

A Predictive Analytical Study on Detecting Fraudulent Transactions in E-Banking

S. Nithyasri¹, Sachin A.S², Tejas. M³, Prajwal⁴

^{1,2,3,4}Student, Don Bosco College, Department of Management, KR Puram, Bangalore (Affiliated to Bangalore North University)

Abstract

With the massive shift toward digital and mobile banking, the methods used by criminals to steal money have become increasingly advanced. Traditional security measures, which usually rely on simple "if-this-then-that" rules, are no longer enough to stop modern hackers. This study looks into how we can use predictive analytics—like machine learning and deep learning—to build a more proactive defence. The real issue here is that fraud isn't usually a single, glaring event; it's tucked away in **underlying** patterns that get lost among millions of ordinary, everyday purchases.

This research has analysed different models, like Hidden Markov Models and Neural Networks, handle the "imbalanced data" problem. This is a huge issue in banking because 99% of transactions are legitimate, making the 1% of fraud very hard for a computer to "learn" how to spot. By looking at a user's historical spending habits and behaviour, these predictive tools can flag a transaction as suspicious the moment it happens. It also looks at the trade-offs involved—while deep learning is incredibly accurate at catching tricky scams, it requires a lot of computing power and can sometimes be a "black box," making it hard for bank staff to explain why a transaction was blocked.

Ultimately, this study shows that the best approach for e-banking security isn't just one single tool, but a hybrid system. By combining behavioural analysis with real-time AI monitoring, banks can stay one step ahead of evolving fraud tactics like phishing and identity theft. The goal is to create a system that is fast enough to protect the bank but smooth enough that it doesn't annoy regular customers with constant "false alarms." This paper highlights that while the technology is getting better, we still need to balance technical accuracy with the practical costs of running these systems in a real-world banking environment.

Keywords: Predictive Analytics, E-Banking Fraud, Machine Learning, Anomaly Detection, Class Imbalance

Introduction

In the world of fast growth of e-banking in today's life where the digital technology plays a major role in financial innovation which changed the banking industry all over the world it also plays an important role in today's generation where it allows the customers to do financial transaction easily using the internet and mobile phones, without visiting banks physically. E-banking offers many benefits like 24/7 service, quick transactions, lower cost for banks, and better customer satisfaction. Online banking also increased the risk of cybercrimes and fraud activities. Therefore, detecting has become a major concern for banks

and financial institutions.

With the help of advanced **Information and Communication Technology (ICT)**, where banking transaction has become more automatic and connected. Even though it improves efficiency and convenience, it also creates a new weakness that fraudsters can exploit. In E-banking, fraud is an illegal or unauthorised activities that is done to gain financial benefits and also it reduces the customers trust, harm bank reputation, and affect the stability of the financial system.

Many research paper highlight that although e-banking improves banking services and competitiveness, weak security and changing of cyber-attack methods make online banking vulnerable to fraud. Digital banking fraud usually stems from things like tech loopholes, people not knowing the risks, or just plain weak passwords and bad authentication. To fix this and keep things reliable, banks have to step up their game with much stronger prevention systems.

Another major issue in e-banking is the fraud patterns which keeps changing continuously. Unlike old-fashioned scams, digital fraud moves fast and changes with every new piece of tech, making it a nightmare to catch. Scammers are constantly finding ways around the usual security, so banks need flexible, smart models that can flag suspicious activity the second it happens. Tools like watching transaction patterns, analysing how users behave, and using risk-based logins are now the standard for stopping fraud before it starts.

Where these tools use large amounts of transactional data and identify unusual patterns which may indicate fraud behaviour. The major challenges in the fraud transaction is very few compared to genuine transactions, that creates an imbalanced dataset problem because of this, the traditional models sometimes fail to detect fraud accurately and may give too many false alarms.

Researchers use advances analytical techniques such as **Hidden Markov Models (HMM)**, clustering methods like **DBSCAN**, and oversampling techniques like **SMOTE** to balance the data and to improve the model accuracy. New deep learning models like Tab net are also used to make fraud detection more accurate, scalable and easy to interpret. These models help banks in better decision-making by identifying suspicious transaction patterns while following financial rules and regulations. One of the most important concepts is **Explainable Artificial Intelligence (XAI)**, which helps banks understand how a fraud detection model works. Explainable AI increases trust in automated fraud detection systems and supports financial analysts in making correct decisions.

Security issues remain a major challenge in e-banking adoption. Common cyber threats include phishing, keylogging, identity theft, card fraud, and hacking of accounts. Transaction monitoring systems are one of the most effective ways to detect fraud, as they continuously track customer behaviour and transaction patterns to find irregular activities.

Apart from these technology, human factors play a big role in e-banking fraud. Ethical responsibility and proper digital behaviour are important to reduce fraud. Banks should focus on customer education and awareness program.

Government initiatives promote digital payments and financial inclusion have encouraged many people to use online banking. Hence, it is necessary to implement smart, fast, and scalable fraud detection system that can handle larger transaction volume efficiently and accuracy.

Fraud detection in e-banking is a complex issue that requires a mix of technology, users' behaviour analysis and regulatory measures. Old-school, rule-based methods just aren't cutting it anymore. That's why we're seeing a shift toward smart, data-heavy tools like machine learning and deep learning to

actually catch fraud in the act.

At the end of the day, e-banking has completely changed the game by making life easier for customers, but it definitely comes with higher risks. Between the sheer volume of digital deals and the cleverness of modern hackers, we need security systems that can learn and adapt on the fly. If banks mix high-tech AI with better security rules and some basic customer education, they can stay ahead of the curve and keep digital banking safe

Thus, this study focuses on using predictive analytics techniques to detect fraudulent transactions in e-banking and improve financial security and customer trust.

Objectives

1. To Evaluate Advanced Deep Learning Architectures in e-banking
2. To Mitigate Data Imbalance Issues
3. To Analyse Ethical and Behavioural Factors
4. To Enhance Detection of Phishing and Social Engineering

Review of Literature

Nuha, M., Mahmud, S., & Sattar, A. (2022) in the era of fast-changing sector of banking in Bangladesh they found the fraud in e-banking has increased a lot. Their research was about why this fraud occurs and especially how lack of user awareness makes people more likely to become victims. They used both primary and secondary data for their study. Which resulted into a strong link between lack of awareness and fraud that is about 76% of people did not know about e-banking fraud techniques and 86.3% of fraud victims had never received any formal awareness training. So, they analyse because of low awareness is a major reason behind the fraud incidents.

Kian, R. & Obaid, H.S. (2022) they explained that the banks are very important for the economy, but fraud transaction affected badly with their profits and performance. The study focused on banking transaction using big data clustering techniques based on customer behaviour patterns. They used unsupervised methods that can find suspicious patterns without calling the pre-labelled data. The result is very effective for finding unusual or any other fraud behaviour in large data sets that helps the banks to protect their financial transactions

Ahmad, I., Iqbal, S., Jamil, S., & Kamran, M. (2021) they stated that in the era of digitalization which has created a worldwide risk in digital fraud, even the cybercriminals use tools like malware, phishing, and identity theft. This research aimed to review the existing security technologies which is used by the banks and to find the remaining weakness they also analysed many research articles from other online databases. At the end they concluded that the technologies like multi-factor authentication and biometrics are very helpful, but human mistakes are still the biggest weakness. Hence, the security technology must always be combined with user education.

Danaa, A. A. A., Daabo, M.I., & Abdul-Barik, A. (2021) this research discussed how Machine Learning can detect new fraud attacks but it also faces a problem called “imbalanced data,” where the genuine transaction is much more than fraud ounces. They aimed to design an accurate fraud detection system that reduces false alarm while catching the fraud activity. Where they used combined framework including Hidden Markov Models (HMM), DBSCAN clustering, and SMOTE sample techniques. Where the study concluded with the integrated model improved for detecting fraud accuracy and to perform

better than standard machine learning models for imbalanced datasets.

Aluko, S.E. (2017) the author explained that the e-banking has increased the competition among banks but it also increased cybercrime risks. This research focused on using the data mining and unsupervised machine learning for detecting uneven transaction. The methodology uses K-mean clustering and Principal Component Analysis (PCA) on 8,641 e-banking transaction records. The result is combining multiple techniques which works better than using only one method and the clustering methods are very useful for larger, and unlabelled datasets in real-time fraud detection.

Abreu, R., David, F., & Segura, L. (2016) this paper argued the e-banking services which requires a major change in traditional banking practices and also need an ethical approach of security. They focused on understanding fraud from the perspective of ethics and daily moral issues where they used ethics literature and public data from Portuguese Banking Association. This concluded with identifying the ethical risks and security threats which helps to reduce fraud, and bank management it must continuously reassess these risks.

Latif, R. M. A, et al. (2018) these authors described phishing as a method of fake websites where copy of real banking websites to steal user information. This research focused on identifying the features of differentiate legitimate, suspicious, and phishing websites to improve automated detection where they used machine learning models like J48, Naïve Bayes, Random Forest, and LMT with the WEKA tool. This showed the result of random Forest algorithm provided the highest accuracy in detecting phishing websites compared to other models.

Aburrous, M., Hossain, M. A., Dahal, K., & Thabtah, F. (2010) The author explained the real-time phishing website detection is difficult because it involves many dynamic and subjective factors. They aimed to develop a new smart and reliable model using fuzzy data mining techniques the methodology was to combine fuzzy logic and data mining algorithms and to evaluate phishing attacks using six criteria in layered form. The study concluded that “URL and Domain Identity,” is the most important factor in deciding whether a website is phishing or legitimate.

Afolabi, O. (2024) he stated that traditional rule-based fraud detection systems cannot keep up with constantly changing fraud techniques. They focused on combining predictive analytics with Explainable Artificial Intelligence (XAI) for fraud detection in banking. The methodology reviewed existing XAI approaches that detect anomalies and also explain the decision-making process. It concluded that using XAI improves fraud prevention, increases transparency, helps meet regulatory requirements, and builds customer trust.

Guo, C., Wang, H., Dai, H. N., Cheng, S., & Wang, T. (2018) this highlighted the e-banking fraud is growing quickly in China and traditional rule-based detection is not enough. They focused to improve fraud detection using both real-time transaction monitoring and historical transaction analysis and they proposed a big data framework using Kafka, Spark, and MPP Gbase. So, the conclusion was to combine real-time monitoring with offline historical analysis is very effective for detecting fraud in real-world datasets.

Barker, R. (2018) he discussed on the rise of e-banking has allowed fraudsters to target customers more easily than sometimes the responsibility shifts from banks to customers. The author focused on the role of customer negligence and knowledge management in improving security. The methodology used to analysis critical analysis of legal issues and knowledge-sharing frameworks. This study concluded to proactive communication and awareness sharing are the essential empower customers and to maintain

trust in the banking system.

Abu-Shanab, E., & Matalqa, S. (2015) they explained that the e-banking improves the customer satisfaction, many users still does not trust it due to security concerns. This paper focused on identifying the security issues, attack types, and detection and prevention strategies where they used the expert's opinion to rank different security techniques. They study concluded that the transaction monitoring is the most effective security method, while the virtual keyboards and device identification were to considered less effective by experts.

Ramesh, D., Reddy, K. S. K., & Yamuna, L. (2025) this paper explained the rapid global use of e-banking which has transformed the financial sector but also introduced the risks like identity theft and cyberattacks. It also focused on how machine learning and deep learning models can detect and reduce these risks. The methodology focused tested models like Random Forest, Support vector Machine (SVM), and LSTM neural networks on real-time transaction. So, they concluded this with integration of intelligent models into e-banking systems it improves security, increases customer trust, and strengthens overall operational performance of banks.

Methodology

The methodology begins with multi-source data acquisition and feature engineering, where the transaction data, User Behaviour Indicators, and web features such as SSL status are collected. Unsupervised clustering is applied to build behavioural profiles and detect early anomalies.

Data are pre-processing and imbalanced correction are performed, as fraud transactions from a very small portion of the dataset. This stage includes data cleaning, normalization, categorical encoding, and the use of SMOTE to generate fraud samples for balanced model training.

The system is hybrid modelling engine that integrates TabNet for a deep learning-based pattern recognition, Hidden Markov Models (HMM) for a sequence-based anomaly detection, and Fuzzy Logic to handle uncertain and borderline fraud cases. This combination improves detection accuracy across different fraud patterns.

It uses to ensure transparency and trust, Explainable AI (XAI) techniques such as LIME and SHAP which are incorporated to provide interpretable reasoning for flagged transactions. Finally, the main evaluation measures used are ROC-AUC and recall as key matrices, and a Human-in-the-Loop approach is used to manually verify medium-confidence cases and continuously refine the system.

Analysis and Findings

The analysis of this study was done by comparing various research papers which is related to fraud detection. Also, by using predictive analytics in order to know the common trends in them. There was one of the observations which was repeated in the studies. Shows that the fraud transactions are very less compared to genuine transactions. In various banking information fraud cases are only a small set of data. Because there is not a proper stability. The predictive models sometimes face difficulty in detecting the fraud cases or patterns.

So, when we compare the traditional ways of identifying the fraud detection the predictive analytics does it better. When we see this, it becomes clear that modern methods of detecting fraud are better than traditional methods. And it also provides the best performance in most of the time. Earlier in Banking sectors there was a fixed rule-based detection was usually used in order to identify the fraud or

suspicious activity in Banking. And these models were not that strong in terms of detecting fraud. Some people can easily hack the system and change the settings where it can easily bypass without getting caught. In various studies it shows that predictive models are far better. Because they use the past or historical data to identify the different patterns by using previous patterns of data. It can detect easily. Because of this it becomes more effective by learning all previous patterns.

In many studies they have compared various algorithms of fraud detection model like logistic regression. This was commonly used in earlier research. Because of its simplicity. In recent studies we have seen the models that are Random Forest and Gradient boosting. It shows that it provides better results. These models are very strong because it can identify and detect the complex patterns of fraud in the transactions data. It can also detect the nonlinear relationship between the variables. These are the reasons which makes it more suitable for fraud detection.

In another study we have seen that fraud does not happen randomly. When it happens usually it shows some behaviour. For example, multiple transactions, large number of transactions and unusual timings of transactions or user spending behaviour which was mentioned in various research. So here the predictive analytics plays an important role in identifying the patterns which are difficult to detect or identify manually. It can detect the complex patterns easily by noticing unusual behaviour in transactions.

Even though the predictive analytics provides the better results. It has some limitations like it detects the transactions which are genuine. Which can make the people worried and reduce the trust in digital banking. So, it can affect customer experience and they may stop using digital banking. Here this is one of the challenges faced in predictive models.

There was another issue which was noticed in performance of the models. Because they are not consistent in terms of providing accurate result, some models give different results -because the fraud happens rarely. Many research suggests of using additional metrics like precision, recall and F1 score. So that we can get more realistic output from the model.

Overall, when we see the various research regarding predictive analytics. We can notice that it has been improved in fraud detection system in digital banking. Also, we have some challenges like inconsistent result, false detection. It is an important note to be made in real-world application.

Key Findings

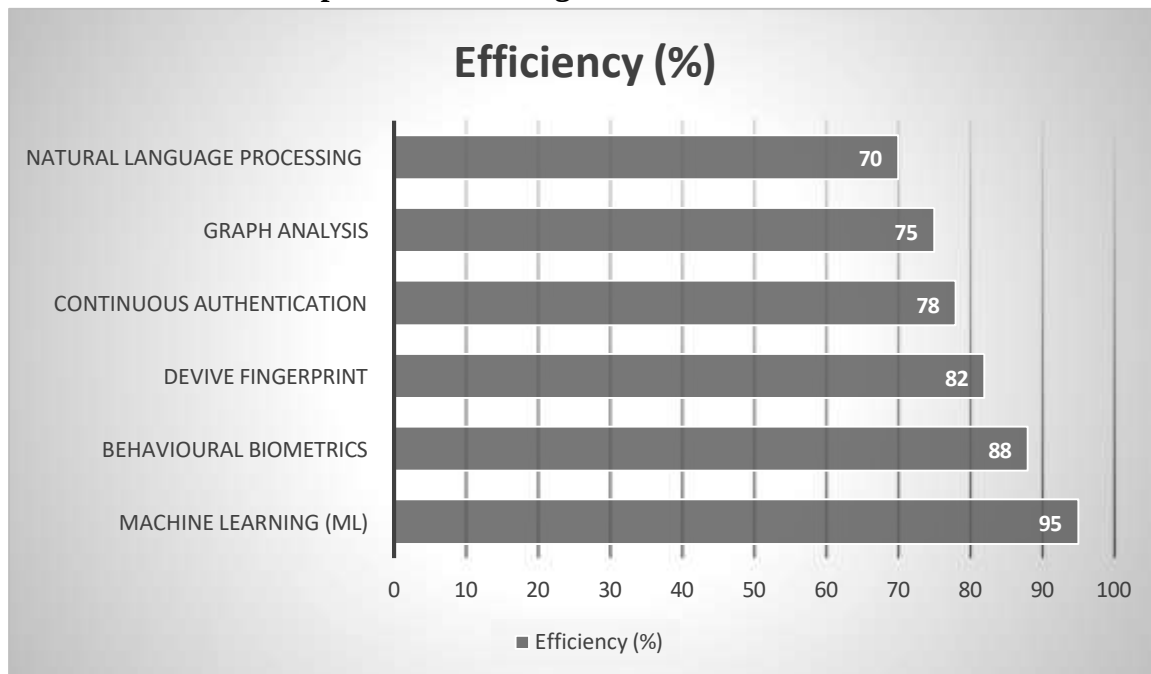
It is based on observation of previous research papers, analysis and comparison that has been written in this section.

1. In various studies it is found that fraud transactions are just a portion of the total transactions which affects a lot.
2. We have also seen that predictive analytics provides better results compare to traditional methods which use rule-based fraud detection system.
3. The models known as Random Forest generally gives better result and also when it comes to handling the complex tasks.
4. We have also seen that the fraud does not happen randomly. It shows some unusual behaviour like large number of transactions, unusual timings etc.
5. There are also some of the advanced models. Which provides accurate result and handles various and large complex patterns. But it slightly cost more and small financial institutions cannot afford it.

These findings shows that the predictive analytics is strong and provides good support for fraud

detection in digital banking but optimization is required for the real-world application.

Graph Chart Showing Modern Fraud Defences



Explanation

This chart illustrates how banks have shifted away from simple, fixed passwords toward a more dynamic "living" defence that prioritizes behavioural patterns and real-time machine learning. By analysing how a user actually interacts with their device rather than just checking a static credential, these advanced systems can identify fraudulent activity with far higher accuracy before any money leaves the account.

Conclusion

This study shows that the role of predictive analytics in fraud detection in digital banking. Where it goes through the various patterns to Identify the fraud. By using the historical patterns. It can detect the fraud easily. Overall, when we see the current system. We can say the predictive analytics has become important tool in modern fraud detection system.

In Advanced model we can see that it is much stronger in terms of identifying patterns and fraud detection. Also, it is very much consistent in providing result. Where in traditional methods it uses rule-based fraud detection which can be bypass easily and it is fail to detect.

The key findings shows that how important the predictive fraud detection in day-to-day life of the people. We have seen a rapid growth in digital banking and large number of transactions are made in a single day. So, in order to be secure or the people transactions these models should be used for fraud detection or any suspicious activities. Even though the predictive analytics is strong and support e-banking, it has some limitations like detecting genuine transactions, inconsistent performance, false positives which makes concern for its application.

In future research we can focus on improving the fraud detection and making it more strong, accurate and stable. By integrating the hybrid machine learning methods and real time analytics. The Artificial Intelligence (AI) can contribute a lot automated fraud detection. Which makes it more advanced in

predictive analytics fraud detection system. Also it can become more powerful and reliable which can become more accurate in future.

References

1. https://link.springer.com/chapter/10.1007/978-981-15-7394-1_6
2. <https://d1wqtxts1xzle7.cloudfront.net/96035708/228181713-libre.pdf?1671459050=&response-content-disposition=inline%3B+filename%3DAn+Examination+of+E+Banking+Fraud+Preven.pdf&Expires=1772035478&Signature=PltMLGQKnGUT7rF~eIx~PacGAFjvd98iZC~7SHVWZRqg8msuinSJD6ryMWh9r8oAIPiwbJBbGoi8frzWn-uzu5m1BQagglvNqQBVM03KBxwZbX~OKG0D2fKAkfhGun0YX99H6H5luIYVfW089oUJE57oKH056ppOozvVCiby2~fL53lNZgbMnRwHeNxbZcyLw2zaJ658Az9Y~nPgCtMGrtyZpRYdRElNkDPeweIrjWl9vslz-DN-Z9Y5K6m478b2X6Lq8Wqdwg0Qgbg8~SjkcCtqntXYNr0Y4~5RdGzFdckEy4Ky-c9AG2b5V6eFADSVsEkGseLKkzWswiDsESnmw &Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA>
3. <https://ieeexplore.ieee.org/abstract/document/8511873>
4. <https://ieeexplore.ieee.org/abstract/document/8667255>
5. https://www.academia.edu/33450273/Security_and_Fraud_Issues_of_E_banking
6. https://link.springer.com/chapter/10.1007/978-981-15-7394-1_6
7. https://www.academia.edu/95271038/Detecting_Electronic_Banking_Fraud_on_Highly_Imbalanced_Data_using_Hidden_Markov_Models
8. <https://irep.ntu.ac.uk/id/eprint/47097/>
9. <https://ieeexplore.ieee.org/abstract/document/10532788>
10. <https://doi.org/10.38140/com.v23i0.3721>
11. https://www.researchgate.net/profile/Shahid-Iqbal-29/publication/352668394_A_Systematic_Literature_Review_of_E-Banking_Frauds_Current_Scenario_and_Security_Techniques/links/60d28a5d299bf19b8d9d_9446/A-Systematic-Literature-Review-of-E-Banking-Frauds-Current-Scenario-and-Security-Techniques.pdf