Real Time Credit Card Fraud Detection using Fuzzy and Deep Neural Network

N. Prabha
Research Scholar, School of Computer Science, TheivanaiAmmal College for Women, Thiruvalluvar University, Serkade, Vellore, India

S. Manimekalai
Department of Computer Science, TheivanaiAmmal College for Women, Villupuram, Tamil Nadu, Inida

ABSTRACT

The financial fraud prediction is important research is needy one for current research. Because every minute's lakhs of fraud activities are happening in the worldwide. The different approaches are used to predict and analysis the credit card fraud activities. In credit card fraud prediction behaviour analysis is an important one. In this work, presented new prediction model based on the real time behaviour analysis and with the help of history transaction dataset. The behaviour analysis is performed based on the proactive data such as user typing speed of username and password, normal, abnormal activities and other properties. The proposed model prediction is performed based on the fuzzy logic and deep neural network system. The fuzzy logic used to check the behaviour of human and find the member activities. The deep neural network is used to identifying anomalous behaviour of credit card activities. For implementation, a real-time dataset from a commercial bank used static and dynamic features are used. The accuracy, sensitivity, and specificity metrics are used to measure the performance of proposed work.

Keywords

Credit card fraud Detection - fuzzy logic – deep learning - real time prediction.

1. INTRODUCTION

The financial fraud prediction is important research is needy one for current research. Because each any minute's lakhs of fraud activities are happening in the worldwide. The different approaches are used to predict and analysis the credit card fraud activities. In credit card fraud prediction behaviour analysis is an important one. In this work, presented new prediction model based on the real time behaviour analysis and with the help of history transaction dataset. The behaviour analysis is performed based on the proactive data such as user typing speed of username and password, normal, abnormal activities and other properties. The proposed model prediction is performed based on the fuzzy logic and deep neural network system. The fuzzy logic used to check the behaviour of human and find the member activities. The deep neural network is used to identifying anomalous behaviour of credit card activities. For implementation, a real-time dataset from a commercial bank used static and dynamic features are used. The accuracy, sensitivity, and specificity metrics are used to measure the performance of proposed work.

The term "fraud" can be interpreted as dishonesty for personal life and advantage. As the Internet has taken over the lives, many individuals and organizations have been the subject of fraudulent activity. According to multiple reports, the number of commercial fraud attempts increased in 2018 compared to 2016. In these years, frauds have surpassed each other by an astounding 83%. According to the E-commerce Fraud Index, the retail fraud rate increased from 0.06% in 2016 to 0.23% in 2017. 10% of all scams involve credit cards entirely, resulting in enormous financial losses that alarm businesses. Since the majority of transactions are now digital, the number of active cards has increased, and their transaction data has multiplied faster than before. Consequently, the volume of data to be examined throughout the detection process has grown substantially. Researchers mostly employ ML Algorithms, Neural Networking models, and Classification and Clustering approaches. Numerous studies are now focusing on the early detection or prevention of credit card fraud. Other researchers have also investigated viable and effective Fraud detection tools and techniques. ML and other linked methodologies, such as ANN (Artificial Neural Network), the method of rule induction, Logistic Regression (LR), Decision Trees (DT), and Support Vector Machine, are typically employed. ML algorithms are AI techniques that can tackle a variety of problems from a variety of disciplines and fields that typically contain vast volumes of data. Although numerous concepts and methods have been offered to prevent and detect fraud, there is still a significant need to apply and evaluate the efficacy of machine learning (ML) algorithms. Clusters are formed by dividing a huge group of data into smaller, related groups based on the similarities in their nature. Items in distinct clusters may not share the characteristics of other cluster pieces, as depicted in Fig.1. When certain values from a massive quantity of data are provided as input, these algorithms find a common interest or conclusion based on it. This means that these methods attempt to extract one or more outputs from the given input. ML algorithms are beneficial when solving certain tasks.

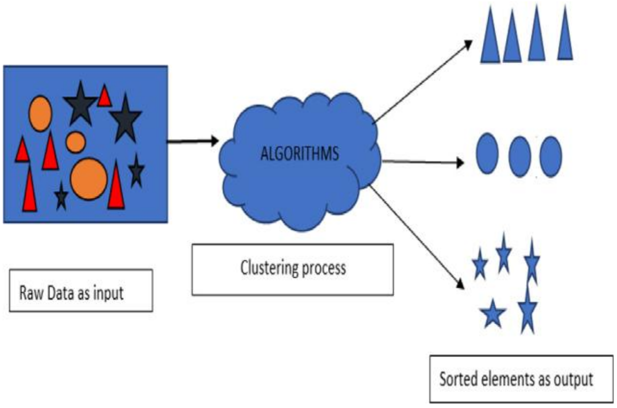

Fig 1 Data points and cluster of predictions

supervised learning strategy that determines the output directly based on the user-provided input data. This is typically employed for continuous data that cannot have discrete values. In such forecasts, ML algorithms are used. The purpose of this study is to evaluate the performance, precision, and efficacy of various Machine Learning classifiers when used to the prediction analysis and preventative analysis of particular algorithms. Key elements are additive approaches such as oversampling and binary classification, etc. It has been shown that when partnered with these strategies, typical machine learning classifier algorithms produce better results and boost the efficiency of the process.

2. RESEARCH ANALYSIS

The research analysis based on the two parts such as supporting methods and supporting previous methods. The real time fraud detection two methods used such ad deep neural network and fuzzy logics.

Deep Neural Network

A Deep Neural Network (DNN) is a type of artificial neural network with multiple layers between the input and output layers. The "deep" in deep learning refers to the use of more than one hidden layer in the network. DNNs are designed to model complex, high-level abstractions in data by using multiple processing layers, composed of neurons with nonlinear activation functions. The structure of the deep neural network shown in Fig2.

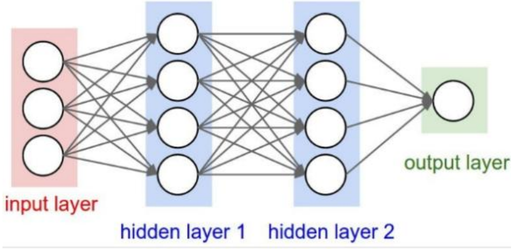

Fig 2 Structure of Deep Neural Network

Key Components of a Deep Neural Network

Input Layer: This layer receives the input data. The number of neurons in this layer corresponds to the number of features in the input data.

Hidden Layers: These are the intermediate layers between the input and output layers. DNNs typically have multiple hidden layers, where each layer consists of neurons that process and transform the input data. The depth of the network refers to the number of hidden layers.

Output Layer: This layer provides the final output of the network. The number of neurons in this layer corresponds to the number of desired outputs.

Neurons: Also known as nodes or units, neurons are the fundamental processing elements in each layer. Each neuron receives input from neurons in the previous layer, processes it through a weighted sum and an activation function, and passes the output to the next layer.

Weights and Biases: Each connection between neurons is associated with a weight, which is adjusted during training to minimize the error in the network’s output. Biases are additional parameters that help in adjusting the output along with the weighted sum.

Activation Functions: These functions introduce non-linearity into the network, allowing it to learn and model complex patterns. Common activation functions include ReLU (Rectified Linear Unit), sigmoid, and tanh.

Loss Function: This function measures the difference between the predicted output and the actual output. It guides the training process by providing feedback on how well the network is performing. Common loss functions include Mean Squared Error (MSE) for regression tasks and Cross-Entropy Loss for classification tasks.

Optimizer: The optimizer adjusts the weights and biases of the network to minimize the loss function. Common optimization algorithms include Stochastic Gradient Descent (SGD), Adam, and RMSProp.

Working of a Deep Neural Network

Forward Propagation: Input data is passed through the input layer, through the hidden layers, and finally to the output layer. At each neuron, the data is transformed by weights, biases, and activation functions.

Loss Calculation: The output of the network is compared with the actual target values using the loss function to compute the error.

Backward Propagation (Backpropagation): The error is propagated back through the network to update the weights and biases. This is done using the gradient of the loss function with respect to each weight and bias. Optimizers use these gradients to perform the actual updates.

Iteration: The process of forward propagation, loss calculation, and backward propagation is repeated for many iterations (epochs), with the weights and biases being adjusted each time to minimize the loss.

Activation Functions

Activation functions help map input values to a known range, which helps stabilize training and helps map values to a desired output in the last layer. The most popular activation functions include binary step, linear, sigmoid, tanh, and ReLU. In this work ReLU activation function is used.

Supporting Previous Methods

Credit card transactions are classed as either fraudulent or legitimate and are primarily a problem of binary classification. Fundamentally, the classification of data mining includes challenges such as Fraud detection, which is used to determine if credit card transactions are fraudulent or legitimate. In addition to data mining, some additional techniques and factor methods involved in fraud detection are Web- services based collaborative schemes in which private bodies such as banks can share information about fraud patterns and frequencies to improve fraud detection capability and reduce financial loss. The fundamental technique for constructing any Machine Learning model [1] is depicted in Figure 3 below. Numerous Machine Learning approaches have been created and utilized in experimental investigations addressing the subject of credit card fraud detection. In the [1], a data-driven method for establishing fraud alerts based on certain criteria such as Oversampling under sizes and SMOTE methodology is proposed. In their research [2], a few authors have encountered a comparison of numerous models and their resulting analyses, such as XGBOOST, Random Forest, Decision Trees, etc. These are the most prevalent fraud detection methods now in use. There has also been research into new techniques, such as Adaboost and Majority Voting, that add to or improve the performance of ML algorithms. [3 , 4] It has been observed that Artificial Neural Network (ANN) uses Feature Selection (FS) with optimization to choose the required characteristics when implementing the algorithms [5]. When multiple valid parameters are present, it is essential to choose the most effective one. Since most models are irrelevant in transaction sequencing, they cannot learn from a single level of input; hence, a new structured sequenced learning ensemble classifier that enhances performance is also observed [6, 7]. In another paper [8] by S. Venkata Suryanarayana et al. Numerous classes, as well as their metrics and performance, have been examined. This gives an indication of the number of metrics that can be considered while determining the optimal algorithm. Adaptive features selection processes may be analysed by comparing five distinct ways [9]; these adaptive characteristics make it easier to divide and eliminate the unimportant ones. The graph below illustrates the precision and accuracy found in [10, 11].

The research by Priya Gupta et al. that analyses in depth the major elements of Random Forest and its limits discusses the systematic, comprehensive investigation on the application of Random Forest. In addition, the concept of real-time deep learning and binary data categorization by multiple approaches is described in the referenced study [12], which compares these techniques. Hassan Najadat and others believe that the bidirectional Long short-term memory (BiLSTM) and bidirectional Gated recurrent unit (BiGRU) are a powerful and novel tool for enhancing performance [13].

3. METHODOLOGY

Materials: For implementation, the default credit card dataset used form the UCI [14]. The dataset consists of different features such as transaction timing, residence country, gender details, card type, card transaction type, card entry mode and amount details. Apart from above feathers in this work the following new features are included such as Transaction speed, Transaction min and max time, keyword accessing time to enter password and user name, and minimum and maximum amount for last few transactions are considered for the implementation. The main spatial and temporal feathers for implementations are shown in the Table 1.

Table 1 Various Features of transaction

S.No	Features	Description
1	Time and Dates	Transaction Time and Dates
2	Accessing Country	Residence of Transaction Country
3	Entry mode	magnetic stripe, contactless
4	Transaction type	Internet, PIN, POS, 3D Secure
5	Max and Min Limit of Transaction	Last one-year min and maximum transaction.
5	Speed of Transaction	Min and max time to taken for transaction on online and offline.
6	Speed of Utilities	All the types of human activities related to inputs.
7	Patten and speed of accessing Patterns	Password typing speed, accessing card patterns and etc.

Methods: The basic working flow of credit card normal transaction and fraud detection model shown in Fig 2.

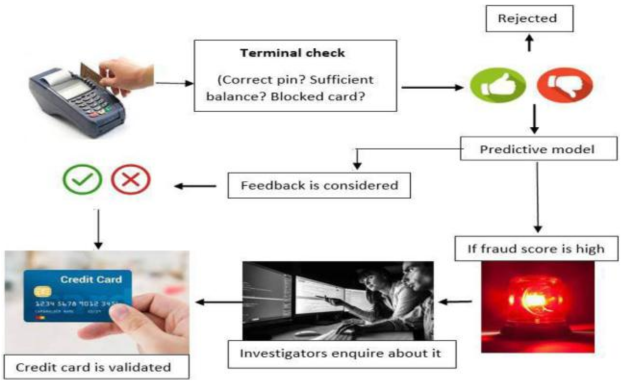

Fig.2 Credit card transaction and fraud detection flow

The structure of the proposed method has been represented in the Figure 2. The proposed structure consists of two techniques used for analysis and prediction. The combination of fuzzy logics and Deep neural network methods are used to analysis and predictions of credit card frauds detection. The fuzzy logic is used to find the relationship and similar possibilities of prediction of fraud detections and also DNN is used to predict the fraud activities. The structure of the prediction and analysis of the proposed work shown in the fig 3.

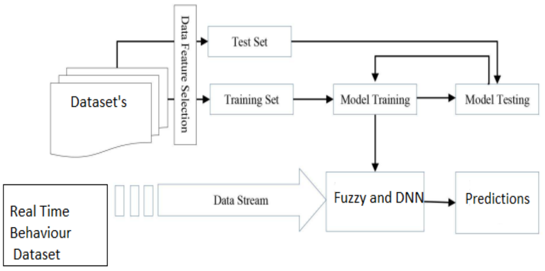

Fig 3 Proposed model Prediction process

Fuzzy logic

Fuzzy logic concept is based on the combination of probabilities theory and accept equivalent impression for decision making. Basically, bivariant approach for finding the values of prediction is true or false. The fuzzy rule-based inputs are transferred the input into logical output using the linguistic rules and transferred the true or false output. Specifically, the rule-based inputs such as if- then rules are transfer the output. Based on the if-then and crisp outputs inference results are created. In this work based on the different features are mentioned in Table 1, different inference results are created. The two types of inference results are created such as based on the if then operations and reach conclusion.

Deep Neural network

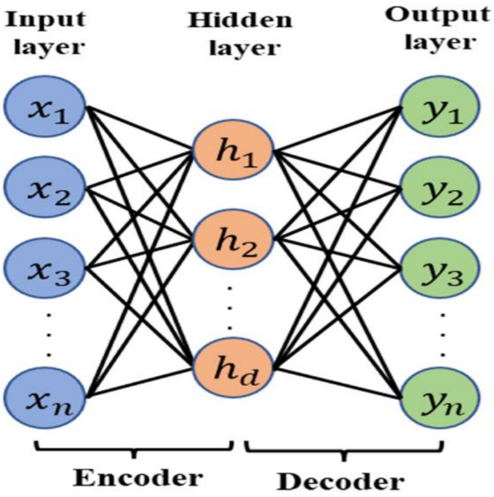

Fig 4 Deep neural network works

Deep neural network used auto-encoder for getting equal input and output. The deep neural network basically having two Restricted Boltzmann for parallel inputs and it shown in the Figure. The auto encoder having two parts such as encoder and decoder. The encoder compresses the set of inputs into fewer parts and decoder reconstruct the input parts. This deep neural network used three encoders and three decoders with 6 hidden layers. For getting higher metrics, the usage of tanh activation functions. The tanh function is the hyperbolic tanget and the process the elements of X set inputs. And also, it processes the real and complex input for getting the valid output for achieving target output. The structure of the encoder and deep neural network shown in Figure 6.5. The processing steps for proposed work mathematical representation are as follows. The equation 1 represent the set of input process

$X= \{X1, X2…. Xn\}$

X – denotes the set of inputs, n- represents the number of features for processing the inputs. The equation 2 represent the output process equation using the different inputs and output.

$Y=\{X_1n,X_2n_2,\ldots,X_nn_n\}*Z$

Y- denoted the set of targeted output using the combination of different inputs and outputs, Z represent different weight based on the input combinations.

The input and possible features are connected cardinality using the fuzzy logic and related membership. The equation 3 represent the combination of different fuzzy relationships with cardinality

$$|A| = |B|$$

The A and B having the same cardinality with set of features. The equation 4 represent the set of input and weight process of mapping with output Y

$$Z = W_0 + W_1X_1 + W_2X_2 + \ldots + W_nX_n$$

W represent different weighted averages and X is the number of continuous inputs. The activation function representation shown in the equation 5.

$$Y = \tanh(X)$$

The entire process of inputs and output processed continuously using the above-mentioned equations binary classification of fuzzy set and prediction of fraud detection.

4. RESULT AND DISCUSSIONS

Experimental Setup: The proposed work is used open-source dataset from the UCI [14] and implemented using the python programming language. The different static and dynamic features are mentioned in the Table 1. The time, date country, typing speed and set of other dynamic features used for implementations. Initially, the dataset is processed using the LSTM algorithm and balanced and unbalanced data are processed [15]. The results of balanced and imbalanced data representation have shown in the figure 3. Initially the data set is classified as valid and invalid. Using the total number of records (284807), the 511 records are considered as the invalid. The experimental setup used different folds such as 1, 5, 10, and 15 and 20 folds. Each process of folds data is equally trained and tested in the proposed method. The training, test and validation methods for the proposed work is 60%, 15% and 15%.

Evaluation Metrics: The different evaluation metrics are used for comparing the proposed work. In this work three main parameters are used for evaluation such as accuracy, sensitivity, and specificity [15]. The F1 score is used to find the test accuracy based on the binary classifications. The F1 score is used to find using recall and precision using the dynamic features. And also, relevant true positive and false positive values are mapped using the set of features. The basic accuracy and F1 score values are compared with different methods such as logistic regression, decision tree, random forest, K-NN and ANN-DL algorithms [16]. In this work ,the comparation of recent three main methods and presented in the Table 2.

Table 2 F1-Score and Accuracy of prediction

Method	**F1-Score**	**Accuracy**
Random Forest Classifier	0.446	80.16
Logistic Regression	.015	78.18
K- Neighbour’s Classifier	0.23	72.50
ANN-DL	0.44	77.63
FL+ DNN	**0.33**	**88.10**

Comparison: The proposed work result is performed using the different iterations and results are consolidated using the above-mentioned comparison parameters. In this discussion three 10, and 15- and 20-folds results are summarised and comparison results are plated in Table 3,4 and 5. The Neurl Network, KNN and Logistic Regression results are tested in the three mentioned folds. The above-mentioned folds proposed results are shown the better accuracy in term of binary classifications. Because the different features are checked and verified with the help of fuzzy metrics. The fuzzy metrics verifies the relation between each other variables such as A and B. The three-fold values of sensitivity and specificity is represented in the Table 3.

Table 3 10 folds Results

Methods	**Folds**	**Accuracy**	**Sensitivity**	**Specificity**
Logistic Regression	10	96.31	96	80.4
K-Neighbour’	10	95.8	96. 6	74.4

s Classifier				
ANN-DL	10	93.99	95.5	68.5
FL+ DNN	**10**	**97.97**	**97.2**	**88.4**

Table 4 15 folds Results

Methods	**Fold s**	**Accurac y**	**Sensitivit y**	**Specificit y**
Logistic Regression	15	0.9624	96.85%	80.62%
K-Neighbour's Classifier	15	0.9585	96.71%	74.50%
ANN-DL	15	0.9372	96.40%	80.31%
FL+ DNN	**15**	**98.97**	**97.2**	**86.4**

Table 5 20 folds Results

Methods	**Fold s**	**Accurac y**	**Sensitivit y**	**Specificit y**
Logistic Regression	20	0.96238	96.85%	80.49%
K-Neighbour's Classifier	20	0.95838	96.70%	74.44%
ANN-DL	20	0.92813	96.61%	74.33%
FL+ DNN	**20**	**98.0**	**96.2**	**86.4**

5. RESULTS ANALYSIS

In card fraud detection systems, when a transaction raises a fraud flag, the transaction is refused, and the user must go through a verification process to verify if it's the case of a false flag or a real fraud. This verification processes vary from a phone call to a series of verification forms. The cost of a false flag is then equivalent to the cost of these processes, which is significantly lower than the cost of a fraud case. However, when the number of false flags is important, purchases get blocked more frequently by error, making one's use of a credit card tedious and time consuming, moreover it can generate big loses for either sides of the transaction. thus, the model should have a balanced mount of fraudulent transactions caught and false flags raised.

Table 6: Confusion Matrix

Algorithms	**TB**	**FP**	**TN**	**FN**
Linear SVM Regression	341	2111	115173	234
Logistic Regression	302	1290	116002	270
NN Based Classification	384	2229	115049	198
Non Linear Auto Regression	430	3871	113408	151
Deep NN With fuzzy logic	370	1557	118793	260

Table 6 shows the experimental results of the implemented algorithms. Each algorithm was tested four times and for each algorithm the creation of confusion matrix, and represented the TP FN FP TN in the table. The non-linear auto regression has caught the most amount of the fraudulent transactions, but at cost of the false flags. Logistic regression raised the least amount of false flags but is not the best at catching the fraudulent transactions. Deep Learning based on the auto encoder has balanced results, with a good amount of frauds caught and a fair amount of false flags. First results seem promising for the Deep neural network model. Let's see the accuracy of the models.

The accuracy is defined as correct predictions made divided by the total number of predictions made. The data shows that Logistic regression followed by the Auto-encoder have the best accuracy and the typical neural network classification method has the worst. Because the data is not well balanced, their may not be draw conclusions from accuracy only. Accuracy can be misleading due to its dependencies. The model Non-linear auto regression has the best Recall, but at the cost of precision. Deep NN auto encoder, in the other hand, has the best precision, with results close to logistic regression. In this case of study, Deep NN with fuzzy logic has the overall best F1 score followed by logistic Regression, which confirms that Deep NN auto-encoder is the best fitted algorithm amongst the tested ones. Moreover, the deep learning algorithm used in this study is very simple; with more tuning of the parameter (Hyper-parameter Tuning with Grid Search) the better results finded. This provides good insight in which algorithm should be chosen in building the prediction model, and Deep Neural network is choosen with auto-encoder.

6. CONSULTION

The different approaches are used to predict and analysis the credit card fraud activities. In credit card fraud prediction behaviour analysis is an important one. In this work, presented new prediction model based on the real time behaviour analysis and with the help of history transaction dataset. The behaviour analysis is performed based on the proactive data such as user typing speed of username and password, normal, abnormal activities and other properties. The proposed model prediction is performed based on the fuzzy logic and deep neural network system. The fuzzy logic used to check the behaviour of human and find the member activities. The deep neural network is used to identifying anomalous behaviour of credit card activities. For implementation, a real-time dataset spatial features and temporal features are used. The accuracy, sensitivity, and specificity metrics are used to measure the performance of proposed work. The results indicate that the better results than in terms of accuracy, sensitivity and specificity.

7. REFERENCES

[1] Vinod Jain, Mayank Agrawal, Anuj Kumar, "Performance Analysis of Machine Learning Algorithms in Credit Cards Fraud Detection, 2020 at 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2022.

[2] D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic and A. Anderla, "Credit Card Fraud Detection - Machine Learning methods," *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2019, pp. 1-5, doi: 10.1109/INFOTEH.2019.8717766.

[3] Krishnaiah, P.B. Divakarachari, Automatic Music Mood Classification using Multi- class Support Vector Machine

based on Hybrid Spectral Features, 2022.

[4] Gurumurthy Krishnamurthy Arun ∗ , Kaliyappan Venkatachalapathy, Intelligent feature selection with social spider optimisation based Artificial Neural Network Model for Credit card Fraud detection, in: | Arun & Venkatachalapathy, 11, IIOABJ, 2020, pp. 85–91.

[5] T.G. Nguyen, T.V. Phan, D.T. Hoang, T.N. Nguyen, C. So-In, Efficient SDN-based traffic monitoring in IoT networks with double deep Q-network, in: International conference n computational data and social networks, Springer, Cham, 2020, pp. 26–38.

[6] Xurui Li, Wei Yu, Tianyu Luwang Jianbin Zheng, Xuetao Qiu, Jintao Zhao, Lei Xia Yujiao Li in " Transaction Fraud detection using GRU-Centered Sandwich-structured Model at, in: Proceedings of the 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design, 2022 .

[7] K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system, IEEE Trans. Intell. Transp. Syst. 22 (7) (2020) 4337–4347.

[8] S. Venkata Suryanarayana ∗ , G.N. Balaji, G. Venkateswara Rao in "Machine learning approaches for credit card fraud detection " at, Int. J. Eng. Technol. 7 (2) (2018) 917–920 .

[9] Ajeet Singh, Anurag Jain, Adaptive Credit Card Fraud Detection Techniques Based on Feature Selection Method, University Grants Commission (UGC), Delhi, India,2022.

[10] Vaishnave Jonnalagadda, Priya Gupta, Eesita Sen in "Credit card fraud detection using Random Forest Algorithm" International Journal of Advance Research, Ideas and Innovations in Technology, Volume 5, Issue 2.

[11] B.D. Parameshachari, K.M. Keerthi, T.R. Kruthika, A. Melvina, R. Pallavi, K.S. Poonam, Intelligent Human Free Sewage Alerting and Monitoring System, in: 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021), Atlantis Press, 2021, pp. 480–486 .

[12] Hassan Najadat, et al., Credit Card Fraud Detection Based on Machine and Deep Learning, 2020 11th International Conference on Information and Communication Systems (ICICS), 2022.

[13] Anuruddha Thennakoon, Chee Bhagyani, Sasitha Premadasa, Shalitha Mihiranga, Nuwan Kuruwitaarachchi in "Real-time Credit Card Fraud Detection Using Machine Learning, 9th International Conference on Cloud Computing, Data Science & Engi-neering (Confluence), 2022

[14] https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients

[15] N. Prabha and S. Manimekalai, "Imbalanced data Classification in Credit Card Fraudulent Activities Detection using Multi-Class Neural Network," *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, 2022, pp. 131-138, doi: 10.1109/ICAIS53314.2022.9742878.

[16] Madhurya, M. J., H. L. Gururaj, B. C. Soundarya, K. P. Vidyashree, and A. B. Rajendra. "Exploratory analysis of credit card fraud detection using machine learning techniques." Global Transitions Proceedings 3, no. 1 (2022).