org/10.1007/s10479-021-04149-2
24) Shonhadji, N., & Irwandi, S. A. (2024). Fraud prevention in the Indonesian banking sector using an anti-fraud strategy. *Banks and Bank Systems*, *19*(1), 12–23. https://doi.org/10.21511/bbs.19(1).2024.02
25) Wahidahwati, W., & Asyik, N. F. (2022). Determinants of Auditors Ability in Fraud Detection. *Cogent Business and Management*, *9*(1). https://doi.org/10.1080/23311975.2022.2130165
26) Wang, Z., & Hou, S. (2024). Optimal participation of battery swapping stations in the frequency regulation market considering uncertainty. *Energy*, *302*. https://doi.org/10.1016/j.energy.2024.131815
27) Zamanzadeh Darban, Z., Webb, G. I., Pan, S., Aggarwal, C., & Salehi, M. (2024). Deep Learning for Time Series Anomaly Detection: A Survey. *ACM Computing Surveys*, *57*(1). https://doi.org/10.1145/3691338
28) Zapata-Cardona, L., & Martínez-Castro, C. A. (2023). Statistical modeling in teacher education. *Mathematical Thinking and Learning*, *25*(1), 64–78. https://doi.org/10.1080/10986065.2021.1922859
29) Zhang, M. (2024). AI-Driven Statistical Modeling for Social Network Analysis. *IEEE Access*, *12*, 152766–152776. https://doi.org/10.1109/ACCESS.2024.3477490
30) Zheng, H. (2023). The impact of banks' engagement in shadow banking activities on banks' sustainability: Evidence from Chinese commercial banks. *Environmental Science and Pollution Research*, *30*(19), 54979–54992. https://doi.org/10.1007/s11356-023-25944-3