Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

Analysis of Artificial Intelligence Techniques for Prevention of Financial Fraud

Rohit Raj
Department of Computer Science and Engineering
Amity University Jharkhand
Ranchi, India

Shikha Priya Choudhry
Department of Computer Science and Engineering
Amity University Jharkhand
Ranchi, India

***Abstract*— In today's era, fraud is very common in all aspects of life. With the advancement of technologies like artificial intelligence or big data are rising in detecting the financial fraud with the help of machine learning.**

This research paper will be about the analysis and implementation of the AI technology to prevent financial fraud.

Robbing a bank with gun has now become obsolete. Now the fraudster commit robbery just by seating at their home. Frauds are the biggest challenge for the finance industry. Credit card fraud is the most common type of fraud and as per report 270,000 cases were reported in 2019. It has now become very important to detect the financial fraud and to be aware of safe future transactions. AI technologies are used to prevent and detect financial frauds.

In the past few years, many studies have used machine learning techniques to detect the fraud.

In my whole paper I will work on how the technology of Artificial Intelligence can be useful in preventing the financial fraud.

I. INTRODUCTION

→Robbing a bank with gun has now become obsolete. Now the fraudster commit robbery just by seating at their home. Frauds are one of massive challenge for finance industry. Credit card fraud is the most common type of fraud and as per report 270,000 cases were reported in 2019. Some studies suggest that in USA alone a loss of 17-billion-dollar credit card fraud was associated.

Fraud has become very common in many industries. The 'Juniper Research' has found three industries where the fraud is at peak level –

1. $16.6 billion (e-retail) - 65%
2. $6.9 billion(baking) - 27%
3.$1.5 billion (Airline ticketing) – 6%

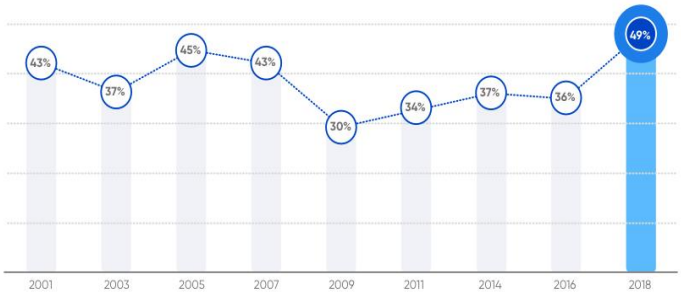

Figure-1

The financial institution try many methods to prevent fraud. But, fraudster are very adaptive to the methods, over the time they learn how to overcome this protective models. Fraudsters are very intelligent and fast learner. In short, we can say that the best methods implemented by financial institutions the for the fraud detection fails and fraud continues.

Development in the technology in AI and ML (which is part of AI) is playing important role in detecting and preventing fraud. In the upcoming sections We will learn that what exactly is AI and financial fraud and how AI can be implemented to prevent financial fraud.

II. LITERATURE REVIEW

→The 21st century is a century of different modern inventions and technologies. This century has witnessed many discoveries and inventions and will witness many more. A technology which progressed in every field was AI from engineering to medicine, education to sports it covered all the dimensions**. *(Halal (2003),Masnikosa (1998), Stefanuk and other scientists*.** It is believed AI will continue to grow and will play an important role in various fields.

It involves complex things such as feeding a particular data into the machine and making it react as per the different situations. The basic concept of AI was to create self-learning patterns where the machine can give answers to the never answered questions like human would ever do. According to Canada based company report, Global AI report 2019, India stood ninth in number in AI specialists working in the field. The US, China and UK topped the list.

The field of AI has grown extremely to an extent that tracking expansion of studies becomes a strenuous task**, *Cristani (1999) and Goyache(2003))***

It has brought revolution and provided us with some fascinating facilities like Gaming, Natural Language Processing, Image Processing and Vision System, Virtual Personal Assistants, Self Driving Cars but fraud detection has been a much needed one. There's no doubt that AI has given many satisfactory services to us, but it has also invited the fraudsters to commit modernized fraud.

(***Raghavan and Parthiban, 2014)*** There have been significant rise in cases of financial frauds committed in e-banking sector, credit card frauds such as phishing attacks, which is results into a sense of insecurity among users. It's also a huge financial loss to the country.

It has become a necessity to create a secure environment to control frauds, and AI can play an important role in this. Work on learning algorithms and methods have been applied for data analysis and anomaly detection. Attempts are made for learning method of supervised, unsupervised, and artificial neural networks approach Also proposal has been made for creating fraud detection system which can be compatible to the behaviour changes.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

III. FINANCIAL FRAUD

→ When someone stoles the money digitally which harms our financial condition through illegal practices are known as financial fraud.
There are many types of financial fraud, some are explained below: -

3.1 Credit Card Fraud

→Credit card fraud is consider to be common type of cybercrime. The personal data of the users can be hacked by phone calls, wi-fi hotspots and emails by the fraudsters.
Fraudster just steals the cardholder information by the online means, and they do not need to hold your card in their hands.

Figure-2

3.2 Mobile Fraud

→ Implementing the bots are the most common types of mobile fraud. In the year 2016, bot was created by fraudster on Facebook through which about 10,100 accounts were hacked by offering the users to install the browser extensions. User's personal data including financial sites were hacked.
Fraudster used to make phone calls to the users and claim that you have won the lottery and asks you personal details to transfer the money but at the end the money from our account are transferred to fraudsters account.

Figure-3

3.3 Internal Fraud

→In this, the fraud is committed by an employee against the organization. The employee uses the resources and assets of other employees for their own personal use.
As stated by NetGuardians, 70% of banking fraud is internal and most of them remains undetected. This happens because an employee is trusted person who has little supervision.

Figure-4

3.4 Money Laundering

→Money laundering is the illegal process in which large amount of money is generated by criminal activity like terrorist funding and drug trafficking. The money from the criminal activity is illegal but the launders make it look legal.
The origin of the money is illegal, but it looks like legal.

Figure-5

3.5 Identity and Social Fraud

→Fraudsters commits identity fraud when user uses internet banking via the internet. They obtain the personal data which is uses to steal the money digitally.
In social fraud fraudster uses social media sites to commit fraud. They try to send direct messages to people to extract personal details through which bank details can be accessed.

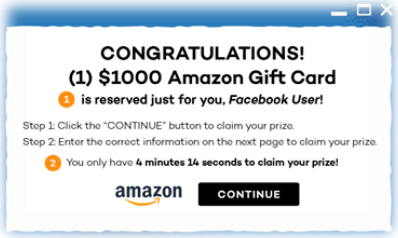

Figure-6

IV. ARTIFICIAL INTELLIGENCE

4.1 What is AI?

→Before analyzing the problem statement and finding the solution we will learn that what exactly is AI and how does it work?
AI is the simulation of human intelligence processed by the machine. It has many applications, and it is becoming important to us as it can be useful to solve composite problem accurately. Some of the applications include- AI in healthcare, astronomy, gaming, education, data security, social media, detecting financial fraud and many more.

4.2 How does AI work

→ AI works on a foundation of special software and hardware for training and writing the machine learning algorithms with the help of programming languages like python, R and java.
AI works on combing the large amount of data with intelligent algorithms which allows the software to lean automatically from the patterns which is present in data, there are various subfields of AI which include:

1.Deep Learning
2.Computer Vision
3.Congnitive Computing

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

4.Neural Network
5.Machine Learning

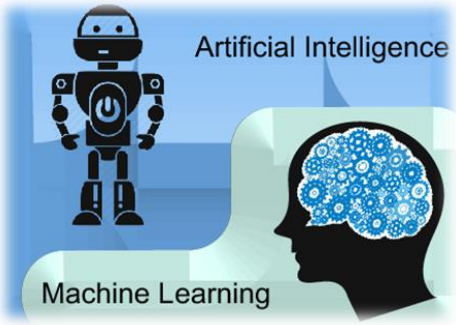

Figure-7

In my paper I will be applying machine learning to for the detection and prevention of fraud.
Machine Learning is the branch of AI which makes the decision based on the past experiences. It identifies patterns and analyze the past data to reach a possible conclusion without the involvement of human and works upon on that.

AI mainly focuses on three cognitive skills:
1.Learning process
→ It focuses on obtaining data and creating the rules to turn into accountable information. These rules are called algorithm.
2.Reasoning process
→ It focuses on choosing the right algorithm for the particular task
3.Self-correction process
→ It focuses on fine-tuning the algorithm so that they can provide accurate results.

V. RISE OF FINANCIAL FRAUD IN COVID-19 PANDEMIC

→The world has experienced the first covid-19 case on 31 December 2019 and due to the lockdown, lots of people have pushed themselves into the digital world. Lots of people have adopted various digital tools to study, paying bills, banking transactions and many more.
According to recent studies it has been found that the lockdown period acted as boost for digital banking services. The investments in the digital bank initiatives fall into two categories- specialized features and enabling hygiene. With the outbreak of covid 19 pandemic these investments have accelerated a lot.
This accelerated digital activity and attracted the attention of many criminals which lead to increase of financial fraud cases. According to the reports of VMW from last year, between February and March, attacks to the financial sector have been increased to 238% globally. From the recent reports it has also been found that in India, the fraud was committed from credit card and debit card details which was due to online banking and shopping.

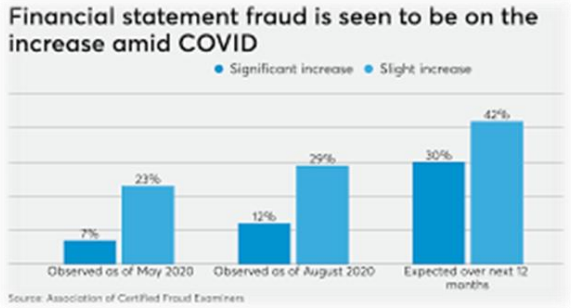

Figure-8

The financial services are always prone to the cyberattacks and the emerging economy and legal infrastructure are more prone to cyberattacks. The cyber criminals are very hard to be traced out and the victims rarely get the compensation. This is the main reason of the bank fraud.
The most banking fraud in our country India consist of credit card fraud which consist of card cloning which is rising significantly over the years. India was one of the first country who adopted the cashless payments with the help of QR code-based system which is consider to best in terms of security both for bank and the customers. But the lack of end point security at the device level has been the reason for the fraud transactions. Impersonation fraud is also increasing in India over the years in which the fraudster obtains the KYC details of the victims via online phishing from social media and email
According to the recent report published by McAfee the person who use cloud-based email services in the finance sector are the targets of the fraudsters.

VI. WORKING OF MACHINE LEARNING IN DETECTING FINANCIAL FRAUD

Figure-9

→First of all to detect the fraud, machine learning model first requires to collect some of the data. The model then analyzes and collects all the data which is gathered, segments, and extracts the require features from it. At next step the machine learning model receives the training set that teaches it to predict the probability of fraud and then it creates fraud detection machine learning models.
In first step there is data input, which is differs in machine learning and humans. Humans' can sometime struggle to understand that massive amount of data, but it is just a piece of cake for machine learning. The more data ML receives the better it can learn and improve the fraud detection skills, that's the positive side of machine learning.
The next step is all about extraction. In this step, features describing the good behavior and fraudulent behavior are added. These features usually include the location of customer's, identity, orders, networks, and choses payment

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

method. Bases upon the complexity of the fraud detection, the list of investigated features can also be differed.
In next step training algorithm is launched. This algorithm is a set of rules that ML model has to be follow to decide that an operation is legitimate of fraudulent. The more data ML is provided the better it will work.
Finally in the last step as soon as training is over, the company receives a machine learning model which is capable of detecting fraud is suitable for their business. This model can detect the financial fraud in no time with higher accuracy. To detect credit card fraud, a machine learning model needs to be constantly improved and updated. It is very important to update and enhance the methods and algorithms timely because fraudsters always comes up with new trick to commit fraud.

VII. ARTIFICIAL & COMPUTATIONAL INTELLIGENCE TECHNIQUES

→There are various artificial and computational intelligence model to detect financial fraud, they are set of nature-inspired methodologies to address the real world related problem.
Some of the models are explained below-

7.1 Genetic algorithm

→ It is an evolutionary algorithm which basically aims at obtaining better solution as the time passes. It was first introduced by Holland. This algorithm has been successfully applied to many problems in sectors like astronomy, sports and also in detecting financial fraud. They are also used in data mining for variable selection. This algorithm can be implemented with the help of Java.

7.2 Genetic programming

→It is an extension of genetic algorithm. It is population of random classification rules(GP individual) which is generated first than each rule is evaluated in order to receive a fitness value and the better rules have more and more chances of survival. In recent years this algorithm is very useful in detecting financial fraud.

7.3 Scatter Search

→It is an algorithm which shares some common characteristic with Genetic algorithm. It operates on a set of solutions and reference set and combines these solutions to create new ones.

7.4 Big data Technologies

→Big data is an indicator for large amount of data, increasing in size and exponentially with time. It can be defined as a software tools which is used for analyzing, processing and extracting data. Bid data technologies can be used to detect financial fraud using data analytics.

7.5 Block chain Technologies

→ It is the type of database and it is defined as the collection of information which is stored electronically in the system. Data is structured in database in such way that allows for easier searching and filtering data for specific purpose. Anti-money laundering can be built with the help of blockchain technologies.

VIII. MODELS AND METHODOLOGY

→ The detection system which exists currently depends upon a defined criteria and learned records, which sometimes makes it difficult to detect new patterns of attack. New patterns can be discovered, and higher detection accuracy can be achieved with the machine learning methods that are based on supervised learning and unsupervised learning. This paper analyses the recent methods for detection of financial frauds studies of machine learning from the year 2016 – 2017. In this research paper machine learning is implemented to detect financial frauds. In ML, the unsupervised learning method could be applied to discover threats for accurate detection of fraud transactions. The overall process of detecting the fraud includes sampling process for imbalance problem and feature selection process of accurate model. This research paper will perform the overall process for detecting the financial fraud based on supervised and unsupervised learning and sampling process and feature selection process will be applied to solve data unbalanced problem and rapid detection in real world.

8.1 Machine Learning

→ It is the field of artificial intelligence in which machines learn several concepts using data using statistical analysis to predict input data as output value. This field is divided into supervised and unsupervised learning depending upon the learning method. In supervised learning the value of input data is predicted and classified with the given label and in unsupervised learning the data is not labelled and is also called clustering process.

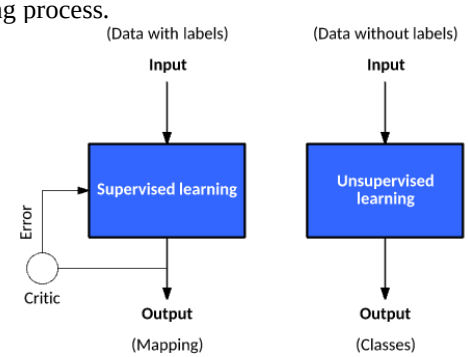

Figure-10

The model which is proposed consist of data processing, sampling, feature selection, application classification and cluster algorithm which is based upon machine learning. The validation step is also performed in each step, to verify the proposed model in financial fraud detection. In the preprocessing process, the correlation analysis of data and cleaning process of data which cleans the noise data are performed. Data transformation, integration and reduction are also includes in the process. The overall proposed fraud detection system is specified in figure-11.

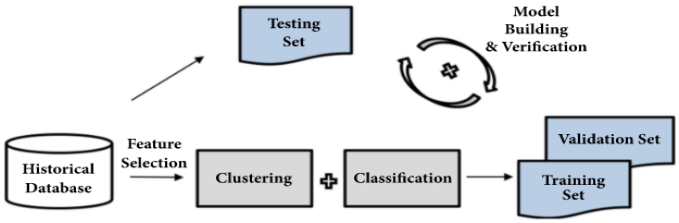

Figure-11

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

8.2 Sampling

→The imbalanced problem in the data can create problem in the process of detection and can misclassify the problem and a real transaction dataset of financial transactions always contains a data imbalanced problem. The previous research consists of oversampling method and cluster based undersampling approach to select the representative data as to improve the accuracy.

To solve this kind of unbalanced problem there are various datasets such that Minority Oversampling Technique (SMOTE) and Random Undersampling (RUS) that can give more accurate results, SMOTE is and kind of oversampling technique that uses method for generating various kind of examples rather than duplication or replacement. RUS reduces the normal transactions by simply extracting the normal sample data randomly for the class imbalance problem. Since the low ratio of anomalous data might lack in accurate results, so SMOTE and RUS both is applies for generating different ratio of sampling dataset for getting accurate results.

8.3 Feature Selection

→It has proven to be very effective and efficient for ML problems and the objective of the feature selections is to build more simple and comprehensible model thus improving the data mining process. It also removes redundancy and irrelevancy of data. There are two methods of feature selection -wrapper and filter.

The wrapper method depends upon predictive or expected performance of learning algorithms to evaluate the feature. This repeats the searching step until the desired learning performance is obtained.

In case of filter method, it is not dependent of any learning algorithm and depends upon the characteristic of data to access the features. Features are score based and hence the lowest scores are removed. This is main reason why filter method is most often used as it is fastest.

In my research paper I have selected 8 subset feature selection algorithm and six ranked selection feature algorithms. Also, the score is assigned to evaluate the feature based on frequency. The ranker algorithm is calculated by higher rank. Figure-12 is feature selection process flow chart.

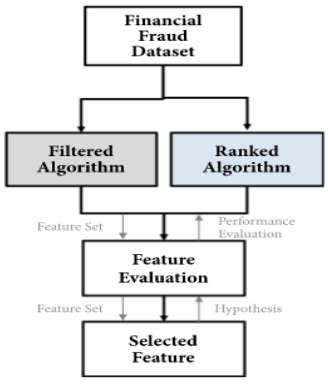

Figure-12

8.4 Deep Artificial Neural Networks

→It is the subfield of machine line and it is inspired by the functions and structure of brain called ANN(Artificial Neural Networks). AI works similar to human brain in creating patterns which is used in decision making through the capability of unsupervised learning from the data. Artificial Neural network are also called multilayer perceptrons. It is single neuron model that came before an large neural network an In neural network, the capability of prediction came from multi-layered structure of the networks. Figure-13 is an example of simple neural network. Hidden layer is the layer between output and input layer. This network is attached in the direction of the layer of input and hidden layer. Output layer is a feedforward network layer and in this layer there is indirect connection from output layer to input layer.

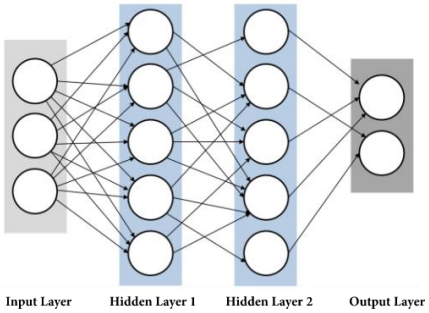

Figure-13

IX. ANOMALY DETECTION FOR PAYMENTS FRAUD

→The anomaly fraud detection and prevention solutions are more effective than predictive and prescriptive solutions. This application mostly requires a machine learning model that is trained on a continuous stream of the incoming data. This model is trained such that there should be sense of normalcy of bank such as transactions, loan applications, contents of banks or opening new account.

Then the software can alert the systems for any deviations from the normal pattern so that it can reviewed further. The system must accept or reject this alert, which then sends signals to machine learning model that it a fraud transaction or customer's information is correct or not.

This would further train the ML model to understand the deviation that it was fraud or new acceptable deviation.

This kind of model can also be applied for interactions with various banking operations Basically the fraud comes the frauds can come for merchants and the issuers and their transactions information can be further used to train and learn in machine learning model to recognize secure transactions.

Another possibility is the spending behaviour patterns that allows the machine learning model to detect the fraudulent details of e-commerce websites. The geolocational data can play important roles on this model as it is a known fact that fraud transactions always occur far away from the account holder.

X. SURVEY REPORT

→A survey was conducted my from around 100 people of different age group as to clearly know that how common is financial fraud among the people and how much they know about artificial intelligence. Do they really believe that artificial intelligence can really detect financial fraud or not?

(i)From survey it is clear that 78% of the responders are from age group 15-25 and maximum of them has faced financial fraud as per their response.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

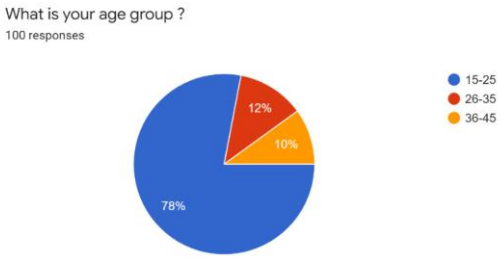

Figure-14

(ii)52% of the respondents have faced financial fraud ion their life which is the concern of problem amongst the age group of 15-25.

Figure-15

(iii)Covid-19 forced the people to switch to online platform for their certain transactions as they were having no choice as there was a strict lockdown. So as per the survey 88% of the respondent switched to online platform for their transactions

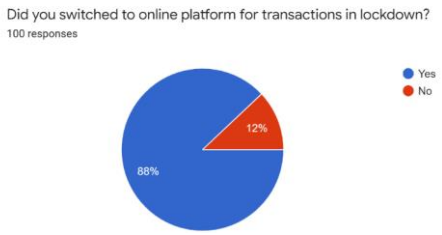

Figure-16

(iv) Before covid to till data people used online platform only for their transactions. Some people feel safe to use and some not. According to survey conducted 70% of the respondents use online platform

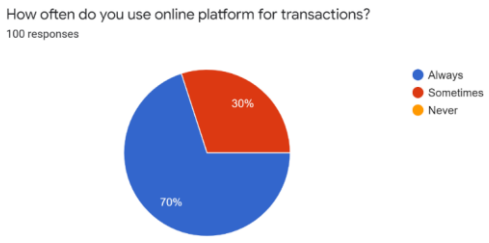

Figure-17

(v) Technology is emerging day by day. Artificial intelligence and machine learning(which is subpart of AI) has passed in detecting the financial fraud. So it was important for me to know the if people are really aware of these technologies or not. 74% of respondent and 73% of them are aware of ML and AI respectively.

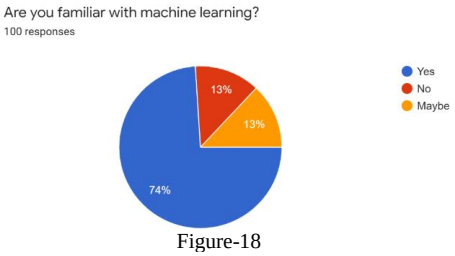

Figure-18

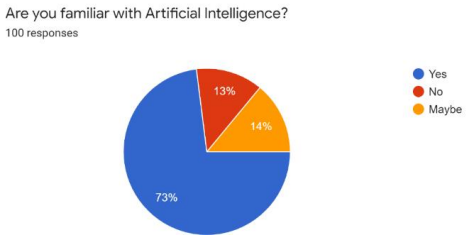

Figure-19

(vi) Now comes the most important part of my research paper. Do the respondents really believe that AI can be helpful in detecting financial fraud or not. I got the positive output 65% respondents believe that yes this technology can be helpful to detect financial fraud.

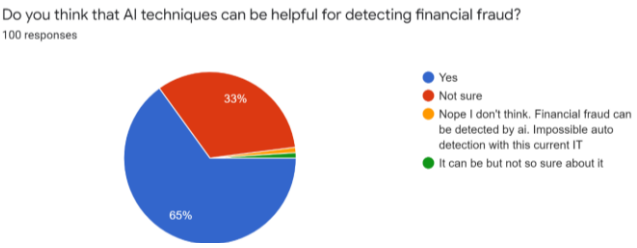

Figure-20

XI. MAJOR CHALLENGES FOR ADOPTING TECHNIQUES TO DETECT FINANCIAL FRAUD

→AI and machine learning has proved to be a successful technique to detect financial fraud over the time. But none of the technology is 100% perfect. They too have some limitations and challenges that are important to mention.
Some of the challenges are-

11.1 Changing fraud patterns

→This is one of the toughest challenges to address since the fraudster always tries to find some new method around the system to commit the fraud. So, it is very important to build such a model that can easily be updated with time to evolve patterns to detect the fraud. This results the efficiency of model.

11.2 Class Imbalance

→To be very practical very few percentages of customers have the intentions of fraudulent. So there is an imbalance in the classification of the detection of fraud models which makes it harder to built it. This challenge results into poor users experience for genuine customers since tracing the fraud declines the legitimate transactions.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

11.3 Model Interpretations

→This challenge is connected to the explainability concept as models gives some certain scores which tells that a transaction is a fraud transaction or not and fails to explain that why this is a fraud transaction.

11.4 Feature generation can sometimes by time-consuming

→Some data can be sometime so huge that it can take long period of time to generate a feature set which slows down the time to detect financial fraud.

XII. FUTURE SCOPE AND CONCLUSION

→The idea of implementing neural networks for detecting fraud can be further explored in the future. The only problem with the neural network is that it requires appropriate amount of training data so that it can work with higher accuracy. More data sets and larger fraud transactions is required to train larger models.

Deep learning networks can be also used for the fraud detections. Generative Adversarial Networks are used to detect for anomaly detection in cybersecurity for the detection of cyber-attacks. GANs are simultaneously used to train the generator and discriminator on the healthy high data which is used to identify anomalies in the eyes of human. This could also be used to detect and prevent fraud.

Fuzzy system for anomaly detection can be paired with neural network for better outcomes and to provide best models out of it. Genetic programming can also combine with neural network to provide good optimization. Fraudsters are intelligent, they find new way to commit fraud. Therefore in the future, development of good models with the help of artificial intelligence will play important role in detecting fraud.

After working on my research paper, I came to conclusion that the technology of artificial intelligence can be useful in detecting financial fraud though it has some limitations too. But the good approach can remove those limitations. I learnt about artificial intelligence, it's working and I also learnt about machine learning which is the subfield of AI which can prevent the financial fraud.

XIII. REFERENCE

[1] Raghavan, Pradheepan & Gayar, Neamat. (2019). Fraud Detection using Machine Learning and Deep Learning. 334-339. 10.1109/ICCIKE47802.2019.9004231.
[2] JOUR ,You, Ilsun, Choi, Dahee, Lee, Kyungho,(2018),2018/09/25,An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment: A Survey and Implementation,5483472
[3] Alsedrah, Mariam. (2017). Artificial Intelligence. 10.13140/RG.2.2.18789.65769.
[4] Ayodele, Taiwo. (2010). Machine Learning Overview. 10.5772/9374.
[5] Reurink, Arjan. (2016). Financial Fraud: A Literature Review. MPIfG Discussion Paper Series. 16. 10.1111/joes.12294.