IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Multi Level Associated Weighted Feature Vector For Online Fraud Detection Using Machine Learning

[1]Dr.K.Mandakini,[2]Mrs.G.Prashanthi

[1]Assistant Professor,[2]Assistant Professor

Computer Science and Engineering,

[1]Anurag Engineering College, Kodad,India

Abstract: According to studies of international monetary data, the sum of money lost due to fraud in all transactions around the world keeps rising. In particular, the adoption of the Bank-as-a-Service model by financial institutions will place additional strain on payment processing infrastructure and exacerbate the already serious problem of payment fraud. Using automated machine training and Big Data analysis methods, this research aims to synthesis viable solutions for detection of fraud in electronic payment systems. Online Payment fraud has been given a lot of attention since it can result in significant economic losses and negative consequences for account holders. In order to construct a reliable model for identifying fraud, it is essential to employ efficient feature engineering. It is discovered, however, that the present feature engineering that relies on transaction frequency has flaws. Convenient as they may be, fraudsters are drawn to modern money transfer services to conduct scams in which unsuspecting victims are tricked into sending money to fake accounts. As a result of the limitations of conventional rule-based methods for detecting fraud, machine learning models have found widespread application. Most studies build features by extracting patterns using raw transaction records because learning directly from transaction data is challenging owing to its enormous dimensionality. Recency, frequency, financial, and anomaly detection are some of the common labels used to classify these characteristics in the existing research. Using real-world transaction data, features can be extracted and most relevant features can be selected to train the machine learning model for accurate fraud detection. The temporal properties of user transactions are shown using frequency-based feature engineering, however the characteristics of fraud and the differentiation of transaction behaviors are not taken into sufficient account. In this research a Multi-Level Associated Weighted Feature Vector (ML-AWFV) Model using machine learning is proposed for accurate feature extraction and selection for further processing of online payment fraud detection. The proposed model when contrasted with the existing

Index Terms - Online Payment Fraud, Transactions, Feature Extraction, Feature Selection, Machine Learning, Financial Loss.

I. INTRODUCTION

As Internet technology advances, so does the volume of business conducted over the global network. Simultaneously, the issue of online payment fraud in network transactions has grown in importance. Low cost, broad coverage, and high frequency are hallmarks of network transactions, making fraud detection more difficult than with online transactions [1]. Online payments have rapidly gained popularity and are now accepted by many retailers. The online trading platform regularly processes thousands of transactions. Some criminals take advantage of the prevalence of online commerce by engaging in illicit activities [2]. Theft of private possessions in the complex network environment harms not only the interests of consumers but also seriously hampers the growth of the network economy. Accordingly, online transaction fraud detection is a vital tool for combating the issue of fraud in online financial transactions. Statistics and multi-dimensional

analysis are the mainstays of conventional fraud detection methods [3]. Due to the nature of these verification methods, the rules concealed in the underlying transaction data are notoriously elusive [4]. The detection methods for payment fraud are efficient because of the big data advanced technologies and machine learning algorithms. Machine learning has the potential to represent crucial features through a heap of data, something that cannot be described by more conventional statistical approaches [5]. With the help of the appropriate machine learning approaches [6], researchers can establish a model based on the existing transactions data to realize the detection of network transaction online fraud, thereby lowering the loss caused by fraud [7]. As the online payment market grows and more people use it, fraud in payments has emerged and is on the rise. The most common kind of online payment fraud is the use of a sensitive information to make a purchase. The use of a credit card or other conventional payment mechanism is not essential for the online payment method to work [8]. Instead, fraudulent payment in a mobile payment environment requires sensitive information such a card number, expiration date, card verification code [9], and PIN. Real-time credit authorization, address verification systems (AVS), card verification codes [10], rule-based detection, etc. are just some of the fraud protection technologies used by financial institutions to combat the fast growing online payment fraud problem [11]. Existing detection systems, however, rely on predetermined criteria or learned records, making it challenging to spot novel assault patterns [12]. This research presented a approach to identifying online payment fraud by using a machine learning process to uncover hidden and fundamental fraud risks [13]. As a branch of AI, machine learning has been a popular topic this decade. Machine learning is becoming an increasingly attractive investment option for businesses that want to enhance their offerings [14]. Machine learning is a set of techniques wherein a computer is able to learn and complete tasks on its own, without the need for explicit programming. Using the collected training data, the acquired model would be able to learn [15]. Experiential knowledge can be used to make predictions or take action. An efficient online fraud detection system will reliably flag suspicious financial dealings. It is crucial to make sure that legitimate users are not denied access to the payments system [16], even while doing so is required to prevent criminal actors from conducting fraudulent transactions [17]. A high rate of false positives might negatively impact client satisfaction, which could result in lost sales. Using machine learning for online fraud detection is complicated by the presence of extremely skewed data sets. The overwhelming majority of payments are legitimate, with only a negligible number of fraudulent ones present in the multiple datasets now available [18]. It is a huge problem for academics to design a detection model using best features that is both accurate and efficient, with a low percentage of false positives while still effectively detecting fraudulent activities [19].

II LITERATURE SURVEY :

Credit card payments are popular for online purchases since they are quick and simple to make. As credit card usage has grown, so has the potential for fraudulent activity. Credit card fraud results in considerable losses for the victims and the financial institutions that provide the cards. Alarfaj et al. [2] developed methods for detecting such frauds, taking into account factors such as the availability of publicly available data, the presence of high-class imbalance data, the dynamics of fraud, and the prevalence of false alarms. Many machine learning based methods, such as the Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression, and XG Boost, are available. However, applying cutting-edge deep learning algorithms is still necessary to cut down on fraud losses, as current methods are not very accurate. The recent advancement of deep learning algorithms has been the primary emphasis for this goal. To achieve productive results, a comparison study was conducted between machine learning and deep learning algorithms. Specifically, the European Card Benchmark Dataset for Fraud Detection is used for the in-depth empirical investigation. After applying a machine learning algorithm to the dataset, the author seen a little uptick in the ability to spot potential frauds. Three convolutional neural network–based designs are then used to boost the effectiveness of the fraud detection system. The detection accuracy was significantly improved as more layers were added. By experimenting with different configurations of hidden layers, time periods, and state-of-the-art models, the author performed a thorough empirical investigation. The use of machine learning to detect suspicious financial transactions is on the rise. Most application systems, however, only pick up on dishonest behaviour after the fact, rather than in real time. Because there are so few fraudulent transactions compared to legitimate ones, fraud detection is exceptionally difficult and requires approaches outside of standard machine learning.

An improved version of the Support Vector Machine (SVM) using quantum annealing solvers has been used by Wang et al. [3] to implement the detection framework. QML application's detection performance was compared to that of twelve machine learning algorithms on two datasets: a bank loan dataset that is severely imbalanced, and an Israel credit card transactions dataset that is moderately imbalanced. The outcome demonstrates that, in terms of both speed and accuracy, the quantum-enhanced SVM has clearly surpassed

the competition on the bank loan dataset. When compared to similar methods using data from Israeli credit card transactions, however, its accuracy is on par. Feature selection has also been proven to greatly increase detection speed for both datasets, but at the expense of accuracy. Online advertising is a form of marketing that employs the use of the Internet to reach potential customers. It has been expanding in recent years to meet the needs of e-commerce. Every time a potential customer clicks on an ad, the advertiser has to pay a fee. There are fraudulent clicks that can be made on a PPC platform because it is so easy to imitate real user behaviour. Massive financial losses are incurred by advertisers as a result, and the confidence of internet advertising networks is severely harmed. Common methods for identifying fraudulent clicks include dynamically customizing and interpreting data based on a machine learning model. Due to the fact that these algorithms only consider each piece of data to be a single feature vector or matrix, it is no longer possible to discover the innate connections between individual data sets. The existing fraud prediction system's relatively low performance can be attributed to the million daily fraud clicks on varying types, which further disperse the focus of models. To combat the issue of fraudulent clicks, Zhu et al. [4] implemented a tensor-based system for doing so. As part of its analysis, this research took into account the possibility of reassembling the data into a high-rank tensor, using tensor decomposition and transformation to delve into the latent meaning of each piece of data and investigate the impact of joining them together. The widespread use of the Internet over the course of the previous decade also contributed significantly to the meteoric rise of online card transactions. Increases in the volume of online banking have coincided with a surge in fraudulent activity in the banking industry around the world. Because of this, rule-based systems were developed to flag potentially fraudulent transactions and allow human specialists to verify their legitimacy. As a countermeasure, the most recent attacks take use of the static structure of rule-based systems to sneak past defences. This motivated the study of machine learning methods by Can j et al. [5], especially deep learning, for the purpose of developing adaptive fraud detection systems. To the best of the knowledge, however, none of available models pushed deeper into better understanding the features of fraudulent transactions in order to build more robust models; rather, their attention was narrowly focused on detecting fraudulent activities. Because of this, the author compiled the largest data collection ever utilized in a study, which includes 4 billion legitimate transactions and 245 thousand fraudulent ones from 35 different banks in Turkey. As a result, the author presented three different types of profile-based fraud detection algorithms and evaluated their efficacy.

A rise in the use of credit cards for both online and in-person purchases can be attributed to the proliferation of electronic commerce and the improvement of communication networks in recent years. However, credit card theft has been on the rise, resulting in substantial annual losses for financial institutions. However, most credit card datasets are severely skewed, making the creation of good fraud detection algorithms difficult but crucial in limiting these losses. As another example, traditional machine learning algorithms are inefficient for credit card fraud detection since they are built on a static translation of the input vector to the output vector. For this reason, they are not flexible enough to accommodate the ever-changing purchasing habits of credit card holders. Using a neural network classification method and a hybrid data oversampling methodology, Esenogho et al. [6] offered an effective method for detecting credit card fraud. Adaptive boosting (AdaBoost) employs a long short-term memory (LSTM) neural network as the base learner to produce an ensemble classifier. Synthetic minority oversampling and the edited nearest neighbour (SMOTE-ENN) techniques are used to accomplish the hybrid resampling. Using publicly available, real-world datasets of credit card transactions, the usefulness of the proposed strategy is proved. Support vector machine (SVM), multilayer perceptron (MLP), decision tree, classic AdaBoost, and long short-term memory (LSTM) are compared to the suggested method's performance. Online payment fraud is a growing source for concern as the world of digital money expands at a dizzying rate. Identifying fraud typically requires the use of machine learning or rule-based systems. The sliding time window is a well-known effective tool for this issue since the most important characteristics of such fraudulent transactions are displayed in a sequential way. Sliding a time window back and forth between two dates allows for the extraction of data about the characteristics of the transactions and the capture of latent patterns buried in the transaction records. However, as transaction patterns in real-world application areas are often too elusive to record, the adaptive configuration of sliding time window is actually a huge challenge.

In truth, the practical environment typically requires continuous updates and fine-tuning via human participation. To detect fraudulent online payment activities within automatically sliding time windows, Wang et al. [7] adopted an adaptive learning strategy in this work. The author therefore worked to perfect the window placement and increase the flexibility. The author created a smart window, which named the learning automatic window (LAW). With the help of learning automata, it is able to dynamically and frequently alter the time frame parameters to account for the ebb and flow of unauthorized transaction

patterns. With the rise of new forms of electronic commerce and communication technology, credit cards have become a viable alternative to cash for making in-person and online purchases; however, fraud has expanded dramatically alongside this trend.

Credit card fraud results in significant annual losses for businesses and consumers, and thieves are constantly on the lookout for new technology and methods to conduct fraud. One major barrier to wider adoption of electronic payment is the detection of fraudulent transactions. As a result, methods that are both quick and accurate are needed to weed out fraudulent charges. In this research, Taha et al. [8] offered a novel method for identifying credit card fraud by utilizing a light gradient boosting machine that has been optimized for this task (OLightGBM). For the suggested method, the parameters of a light gradient boosting machine are tuned with the help of a Bayesian-based hyperparameter optimization algorithm (LightGBM). The author conducted trials with two publicly available, real-world credit card transaction data sets, one containing fraudulent transactions and the other containing valid ones, to prove that the proposed OLightGBM is effective at identifying fraud in credit card transactions**.**

III. PROPOSED MODEL

The need for a more rapid payment infrastructure has led to a rise in the usage of online methods across government, business organizations, and other sectors. Banking activity has increased as a result of today's world's heavy reliance on technology. However, banking fraud has increased in tandem with both online and traditional banking transactions. As internet transactions have grown in popularity, so has the amount of study done to protect consumers from financial scams. The current infrastructure cannot effectively process the volume of transactions being processed. According to research, over 83% of merchants manually evaluate between 16% and 23% of orders for fraud detection, despite the fact that doing so is inefficient, expensive, and prone to more false negatives and human errors. As online banking spreads across the industry, it brings with it a host of new problems that must be solved. In spite of the importance of security, users prefer an easy authentication process with as few steps as feasible.

Multifactor authentication solutions as they stand are unable to determine whether or not a given transaction is legitimate. The customer may not be completely satisfied with the current system, but it does offer some measure of protection and can always be made better. Because of the growing volume and variety of online payment services, the application of analytical models to detect fraudulent behavior is essential.

The main aim is to examine online payment data, which includes both legitimate and fraudulent customer actions. The classification model must accurately determine whether a transaction is legitimate or fraudulent, and it must do so quickly, flexibly, and efficiently. As a result, businesses that adopt this solution will save a significant amount of money, as they will no longer lose money to fraud and will not need to invest in staff training to manually detect the same. The proposed model framework is shown in Figure 4.1.

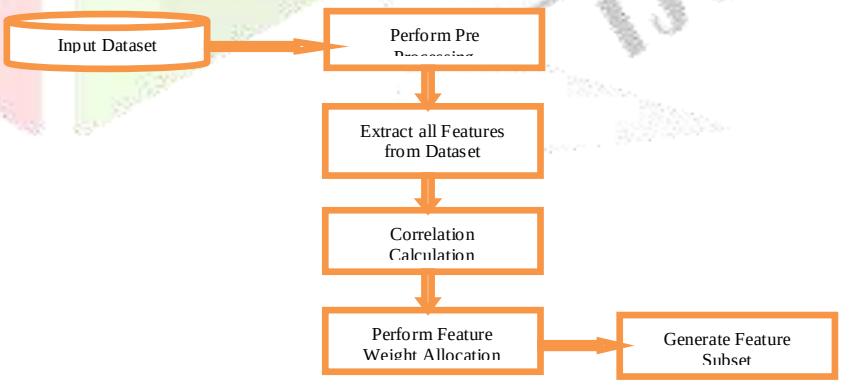

Fig 4.1: Proposed Model Framework

Machine learning-based preprocessing, sampling, feature selection, and classification and clustering algorithms are all components of the proposed model. In this research, the suggested method for detecting fraudulent mobile payments at each stage is tested. Noise in the data is removed and correlation analysis is conducted during the pre-processing phase. This procedure additionally incorporates data normalization and integration. The subsequent step is the sampling procedure, which uses a random over-sampling and under-sampling method to analyze datasets with varying ratios for verification purposes. Feature selection algorithms that rely on filters have been used in the process of extracting and choosing features. Clustering with the suggested approach is performed after feature selection, and the resulting data is used for the classification process's training and validation sets. Higher prediction might be attained by employing supervised algorithms on the previous vote, which was derived in the clustering phase.

For machine learning challenges, feature selection has been shown to be effective and efficient. Feature selection's goals include making models that are easier to comprehend and use, boosting learning efficiency, and making data that is easy to digest by eliminating clutter and removing irrelevant information. The features are evaluated using the prediction performance of a predetermined learning algorithm using correlation factor. In this research a Multi Level Associated Weighted Feature Vector (ML-AWFV) Model using machine learning is proposed for accurate feature extraction and selection for further processing of online payment fraud detection.

Input: Online Payment Transaction Dataset {OPTDset}

Output: Weighted Feature Set {WFset}

Initially consider a dataset that contains online payment transaction data which includes normal and fraud transactions. The dataset records are analyzed by calculating the mean and standard deviations for data processing. The process is performed as

$$Transac[M] = \sum_{i=1}^{M} getattr(i) + \frac{\sum_{r=1}^{M} mean(i, i+1) \varepsilon OPTDset}{len(OPTDset)} + \sum_{r=1}^{M} std(i+1, i)$$

Here i is the instant record in the dataset, r is the instant record in the dataset, mean and std are the derived mean and standard deviation functions.

Each record in the dataset is processed and pre processing is applied. Data preprocessing refers to any operation carried out on unprocessed data in order to get it ready for further processing. The accuracy and integrity of a dataset can be improved by preprocessing to remove missing or incompatible data values due to human or computer error. Pre processing ensures that information is consistent. The pre processing is performed as

$$Norm[M] = \sum_{i=1}^{M} \frac{Th - attr(i)}{\max(attr(i, i+1))}$$

$$Proc[M] = \sum_{i=1}^{M} Norm(i) - drop\big(attr(i)\big) \neq !NULL$$

After pre processing, cleaned dataset is available for feature extraction. Feature extraction is used to describe the procedure of reducing unstructured data to a set of quantifiable features that may be further processed without losing any of the original data's context. By eliminating unnecessary information, feature extraction cleans up the dataset. As a result, the learning and generalization phases of machine learning can go more quickly, and the model can be constructed with less machine effort. The feature extraction is performed as

$$\mathbf{Feat\_Set(Proc[M])} = \sum_{\mathbf{i=1}}^{\mathbf{M}} \mathbf{getatt\big(Proc(i)\big)} * \left(\mathbf{Max(attr(i, i+1))} - \frac{\boldsymbol{\lambda}\mathbf{(i+1)}}{\mathbf{2}}\right)^{\mathbf{2}} + \left[\sum_{i=1} getattr(Norm(i)) - \frac{maxVal(i) - minVal(i)}{mean(M)}\right]$$

Here λ is the model used for considering the normalized values from the record that are relevant in the dataset.

After extracting all the features, correlation factor is calculated among the available features. Correlation factor represents the relation among two features. The proposed model considers the features that are low correlated. The correlation factor is calculated as

$$Corr[Feat\_set[M]] = \sum_{i=1}^{M} \frac{\max\big(attr(i)\big) * M * \sum_{i=1}^{M} \sum(i * i + 1) - \sum i * \sum i + 1}{\sqrt{M * \sum i^2} - \sqrt{\sum i^2} * \sqrt{M * \sum (i+1)^2} - \sqrt{\sum (i+1)^2}}$$

Here i is the instant feature and i+1 is the next neighbor feature in the extracted set.

Feature selection is the process of narrowing down the features extracted used in the model by keeping just the most pertinent information and discarding any extraneous or irrelevant details using the correlation factor. Feature selection in machine learning is the act of selecting features for a model automatically based on the nature of the problem being solved for training the model. The feature selection is performed as

$$\text{Feat\_Subset}(Feat\_set[M]) = \sum_{i=1}^{M}\left(\frac{\text{mincorr}\left(\text{Feat}_{\text{Set(i)}}\right)}{\max(\text{corr(i)})}\right) + \sum_{i=1}^{M}\frac{\sum_{i=1}^{M}\min(Feat\_set(Corr(i,i+1))) - \max(Feat\_set(Corr(i,i+1)))}{len(Norm(M))}$$

This paper employs a technique called feature weighting, which attempts to determine the value of each characteristic and give it a weight accordingly that represents the priority for training. If the features were appropriately weighted, the ones that mattered the most would be given more consideration than the ones that didn't during training phase. The weights are allocated for the selected features and the weighted feature subset is generated as

$$Wei\_Set(Feat\_Subset[M]) = \sum_{R=1}^{M}\frac{\max\left(Feat\_Subset\left(Corr(i,i+1)\right)\right) - \min\left(Feat\_Subset\left(Corr(i,i+1)\right)\right) + setMax(Rand(i))}{\lambda * \text{M}}$$

The methodology section outline the plan and method that how the study is conducted. This includes Universe of the study, sample of the study,Data and Sources of Data, study's variables and analytical framework.

IV.RESULTS:

In recent years, research has proven that machine learning algorithms may be successfully deployed to detect suspicious transactions in massive amounts of online payments data, making detection of this type of fraud an important element of cyber-crime agencies' operations. In addition to catching fraudulent transactions in real time, these methods can catch ones that human auditors might miss. Using available to the public simulated online payment transaction data, a novel machine learning model for feature selection is proposed that considers the best features. Over time, victims of online transaction fraud have suffered enormous financial losses. As cutting-edge technologies and global communication have developed, so too has the prevalence of online fraud.

An essential step in mitigating these costs is the development of robust fraud detection algorithms. Fraud detection relies heavily on machine learning and statistical methods. Due to factors such as sparse data, the potentially sensitive nature of some data, and uneven class distributions, putting into practice a model to identify fraud can be difficult. Due to the sensitive nature of the data, it is difficult to make inferences and develop more accurate models. The proposed model is implemented in python and executed in Google Colab. The proposed model considers the dataset from the link https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset. In this research a Multi Level Associated Weighted Feature Vector (ML-AWFV) Model using machine learning is proposed for accurate feature extraction and selection for further processing of online payment fraud detection. The proposed model generates a feature set that is used for training the model for accurate online fraud detection. The proposed model is compared with the traditional Fine-Grained Co-Occurrences for Behavior-Based Fraud Detection in Online Payment Services (FGCBFD-OPS) [127]. The comparative analysis illustrates that the proposed model performance is high.

Data preprocessing is a crucial part of the mining process, and can involve anything from simple alteration to the complete removal of data in favor of more relevant or useful information. Preprocessing the data is done so that it is of higher quality and more suited for the mining activity at hand for accurate detection of online payment frauds. The Figure 4.2 shows the Pre-Processing Accuracy Levels of the proposed and traditional models.

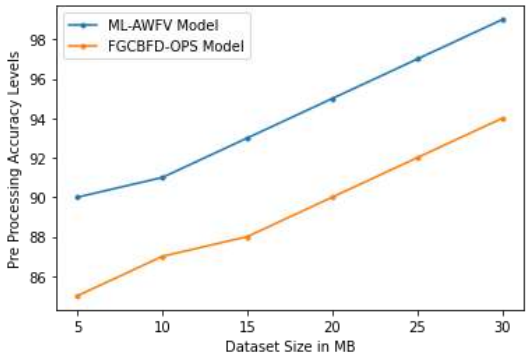

Fig 4.2: Pre-Processing Accuracy Levels

Data redundancy can be minimized with the aid of feature extraction. Data reduction aids model construction with less computational effort and boosts the pace of machine learning's learning and generalization phases. Using the feature extraction method, more features can be generated that are a combination of the preexisting features. The values of the new set of characteristics will be different from those of the old set. The Feature Extraction Time Levels of the proposed and existing models are shown in Figure 4.3

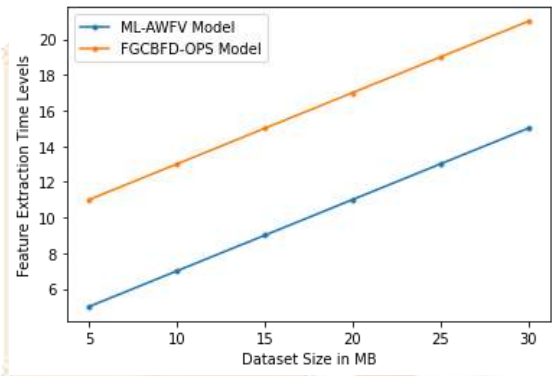

Fig 4.3: Feature Extraction Time Levels

Feature extraction is the procedure of reducing unstructured data to a set of quantifiable characteristics that can be further processed without losing any of the original data's meaning. Better outcomes are achieved than when machine learning is applied to the raw data itself. The Figure 4.4 shows the Feature Extraction Accuracy Levels of the proposed and existing models.

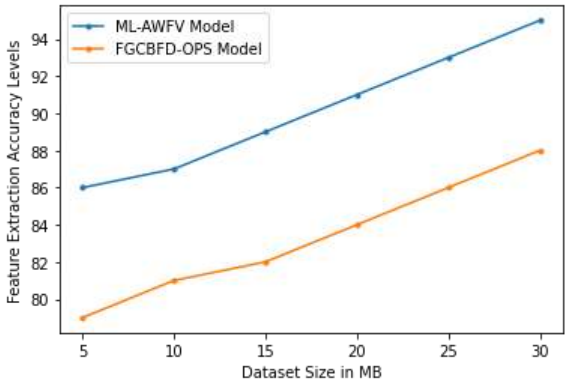

Fig 4.4: Feature Extraction Accuracy Levels

The linear link between two or more variables can be measured using the concept of correlation. One variable can be predicted from another through the use of correlation. The idea behind utilizing correlation to choose features is that superior variables will have a high degree of connection with the end result. Statistics measuring the degree of connection between the input and output variables are frequently used as a foundation for filter feature selection. The Correlation Calculation Accuracy Levels of the proposed and traditional models are shown in Figure 4.5.

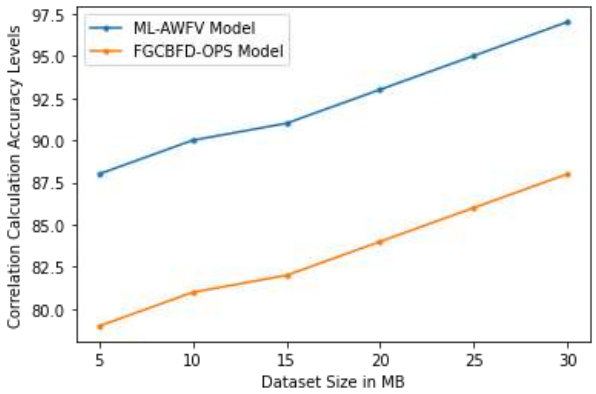

Fig 4.5: Correlation Calculation Accuracy Levels

The extracted features correlation is calculated and the weights are allocated for the features that have less correlation that considers the independent features. The independent features training will result in better accuracy rate. The Feature Weight Allocation Time Levels of the existing and proposed models are shown in Figure 4.6.

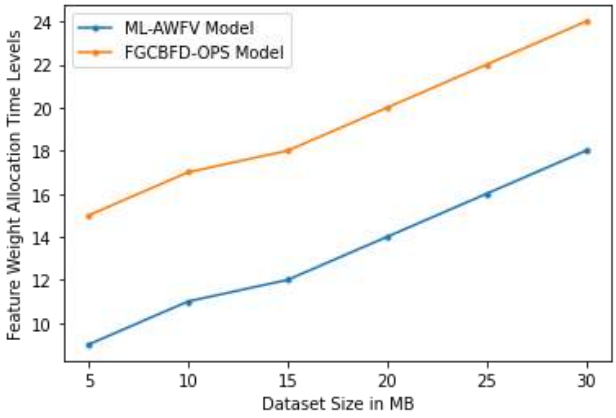

Fig 4.6: Feature Weight Allocation Time Levels

In most cases, feature weights are calculated by applying a learning algorithm to attribute a continuous relevant score to each feature based on the feature's significance in light of the particular context or domain expertise. Feature weighting is employed in an effort to quantify the significance of each feature by giving it a numerical value. When traits are correctly weighted, those with more significance are given more weight than those with less significance or no significance at all. The Feature Weight Allocation Accuracy Levels of the proposed and existing models are shown in Figure 4.7.

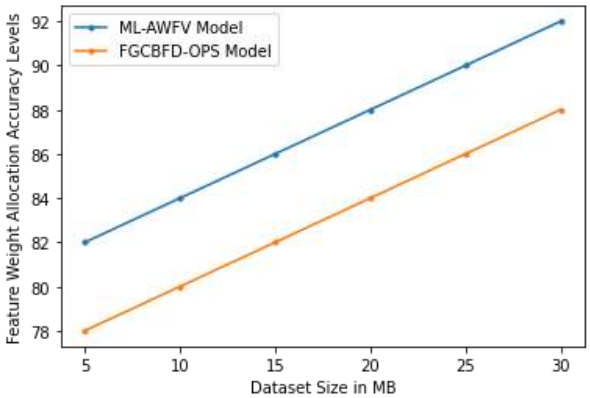

Fig 4.7: Feature Weight Allocation

V. Conclusion

Identifying fraudulent activity in popular payment systems is crucial now more than ever. This research proposed a solution that is a strong, rapid, and accurate method of selecting relevant features that detects frauds that occur in online payments. The solutions are able to reduce the likelihood of fraud occurring because they are built on high-performing Machine Learning algorithms, which produce accurate predictions quickly and at low cost. Financial services providers, banks, financial institutions, etc. may find this method applicable, as it can help them reduce the number of fraudulent transactions their customers

experience. In this research, we offer a method for identifying fraudulent financial transactions made using mobile devices. The field of fraud detection frequently encounters extremely unbalanced datasets. Based on the analysis of the selected dataset, it is demonstrated that the offered methods are effective in detecting fraudulent transactions with a low rate of false positives. In many cases, when trying to identify online fraudulent samples, it is necessary to sacrifice accuracy in favor of avoiding the misclassification of numerous legitimate samples. Every organization that facilitates digital payments must make this kind of design and operational decision often. In this research a Multi Level Associated Weighted Feature Vector Model using machine learning is proposed for accurate feature extraction and selection for further processing of online payment fraud detection. The proposed model achieves 97% accuracy in detecting relevant features to train the model for accurate online fraud prediction. In future, the feature dimensionality reduction strategies can be applied integrated with optimization models for considering the most weighted features for training models.

References:

[1].C. Wang and H. Zhu, "Representing Fine-Grained Co-Occurrences for Behavior-Based Fraud Detection in Online Payment Services," in IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 1, pp. 301-315, 1 Jan.-Feb. 2022, doi: 10.1109/TDSC.2020.2991872.

[2].F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," in IEEE Access, vol. 10, pp. 39700-39715, 2022, doi: 10.1109/ACCESS.2022.3166891.

[3].H. Wang, W. Wang, Y. Liu and B. Alidaee, "Integrating Machine Learning Algorithms WithQuantumAnnealingSolversforOnlineFraudDetection,"inIEEEAccess,vol.10, pp. 75908-75917, 2022, doi: 10.1109/ACCESS.2022.3190897.

[4].F. Zhu, C. Zhang, Z. Zheng and S. A. Otaibi, "Click Fraud Detection of Online Advertising–LSH Based Tensor Recovery Mechanism," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 7, pp. 9747-9754, July 2022, doi: 10.1109/TITS.2021.3107373.

[5].B. Can, A. G. Yavuz, E. M. Karsligil and M. A. Guvensan, "A Closer Look Into the Characteristics of Fraudulent Card Transactions," in IEEE Access, vol. 8, pp. 166095- 166109, 2020, doi: 10.1109/ACCESS.2020.3022315.

[6].E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba and G. Obaido, "A Neural Network EnsembleWith FeatureEngineering forImprovedCreditCard Fraud Detection," inIEEE Access, vol. 10, pp. 16400-16407, 2022, doi: 10.1109/ACCESS.2022.3148298.

[7].C. Wang, C. Wang, H. Zhu and J. Cui, "LAW: Learning Automatic Windows for Online Payment Fraud Detection," in IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 5, pp. 2122-2135, 1 Sept.-Oct. 2021, doi: 10.1109/TDSC.2020.3037784.

[8].A.A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection UsinganOptimizedLightGradientBoostingMachine,"inIEEEAccess,vol.8,pp.25579- 25587, 2020, doi: 10.1109/ACCESS.2020.2971354.

[9].Y. Zhang, J. Tong, Z. Wang and F. Gao, “Customer Transaction Fraud Detection Using Xgboost Model,” 2020 International Conference on Computer Engineering and Application (ICCEA), Guangzhou, China, 2020, pp. 554-558, doi: 10.1109/ICCEA50009.2020.00122.

[10].N. Malini and M. Pushpa, “Analysis on credit card fraud identification techniques based on KNN and outlier detection,” 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), Chennai, 2017, pp. 255-258, doi: 10.1109/AEEICB.2017.7972424.

[11]. A. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga and N. Kuruwitaarachchi, “Real-timeCreditCardFraudDetectionUsingMachineLearning,”20199thInternational Conference on Cloud Computing, Data Science Engineering (Confluence), Noida, India, 2019, pp. 488- 493, doi: 10.1109/CONFLUENCE.2019.8776942.

[12]. S. Khatri, A. Arora and A. P. Agrawal, “Supervised Machine Learning Algorithms for Credit Card Fraud Detection: A Comparison,” 2020 10th International Conference on Cloud Computing, Data Science Engineering (Confluence), Noida, India, 2020, pp. 680- 683, doi: 10.1109/Confluence47617.2020.9057851.

[13].C.Wang,Y.Wang,Z.Ye,L.Yan,W.CaiandS.Pan,“CreditCardFraudDetectionBased onWhaleAlgorithmOptimizedBPNeuralNetwork,”201813thInternationalConference on Computer Science Education (ICCSE), Colombo, 2018, pp. 1-4, doi: 10.1109/ICCSE.2018.8468855.

[14]. R. Rambola, P. Varshney and P. Vishwakarma, "Data Mining Techniques for Fraud DetectioninBankingSector,"20184thInternationalConferenceonComputing

Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-5, doi: 10.1109/CCAA.2018.8777535.

[15]. Massimiliano Zanin, Miguel Romance, Santiago Moral, ReginoCriado, "Credit Card FraudDetectionthroughParencliticNetworkAnalysis",Complexity,vol.2018,ArticleID 5764370, 9 pages, 2018. https://doi.org/10.1155/2018/5764370.

[16].Jain,Y.Tiwari,N.Dubey,S.Jain,Sarika.(2019).Acomparativeanalysisofvariouscredit card fraud detection techniques. International Journal of Recent Technology and Engineering. 7. 402-407.

[17]. I. Benchaji, S. Douzi and B. ElOuahidi, "Using Genetic Algorithm to Improve Classification of Imbalanced Datasets for Credit Card Fraud Detection," 2018 2nd Cyber Security in Networking Conference (CSNet), Paris, 2018, pp. 1-5, doi: 10.1109/CSNET.2018.8602972.

[18]. E. A. Lopez-Rojas , A. Elmir, and S. Axelsson. "PaySim: A financial mobile money simulator for fraud detection". In: The 28th European Modeling and Simulation Symposium-EMSS, Larnaca, Cyprus. 2016.

[19].DalPozzolo,AndreaAdaptiveMachinelearningforcreditcardfrauddetectionULBMLG PhD thesis (supervised by G. Bontempi).

[20].KumarA.etal.(2020)MalwareDetectionUsingMachineLearning.In:Villazon-Terrazas B., Ortiz-Rodr ´ ´ıguez F., Tiwari S.M., Shandilya S.K. (eds) Knowledge Graphs and Semantic Web. KGSWC 2020. Communications in Computer and Information Science, vol 1232. Springer, Cham. https://doi.org/10.1007/978-3-030-65384-2 5

[21].Nerurkar, P., Busnel, Y., Ludinard, R., Shah, K., Bhirud, S. and Patel, D., 2020, August. DetectingIllicitEntitiesinBitcoinusingSupervisedLearningofEnsembleDecisionTrees. InProceedingsofthe202010thInternationalConferenceonInformationCommunication and Management (pp. 25-30). DOI:https://doi.org/10.1145/3418981.3418984

[22].ThusharaAmarasinghe,AchalaAponso,andNaomiKrishnarajah.2018.CriticalAnalysis ofMachineLearningBasedApproachesfor FraudDetectionin FinancialTransactions.In Proceedings of the 2018 International Conference on Machine Learning Technologies (ICMLT '18). Association for Computing Machinery, New York, NY, USA, 12–17. DOI:https://doi.org/10.1145/3231884.3231894

[23].ImaneSadgali, NawalSael, and FaouziaBenabbou. 2019. Fraud detection in credit card transaction using neural networks. In Proceedings of the 4th International Conference on SmartCityApplications(SCA'19).AssociationforComputingMachinery,NewYork, NY, USA,Article95, 1–4. DOI:https://doi.org/10.1145/3368756.3369082

[24].IshanSohony,RameshwarPratap, andUllasNambiar.2018.Ensemblelearningfor credit card fraud detection. In Proceedings of the ACM India Joint International Conference on DataScienceandManagementofData(CoDS-COMAD'18).AssociationforComputing Machinery, New York, NY, USA, 289–294. DOI:https://doi.org/10.1145/3152494.3156815

[25].[17]YounessAbakarim,MohamedLahby,andAbdelbakiAttioui.2018.AnEfficientReal TimeModelForCreditCardFraudDetectionBasedOnDeepLearning.InProceedingsof the 12th International Conference on Intelligent Systems: Theories and Applications (SITA'18).AssociationforComputingMachinery, NewYork,NY,USA,Article30,1–7. DOI:https://doi.org/10.1145/3289402.3289530

[26].Qi Li and Yu Xie. 2019. A Behavior-cluster Based Imbalanced Classification Method for CreditCard FraudDetection.In Proceedings ofthe20192ndInternationalConferenceon Data Science and Information Technology (DSIT 2019). Association for Computing Machinery, New York, NY, USA, 134–139. DOI:https://doi.org/10.1145/3352411.3352433

[27].Bernardo Branco, Pedro Abreu, Ana Sofia Gomes, Mariana S. C. Almeida, Joao Tiago Ascens ˜ ao, and Pedro Bizarro. 2020. Interleaved ˜ Sequence RNNs for Fraud Detection. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery Data Mining. Association for Computing Machinery, New York, NY, USA, 3101–3109. DOI:https://doi.org/10.1145/3394486.3403361

[28].LigongChen, Zhaohui Zhang,Qiuwen Liu,LijunYang,YingMeng, andPengwei Wang. 2019. A method for online transaction fraud detection based on individual behavior. Proceedings of the ACM Turing

Celebration Conference - China. Association for Computing Machinery, New York, NY, USA, Article 119, 1–8. DOI:https://doi.org/10.1145/3321408.3326647