Enhanced Random Forest Algorithm for Faster and Accurate Fraud Detection

Mrs. G. Jaculine Priya[1]
Ph.D. Research Scholar,
Department of Computer Science, VISTAS, Pallavaram, Chennai, Tamil Nadu, India,
Dr. S. Saradha[2]
Assistant Professor,
MEASI Institute of Information Technology, Chennai, Tamil Nadu, India,

Abstract*: *Paper In this Digital Era security is the important concern. The more using digital communication and digital fraud also becomes more. When the technology grows in a faster pace, another side fraudsters also becoming strong and fast. Customers have to be very careful using their user's name, password, pin number, OTP etc. Without Customers knowledge the fraudsters are stealing their secured details. Industry also spending more time, effort and money to stop/reduce these illegal attempts keeping their client's security. But unfortunately, still there's a loop hole we are facing and fraudsters are well ahead in their job. AI&ML is the endowment in this current digital technological world. This niche technology can be used Finance, Medical, Insurance, Ecommerce almost across all the domains. Each company/industry has its own security approaches for providing a safe environment to their clients. Here, suggested solution aims to create a global centralized framework for sharing their fraud patterns by getting into a Digital handshake across the organization from various domains. Here AI&ML algorithms helps to detect and prevent fraud attacks intelligently. These algorithms help to find out all the hidden patterns. Choosing the right algorithm definitely it improves the detecting and preventing the fraud patterns and also gives secured environment to the customers.

Keywords*: *Artificial Intelligence, Machine Learning, Fraud Detection, Prevention, Credit Card

I. INTRODUCTION

Cyber Security is the major concern when an information is conveyed across a medium in our technological era, the applications we use keep on changing in a daily basis. The digital transition is greatly assisting all sectors. At the same time, we leave our digital footprints wherever and whenever we go. It makes users extremely cautious about their credentials. Users want us to be proactive in preventing fraud. In this paper, we will examine how to improve security features. However, scammers do not lag behind when it comes to embracing new technologies. Corporates must make sure that their security measures stay up with their operations with the innovated solutions.

A large number of fraudulent transactions occur in the banking sector on a daily basis. As an example, if we take only credit/debit card fraudulent transactions, the fraud will be notified to the appropriate bank management. As a temporary solution, the bank has blocked the specific customer's card from future transactions and advised the customers to reset his password and other sensitive details.

Only after that did the specific bank/corporate begin investigating how this fraudulent activity occurred, and what strategy the thieves employed to breach the client account. Once the pattern is detected, the banking/corporate sectors will endeavor to close loopholes and implement the required security measures to effectively stop the same pattern-related malicious activity. Meanwhile, criminals would have committed several fraudulent transactions following the same pattern, resulting in massive losses.

Fraudsters are constantly inventing new ways of accessing banks and their customers in the context of a rapidly changing global financial market, where demand for face-to-face banks is falling, volumes of online wallets are expanding, and transactions are made in seconds. Banks must be adaptable in responding to emerging dangers and embracing new strategies and technologies in order to foresee and prevent fraud.

Financial firms' greatest threat is cyber-related theft. To limit fraud risks in the future, financial institutions will need to make a massive change in their approach. Fundamentally, financial institutions must comprehend the rapid digital revolution occurring all around us, recognize the resulting evolving fraud threats, and create fraud risk management practices capable of mitigating these fraud risks in a consistent, efficient, and intelligent manner. Existing financial institution solutions, while expensive to operate, are not capable of coping with rising fraud concerns because they are too scattered and unsophisticated. Computer fraud risk management in the future should be able to adapt to the rapidly evolving digital transformation, discover previously unidentified fraud threats, take advantage of technology, and reduce compliance costs. Fixed deposits, loan disbursement or granting credit facilities for bribery, hacking and other web ATM-based scams are some of the latest fraud instances in India covered by the media. These high-profile incidents in recent years have demonstrated that scams may harm an organization's brand as well as its earnings, operating efficiencies, and customer satisfaction. It can have a severe influence on staff morale and confidence of investors, in addition to potential regulatory sanctions. Although no business can completely eliminate the risk of fraud, it is critical to have procedures in place that can prevent and detect fraud.

The following is the structure of this paper:

The Solving the Global Issue and Data Preprocessing are discussed in section 2. Various Machine Learning Algorithms and Comparison of Algorithms are discussed in section 3. Proposed Enhanced ML Algorithm steps are detailed in section 4. The Results and Discussions are stated in Section 5.

II. SOLVING THE GLOBAL ISSUE

After fraudulent activity reported, fraudulent transactions are identified by the employer of the organization. This information/pattern can be saved in a centralized database. Based on this experience security of the organization can be improved. Each company has its own security approach for providing a safe environment for its clients. In my previous papers suggested a solution aims to create a centralized global framework for sharing fraud patterns, allowing any business to retain and see fresh fraudulent transactions carried out by criminals.

Geotagging, IP Address, Date, Time, as well as other fraud-related information may all be maintained in a record, just like any other fraud transaction. We might reach a digital agreement between organizations across the domains. These organizations can share the knowledge in a standard format by the industry who signed the agreement. Now that all fraud transactions are visible to everyone, the surviving organizations may take preventative security measures to avoid large losses. A fraudster's relationship with a victim may be compared to a cat-and-mouse fight, in which each side is constantly learning and adapting, employing inventiveness and understanding of the other's objectives to develop new fraud prevent and detect techniques. In this paper discussed how to get to fast and accurate fraud detection solution by using the proposed enhanced machine learning algorithms in real time.

A. Importance of AI-ML Algorithms in Solving the Global Issue

Machine Learning is a subset of AI. It has 3 types of algorithms such as supervised, unsupervised and reinforcement learning algorithms. The Supervised Algorithms address regression and classification/anomaly findings problems. The unsupervised algorithms help to cluster unlabeled components. Reinforcement Learning helps the industry to address action-reward-punishment environment-based issues. It helps the model to learn and correct the error on its own. The Credit Card Fraud Detection system uses supervised algorithms such as Logistic Regression, Naïve Bayes and Support Vector Machine (SVM) to identify the fraud transactions. The Anomaly detection is a kind of classification model, where the sensitivity is key to find the anomaly. In general, if the model threshold is greater than 0.5, then it will classify it as +positive case but here the sensitivity/recall/true positive rate is very important to classify +ve/-ve transactions.

Machine Learning Techniques in Ai Technology now give intelligent answers to the majority of issues of human and traditional approaches in FP & FD. In the past, industries relied on their employees and their traditional methodology to FP & FD. The government, as well as all the private sector organizations, is investing more money on fraud prevention. The algorithms are created by fraud domain specialists in the traditional method. The algorithms and procedures are strictly based on rules. Traditional methods are no longer sufficient to address this issue. This is the current state, and it's clear existing FP & FD operate in a soiled approach. Its more over reactive rather than proactive mode. Recovery time of fraud issues are more and not real time. We don't learn from our history. We don't learn from our/another business verticals. Industry has to work well ahead of fraudsters. Because of the popularity and precision of Artificial Intelligence, every sector is currently transitioning from the traditional approach to ML-based solutions for FP & FD.

Fraud detection and prevention algorithms surely find out the hidden and new patterns based on the experience. It is a continuous learning procedure. It needs full life cycle which includes monitoring, learning, identifying, preventing, and decision-making in real time. Fraudulent transactions may be prevented and detected using the correct Supervised and Unsupervised learning algorithms. The benefit is that we may use either of the models individually or in combination to find abnormalities.

B. Proposed Global Fraud Prevention High Level Architecture

Transaction information is published and stored in a centralized database. The Global Fraud Prevention High-Level Architecture is depicted in the above Figure. The source organization with a NO or YES for the "fraud help indication." A "NO" would indicate that the publishing company has previously detected a transaction as fraudulent and shared this information in the centralized database for information purposes. This will help other companies involved in this experiment avoid similar suspicious transactions using the same patterns in the future. A "YES" would indicate that the publishing organization's is looking for an indicator from the Global model to allow or prevent the transaction from being completed in real-time.

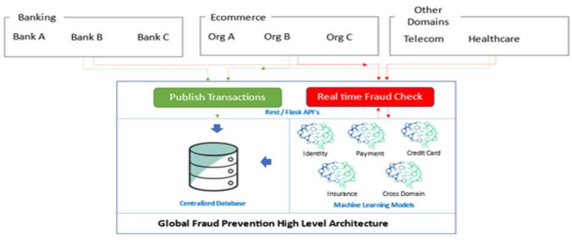

Fig. 1.Proposed Global Fraud Prevention High Level Architecture- Level I

Similarly, remaining participating organizations in this digital handshake can communicate information across sectors. It aids in the prevention and detection of fraudulent transactions worldwide. The accuracy of a model is determined by the number of data sets utilized for training and testing. The model, once trained, can detect and prevent fraud patterns. Based on the fraud indicator signal supplied by the model, parent apps hosted by the company where the fraud originated can prevent the transaction from successfully completing or execute extra security validations before allowing the transaction to pass through. Using the suggested global architecture model, newly discovered fraud tendencies may then be shared across enterprises, allowing for proactive fraud prevention. Thereby its extremely important for organizations to join hands together in proactively preventing fraud Globally.

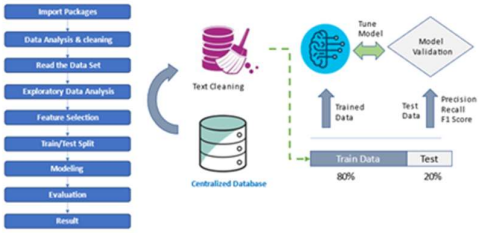

Fig. 2.Proposed Global Fraud Prevention High Level Architecture- Level II

C. Step wise execution for implementing a model

1. Load Dataset into the system.
2. Class Count function to check if the dataset is balanced or imbalanced
3. Scale Non- Anonymised features such as Time and Amount
4. Concatenate both Scaled Amount and Time with Actual Dataset/frame.
5. Drop old Amount and Time features
6. Random under-sampling function
6.1 Split the scaled dataset into Train and Test
6.2.Reset index of scaled train and test dataset
6.3.Find no of fraud transactions in the (random) Training dataset
6.4.Segregate normal and fraud transactions from the training data
6.5.Randomly selecting the same no of normal transactions as fraud transactions
6.6.Reset Index of both selected normal and fraud transactions
6.7.Concatenate both selected normal and fraud transactions.
6.8.Shuffle the subsample/final data frame.
7.Remove extreme outliners and create final dataset
8.Split final dataset into Train and Test
9.Spot-check a couple of Classification Algorithms (using cross validation)
10. Make Predictions on Test data
11. End

D. Data Preprocessing

Steps in Data Pre-processing in Machine Learning

1. Obtain the dataset
2. Add all necessary libraries
3. Load the dataset
4. Recognizing and dealing with missing values
5. Categorical data encoding
6. Dividing the dataset
7. Scaling of features

Because of their varied origin, the bulk of real-world datasets for machine learning are very sensitive to absent, irregular, and incomplete data. As a result, data preprocessing is critical for improving overall data quality. Duplicate or missing data may provide an inaccurate picture of the overall statistics of the data. Deviations and inconsistent data can disrupt the model's overall learning, resulting in incorrect predictions. Quality data must be used to make quality judgments. To obtain this high-quality data, data pre-processing is required.

Implementation of Dataset

Data set represents real-world data. This data set contains all credit card transactions. It's an imbalanced data set.

- It has 30 features and 1 target. Of 30 features, 28 features are labelled V1 to V28.
- The remaining 2 features are Time and Amount.
- The 28 features are in the form of PCA complaint.
- Both Amount and Time are not normalised, that is, not in line with other variables in terms of scale.

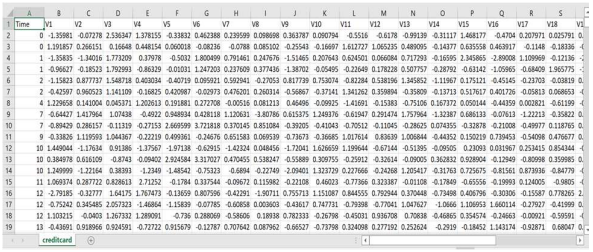

Fig. 3. Data Preprocessing Dataset Representation

Identifying Fraud Transactions

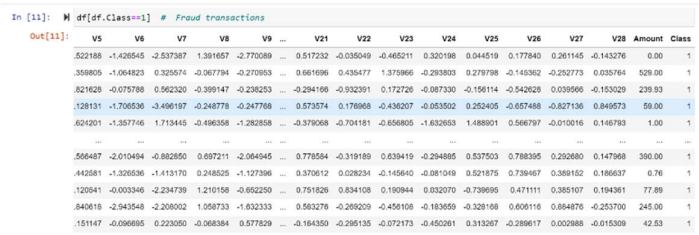

Fig. 4 . Data Preprocessing Fraud Transactions Class 1

Identifying Good Transactions

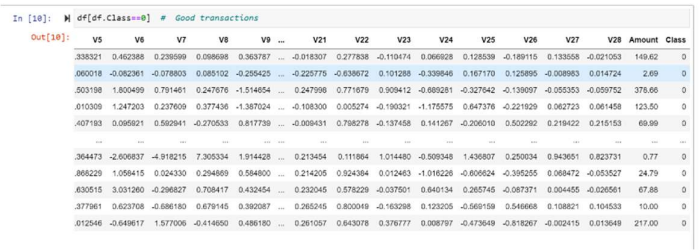

Fig. 5 . Data Preprocessing Fraud Transactions Class 0

Fraud transactions are very less comparing to normal transactions. Feature class is the respective variable it takes 1 in case of fraud transaction and 0 in case of good transaction. We have to handle this imbalanced data set first.

Removing Duplicate Records

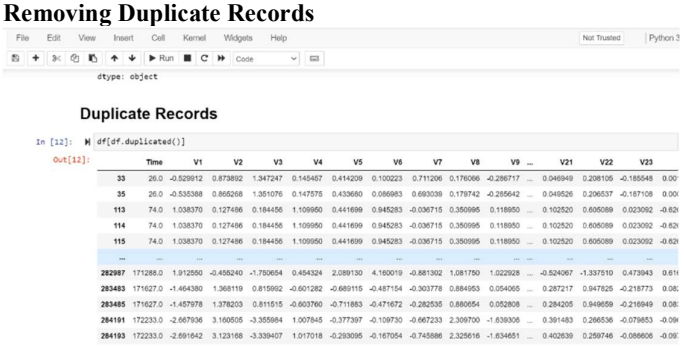

Fig. 6 . Data Preprocessing Removal of Duplicate Record

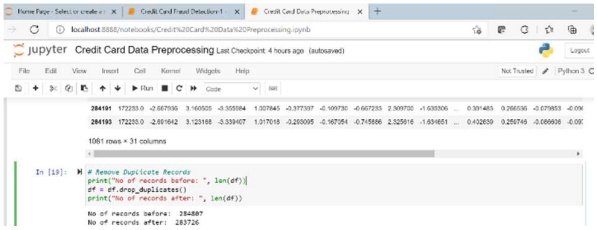

Fig. 7 . Data Preprocessing After Removal of Duplicate Records

In data preprocessing we will often need to find out the duplicate data rows and how to handle them. Finding the duplicate data , counting the duplicate and non-duplicate datas, extracting the duplicted rows with correct location , determining to keep the remaining rows and drop the duplicated rows. Similarly we can consider and do the columns as well by dropping duplicates.

Describing Dataset Descriptive Statistics

Dataset description helps to understand how datas have been collected in a defined structure. Each row and column describes a particular variable like mean, count, standard deviation, percentiles and minimum-maximum value ranges are included. We can get a descriptive statistics summary of given data frame.

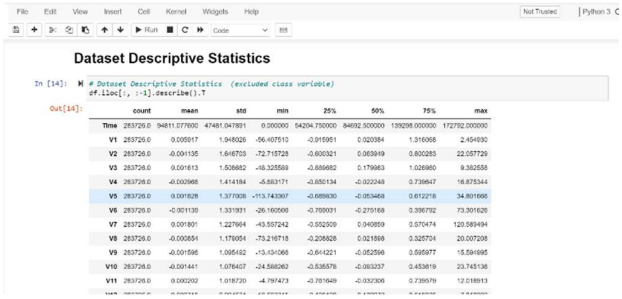

Fig. 8 . Data Preprocessing- Dataset Descriptive Statistics

Heat Map

Heat map can infer features are not correlated mostly that means the variation of one feature minutely affects the other feature that either has a positive or a negative correlation with each other.

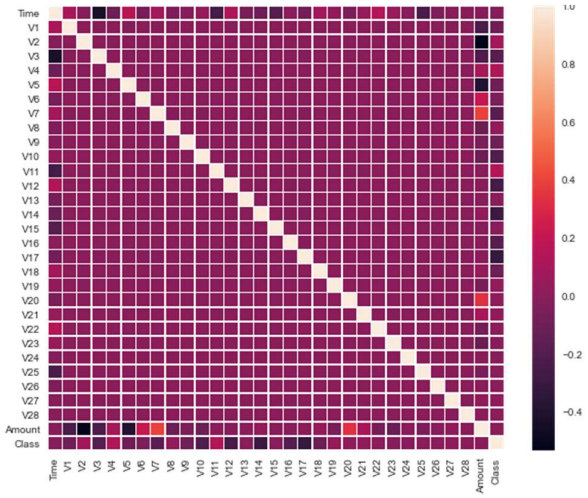

Fig. 9 . Data Preprocessing – Heat Map of the Dataset

III. MACHINE LEARNING ALGORITHMS

A. Support Vector Machine Algorithm:

The SVM method is best understood by concentrating on its fundamental kind, the SVM classifier. The SVM classifier is designed to create a hyper-lane in an N-dimensional environment that splits pieces of data into multiple groups. This hyperplane, however, is chosen based on margin, as the hyperplane with the greatest margin between the two classes is evaluated. These margins are determined using Support Vectors, which are data points. Support Vectors are data points that are close to the hyperplane and aid in its orientation.

If the operation of an SVM classifier needs to be understood theoretically, it may be divided into the following steps -

Step 1: The SVM algorithm predicts the classes. One of the classes is labelled as 1, while the other is labelled as -1. Load the necessary libraries.

Step 2: Import the dataset and extract the X and Y variables independently.

Step 3: Separate the dataset into train and test subsets.

Step 4: Set up the SVM classifier model

Step 5: SVM classifier model fitting

Step 6: Making predictions

Step 7: Assessing the model's performance

When there is no mistake in the classification, the gradients are just updated using the regularization parameter, whereas the loss function is additionally employed when misclassification occurs.

B. KNN Algorithm:

K Nearest Neighbors is a simple method that uses a similarity metric to predict the categorization of unlabeled data. We calculate the distance between the points when two parameters are shown in a 2D Cartesian system to get an idea of how similar they are. The KNN algorithm is based on the idea that similar objects are close to one another.

The K Nearest Neighbors method is one of the most basic and straightforward classification techniques. Based on how it operates, it is also classified as a "Lazy Learning Algorithm." The K-value that everyone achieves while training the model is usually an odd integer, however this is not required. However, there are a few drawbacks to employing KNN. A few of them are as follows:

- It is incompatible with categorical data because we cannot determine the distance between two category characteristics.
- It also does not perform well with high-dimensional data since the algorithm will struggle to calculate the distance in each dimension.

The KNN Algorithm in Python: A Step-by-Step Guide
Step 1: Import Libraries. Importing the libraries required to execute KNN is demonstrated below.
Step 2: Import the Dataset We can see the dataset being imported here....
Step 3: Split the dataset...
Step 4: Design a Training Model.
Step 5: Make Running Predictions.
Step 6: Validation Check.

C. Random Forest Algorithm:

Random Forest operates in two stages: First Step: Generating the random forest by combining N decision trees. Second Step: Making predictions for each tree generated in the first step. Random Forest is a classifier that uses a several decision trees on various subsets of a given dataset and averages them to enhance the predicted accuracy of that dataset. "Rather than depending on a single decision tree, the random forest calculates the forecast from each tree and choosing the final output result based on the majority vote of predictions." The greater number of trees in the forest, the stronger the precision and smaller the chance of errors.
The steps to illustrate the working process:
Step 1: Select M data points at random from the training set.
Step 2: Generate decision trees for the specified data points (Subsets).
Step 3: Construct the N number of decision trees
Step 4: Reverse steps 1 and 2.
Step 5: Find each decision tree predictions for new data points and allocate the new data points to the category that has received the most votes.
Random Forest Algorithm Implementation
1. Pre-Processing of Data
2. RFA fitting to the training set
3. Estimating the Test Set outcome
4. Creation of the Confusion Matrix
5. Visualizing the Training Set's Outcome
6. Projecting the results of the test set

D. Comparison of Machine Learning Algorithms

There are two distinctions in the efficiency of random forest and gradient boosting. The random forest can create each tree individually, but gradient boosting can only build one tree at a time, hence the random forest performs worse than gradient boosting. Random Forest is designed for multi-class issues, whereas SVM is designed for two-class problems. Multiclass problem must be divided into numerous binary classification tasks. Random Forest performs effectively when characteristics are both categorical and numerical. We all know that KNN is a lazy learner; it memorizes the information, resulting in a 0-training time. Due to the fact that it does not train for parameters or weights.
While predicting, it really does all of the work. It has a complexity on the order of n * m *d, where n is the volume of the training data, m is the amount of the test data, and d is the number of operations that can be performed for every test. So, after approximately 10 minutes, it stops making predictions. As seen above, Decision Tree completed instantaneously with 85 percent accuracy, KNN with 92 percent accuracy but significant running time and consuming resources all along and Random Forest with 93 percent accuracy and very little running time.

Table 1.Comparsion of various ML Algorithms for Data Set-1

SNO	ALGORITHM	CLASS	PRECISION	RECALL	ACCURACY	F1 SCORE
1	SVM	0	1.00	1.00	1.00	0.85
		1	0.99	0.75		
2	Logistic Reg	0	1.00	1.00	1.00	0.90
		1	0.94	0.88		
3	KNN	0	1.00	1.00	1.00	0.92
		1	0.99	0.88		
4	Random	0	1.00	1.00	1.00	0.89
		1	0.94	0.77		

Table 2 .Comparsion of various ML Algorithms for Data Set-2

SNO	ALGORITHM	CLASS	PRECISION	RECALL	ACCURACY	F1 SCORE
1	SVM	0	1.00	1.00	1.00	0.85
		1	0.99	0.75		
2	Logistic Reg	0	1.00	1.00	1.00	0.90
		1	0.96	0.63		
3	KNN	0	1.00	1.00	1.00	0.92
		1	0.94	0.88		
4	Random	0	1.00	1.00	1.00	0.89
		1	0.99	0.88		
1	SVM	0	1.00	1.00	1.00	0.85
		1	0.99	0.75		
2	Logistic Reg	0	1.00	1.00	1.00	0.90
		1	0.94	0.88		
3	KNN	0	1.00	1.00	1.00	0.92
		1	0.99	0.88		
4	Random	0	1.00	1.00	1.00	0.89
		1	0.94	0.77		

E. Confusion Matrix

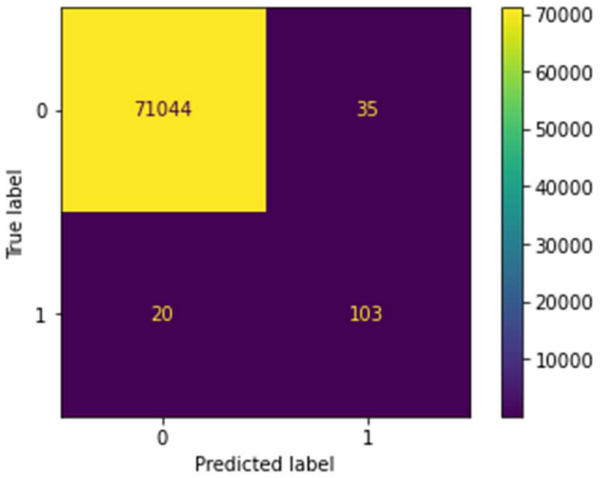

Fig. 10 .Confusion Matrix

Precision:
Precision is one of the indicators of ML Algorithms model's performance i.e., the Quality of the positive prediction made by model. The number of true positives divided by the total number of positive predictions.
Recall:
The number of true positives divided by the number of positive values in the test data. Low Recall indicates a high number of false negatives.
F1-Score:
The weighted average of precision and recall.
Confusion Matrix:
A table showing correct predictions and types of incorrect predictions
.

	Prediction	
Actual	0(Negative)	1(Positive)
0(-ve)	TN (True Negative)	FP (False Positive)
1(+ve)	FN (False Negative)	TP (True Positive)

$$Accuracy = \frac{TN+TP}{TN+TP+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{FP+TP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = 2* \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

Precision: 0.746	Recall: 0.837	Accuracy: 0.999	F1 Score: 0.789

Table 3 . Performance Results

IV. PROPOSED ENHANCED RANDOM FOREST ALGORITHM

Above steps mentioned in II-D more over standard way of how machine learning algorithms works. Since proposing global model solution, it is important that algorithms used here are efficient and speed in real time. Need of the hour is enhanced machine learning algorithm for best performance. Performance has been improved in Random Forest Algorithm by selecting the wise trees alone to predict accurately in terms of identifying fraud attacks.

Revised Steps wise execution for Proposed Model

1. Load Dataset into the system.
2. Class Count function to check if the dataset is balanced or imbalanced
3. Scale Non- Anonymised features such as Time and Amount
4. Concatenate both Scaled Amount and Time with Actual Dataset/frame.
5. Drop old Amount and Time features
6. Random under-sampling function
6.1 Split the scaled dataset into Train and Test
6.2.Reset index of scaled train and test dataset
6.3.Find no of fraud transactions in the (random) Training dataset
6.4.Segregate normal and fraud transactions from the training data
6.5.Randomly selecting the same no of normal transactions as fraud transactions
6.6.Reset Index of both selected normal and fraud transactions
6.7.Concatenate both selected normal and fraud transactions.
6.8.Shuffle the subsample/final data frame.
7.Remove extreme outliners and create final dataset
8.Split final dataset into Train and Test
9.Spot-check a couple of Classification Algorithms (using cross validation)
10. Make Predictions on Test data
11. Tune the Global model with Enhanced Random Forest Algorithm
12.End

V. RESULTS AND DISCUSSION

Early detection, achieved by improving algorithms to detect both developing risks and the actions of fraudster, may be a critical step toward minimizing and reducing losses. On an enterprise-wide scale, incident detection that integrates complicated, adaptive, signaling, and reporting systems may automate the correlation and analysis of enormous volumes of data across the sectors, as well as numerous danger indicators. Monitoring systems in organizations should be operational 24/7, seven days a week, with enough support for fast incident response and remediation processes. A thorough awareness of recognized risks and controls, as well as industry norms and laws, may help financial services businesses protect their systems through the design and implementation of proactive, risk-informed controls. Based on best practices, banks can implement a "defense-in-depth" strategy to combat known and developing threats. This entails sharing common security layers, both to offer stability and to potentially impede, if not prevent, the advancement of ongoing threats.

To summarize, firms cannot manage fraudulent transactions in silos any more. Operating in silos will cost enterprises throughout the world trillions of dollars in lost revenue. Increased preventative measures must be considered and implemented, and the idea of centrally keeping and managing fraud data would meet this demand. Simply allowing organizational systems to communicate and learn from one other fast will dramatically lessen the impact of fraud, therefore actively limiting fraudsters from attacking them.

This proposed framework accomplishes this by allowing organizations to safely access the central database via exposed API (Application Program Interfaces) for every payment in their system, inspecting for fraud while also sharing any new transactions to centralized repository for other organizations to benefit in real time. In this regard, three distinct models – Support Vector Machine, KNN, and Random Forest have been tested by using historical data thus far. Random Forest appears to be producing good results, with the highest success rate among the algorithms tested. With the correct technology in place, firms can work together to avoid fraud by sharing their fraud experiences and employing the latest and best strategies.

REFERENCES

1. G. Jaculine Priya and Dr.S.Saradha, *"Fraud Detection and Prevention Using Machine Learning Algorithms: A Review," 7th International*

Conference on Electrical Energy Systems (ICEES), 2021, pp. 364-368, DOI: 10.1109/ICEES51510.2021.9383631.

2. G. Jaculine Priya and Dr.S.Saradha., *"Real Time Global Fraud Detection and Prevention"., International E-Conference on Advances in Information Technology., June-2020 BIHER., ISBN NO.978-93-5407-796-8.*
3. E. M. S. W. Balagolla, W. P. C. Fernando, R. M. N. S. Rathnayake, M. J. M. R. P. Wijesekera, A. N. Senarathne and K. Y. Abeywardhana, *"Credit Card Fraud Prevention Using Blockchain,"* 2021 6th International Conference for Convergence in Technology (I2CT), *2021, pp. 1-8, doi: 10.1109/I2CT51068.2021.9418192.*
4. B. A. Smadi, A. A. S. AlQahtani and H. Alamleh, *"Secure and Fraud Proof Online Payment System for Credit Cards,"* 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), *2021, pp. 0264-0268, doi: 10.1109/UEMCON53757.2021.9666549.*
5. N. Chumuang, S. Hiranchan, M. Ketcham, W. Yimyam, P. Pramkeaw and S. Tangwannawit, *"Developed Credit Card Fraud Detection Alert Systems via Notification of LINE Application,"* 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), *2020, pp. 1-6, doi: 10.1109/iSAI-NLP51646.2020.9376829.*
6. A. M. Zinjurde and V. B. Kamble, *"Credit Card Fraud Detection and Prevention by Face Recognition,"* 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC), *2020, pp. 86-90, doi: 10.1109/ICSIDEMPC49020.2020.9299587.*
7. H. Dar, A. Abbasi and A. Naveed, *"Credit Card Fraud Prevention Planning using Fuzzy Cognitive Maps and Simulation,"* 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), *2020, pp. 289-294, doi: 10.1109/ICRITO48877.2020.9198002.*
8. N. Boutaher, A. Elomri, N. Abghour, K. Moussaid and M. Rida, *"A Review of Credit Card Fraud Detection Using Machine Learning Techniques,"* 2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech), *2020, pp. 1-5, doi: 10.1109/CloudTech49835.2020.9365916.*
9. V. Shah, P. Shah, H. Shetty and K. Mistry, *"Review of Credit Card Fraud Detection Techniques,"* 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), *2019, pp. 1-7, doi: 10.1109/ICSCAN.2019.8878853.*
10. Mehak Mahajan and Sandeep Sharma., *"Detect Fraud in Credit Card using Data Mining Techniques". International Journal of Innovative and Exploring Engineering, 9, pp.2278-3075. December 2019.*
11. Mandeep Singh and Sunny Kumar and Tushant Garg., *"Credit Card Fraud Detection Using Hidden Markov Model". International Journal of Engineering and Computer Science, 8, pp.24878-24882. November 2019.*
12. Shiv Shankar Singh., *"Electronic Credit Card Fraud Detection System by Collaboration of Machine Learning Models". International Journal of Innovative and Exploring Engineering,8(12S), pp.92-95. October 2019.*
13. Maniraj, S.P., Aditya Saini., Swarna Deep Sarkar and Shadap Ahmed., *"Credit Card Fraud Detection Using Machine Learning and Data Science". International Journal of Engineering Research & Technology,8(9), pp.110-115. September 2019.*
14. Yashvi Jain., Namrata Tiwari., Shripriya Dupey., and Sarika Jain., *"A Comparative Analysis of Various Credit Card Fraud Detection Techniques". International Journal of Recent Technology and Engineering, 7(5S2), pp.402-407. January 2019.*
15. Olawale Adepoju., Julius Wosowei., Shiwani lawte and Hemaint Jaiman., *"Comparative Evaluation of Credit Card Fraud Detection Using Machine Learning Techniques"., Global Conference for Advancement in Technology,IEEE, 2019,978-1-7281-3694-3*
16. Thennakoon, Anuruddha, *et al. "Real-time credit card fraud detection using machine learning." 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2019.*
17. Jhangiani, Resham, Doina Bein, and Abhishek Verma. *"Machine learning pipeline for fraud detection and prevention in e-commerce transactions." 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2019.*
18. Debachudamani Prusti., and Santanu Kumar Rath., *"Fraudulent Transaction Detection in Credit Card by Applying Ensemble Machine Learning techniques".,10th ICCCNT July 2019., IEEE*
19. Anish Halimaa A and Dr. K.Sundarakantham *"Machine Learning Based Intrusion Detection System".,ISBN:978-1-5386-9439-8.,IEEE 2019.*
20. Masoumeh Zareapoor and Pourya Shamsolmoali., *"Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier". Procedia Computer Science,48, pp.679-685. 2015*
21. Hunt,W.(2020).*Artificial Intelligence's Role in Finance and How Financial Companies are Leveraging the Technology to Their Advantage. Available at SSRN 3707908.*
22. Brynjolfsson, E., & Mcafee, A.N.D.R.E.W. (2017). *The Business of artificial intelligence. Harvard Business Review,7,3-11.*
23. Kumarapandian, S. (2018). *Melanoma classification using multiwavelet transform and support vector machine. International Journal of MC Square Scientific Research, 10(3),01-07.*
24. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). *Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.*
25. Bao, Y., Hilary, G., & Ke, B. (2020). *Artificial intelligence and fraud detection. Available at SSRN 3738618*
26. V. Mareeswari and G. Gunasekaran, *"Prevention of credit card fraud detection based on HSVM,"* 2016 International Conference on Information Communication and Embedded Systems (ICICES), 2016, pp. 1-4, doi: 10.1109/ICICES.2016.7518889.
27. K. T. Hafiz, S. Aghili and P. Zavarsky, *"The use of predictive analytics technology to detect credit card fraud in Canada,"* 2016 11th Iberian Conference on Information Systems and Technologies (CISTI), *2016, pp. 1-6, doi: 10.1109/CISTI.2016.7521522.*
28. F. Ghobadi and M. Rohani, *"Cost sensitive modeling of credit card fraud using neural network strategy,"* 2016 2nd International Conference of Signal Processing and Intelligent Systems (ICSPIS), *2016, pp. 1-5, doi: 10.1109/ICSPIS.2016.7869880.*
29. E. Caldeira, G. Brandao and A. C. M. Pereira, *"Fraud Analysis and Prevention in e-Commerce Transactions,"* 2014 9th Latin American Web Congress, *2014, pp. 42-49, doi: 10.1109/LAWeb.2014.23.*
30. Seeja, K.R. and Masoumeh Zareapoor., *"A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining". The Scientific World Journal, Volume 2014, Article ID 252797,10 pages. September 2014*
31. M. D. H. Mahdi, K. M. Rezaul and M. A. Rahman, *"Credit Fraud Detection in the Banking Sector in UK: A Focus on E-Business,"* 2010 Fourth International Conference on Digital Society, *2010, pp. 232-237, doi: 10.1109/ICDS.2010.45*

AUTHORS PROFILE

Mrs. G. Jaculine Priya is currently pursuing PhD in Computer Science at VISTAS, Pallavaram. Her Area of interest and research includes Artificial Intelligence, Machine Learning, Computer Security and Fraud. She has been actively taken part and presented and published various papers in International and National Conferences in his research area.

Dr.S. Saradha, working as Assistant Professor, MEASI Institute of Information Technology Chennai. She has more than 10 years of experience in Educational Institute. Her area of research includes Data Mining, Artificial Neural Networks, Machine Learning, Big Data Analytics. She has published more than 17 Research Papers in National and International journals. She is interested in writing textbooks for the students to make them understand any concept in an easy way. She is a recognised supervisor, guiding M.Phil. and PhD scholars.