

Population based Optimized and Condensed Fuzzy Deep Belief Network for Credit Card Fraudulent Detection

Jisha M.V¹

Research Scholar
Dept. of Computer Science
Nehru arts and Science College
Coimbatore, Tamilnadu, India

D. Vimal Kumar²

Associate Professor
Dept. of Computer Science
Nehru arts and Science College
Coimbatore, Tamilnadu, India

Abstract—In this information era, with the advancement in technology, there is a high risk due to financial fraud which is a continually increasing menace during online transactions. Credit card fraudulent identification is a toughest challenge because of two important issues, as the profile of the credit card user's behavior changes constantly and credit card datasets are skewed. The factors which greatly affects the credit card fraudulent transaction detection are primarily based on data sampling models, features involved in feature selection and detection approaches implied. To overwhelm these issues, instead of using certainty theory, this paper encapsulates with three different empowered models are deployed for intellectual way of fraudulent transaction detection. In this work uncertainty theory of intuitionistic fuzzy theorem to determine the significant features which will influence the detection process effectively. Maximized relevancy among dependent and independent features of credit card dataset are determined using grade of membership and non-membership information of each features. The intuitionistic fuzzy mutual information with the knowledge of entropy it selects the features with highest information score as significant feature subset. This proposed model devised Fuzzy Deep Belief Network enriched with Sea Turtle Foraging for credit card fraudulent detection (EFDBN-STFA). The fuzzy deep belief network greatly handles the complex pattern of credit card transactions with its deep knowledge and stacked restricted Boltzmann machine the pattern of dataset is analyzed. The weights assigned to the hidden nodes are fine-tuned by the sea turtle foraging using its fitness measure and thus it improves the detection accuracy of the FDBN. Simulation results proved the efficacy of EFDBN-STFA on two different credit card datasets with its gained ability of handling hesitation factor and optimization using metaheuristic approach, it achieves higher detection rate with reduced false alarms compared to other existing detection models.

Keywords—Credit card fraudulent; uncertainty; intuitionistic fuzzy; fuzzy deep belief network; sea turtle foraging

I. INTRODUCTION

In modern days, usage of internet for commercial transactions started increasing exponentially because of its availability and flexibility. Usage of credit cards for online or offline transactions is very useful for business peoples [1]. But credit card fraud becomes significant issue in financial sectors, banks and card issuers. This credit card fraud detection

becomes an important and interesting topic of research for the scientific community. To handle such voluminous transaction system, it needs high sophisticated security system to analyze transactions and detect fraud transactions more quickly. This necessitates with the advantage of modern technology utilizing machine learning, mining and artificial intelligence influence's the process of credit card detection more accurately. Detection of fraud has become an important activity which involves in reducing the impact of fraudulent transactions [2]. But credit card fraud detection is very challenging from the perspective of learning process due to its nature of class imbalance.

Generally, classification models are deployed to examine all the approved transactions and alert the utmost suspicious ones. The alerts are investigated by the professionals and intimate the cardholders to discover whether it is genuine or fraudulent transactions for each altered transaction [3]. This provides feedback to the classification system during the training phase to update itself and detect the fraud detection more accurately. Meanwhile, it is very essential to highlight the major difference among user behavior and fraud analysis. The fraud detection models extract the signature of fault tricks pattern and it greatly assist during the testing process. The ultimate goal of this work is to reduce the false detection of fraudulent detection in a more precise way. This is achieved by developing a Population based Optimized and Condensed Fuzzy Deep Belief Network for Credit Card Fraudulent Detection.

The rest of the research is organized as follows: Section I gives the importance of detection of frauds, Section II highlights the related works done before. Section III describes the methodology of the proposed work. Section IV represents the conceptual framework used and Section V gives the results and its discussion. Followed by the conclusion and references.

II. RELATED WORK

This section discusses about some of the existing work which involves in credit card fraudulent detection using machine learning algorithms and mining approaches.

Vaishnavi et al. [4] in their work anticipated a novel approach for fraud detection on streaming transaction data,

which analysis past history of customers transaction details. The behaviour patterns of transaction are extracted and card holders with same patterns are clustered depending on their amount of transaction. They used sliding window concept and transactions are aggregate to determine fraudulent and genuine transaction.

Imane et al. [5] developed a comparative model which comprised of various approaches of machine learning models are deployed for credit card fraud detection. The authors mainly focused on investigating neural network performance. They stated that this study aims to guide the researches to choose best approaches for credit card fraud detection.

Wen-Fang et al. [6] presented an outlier mining model to accurately forecast fraudulent credit card transaction. Distance summing algorithm is used to emulate variation among normal and fraud detection. The outlier approach is mainly used to detect anomalous transactions.

Maniraj et al. [7] devised a recognition model to check whether a new incoming transaction normal or fraudulent. They performed preprocessing and analyzed PCA converted credit card transaction data. They deployed isolation forest model and local outlier detection for classifying multiple type of anomaly detection.

Andrea et al. [8] in their work contributed three different approaches to discover fraudulent transactions. They handled class imbalance problem by designing a novel learning policy. They worked with real time dataset with the concept of drifting and verification.

Navneet et al. [9] investigates fraudulent transactions in banking sectors and analyzed vulnerabilities during online transaction. This work explores possible ways to prevent fraud transaction by developing graph database and finding the patterns of fraud transaction.

Salvatore et al. [10] in their work stated that understanding the purpose of meta learning policies will greatly influence during the process of fraud catching rate and deduction rate. They used skewed distribution to work with balanced data during training that results in better classification.

Dheepa and Dhanapal [11] developed a behavioral classification model using support vector machines. The features are extracted to determine the significant behavioral transaction patterns. If there any conflict occurs then it is examined as suspicious and this is considered to discover the frauds.

Chuang et al. [12] presented a mining model, which uses web services on online bank transaction. The banks which are involved in these scheme shares their knowledge about fraud patterns in a heterogenous environment and with the distributed system it further improves the ability of fraud detection with less financial loss.

Tao Guo et al. [13] developed a neural network model to discover customer's behavior pattern. The significant task is to discover any deviation from the usual transaction pattern. This is achieved by training the neural network with dataset and their confident value is computed. Those credit card

transactions with less confident value is treated as fraudulent transaction.

Suvasini Panigrahi et al. [14] in their work designed a fusion model which comprised of four components like Dempster-Shafer, rule-based filter, transaction history and Bayesian model. Rule filter is used to determine fraudulent transactions. DST is used to compute belief value of each transaction based on its evidence value.

III. METHODOLOGY OF POPULATION BASED OPTIMIZED AND CONDENSED FUZZY DEEP BELIEF NETWORK FOR CREDIT CARD FRAUDULENT DETECTION

This proposed work aims to overcome the ambiguity, vagueness and uncertainty in prediction of credit card fraudulent detection shown in "Fig. 1". This work used two different datasets where one is collected from Kaggle repository and another dataset is collected with a case study of a specific bank. The existing models use the neural networks, support vector machine, random forest and other conventional classification models to perform this fraudulent detection process. Most of them fail to concentrate on handling vagueness and inconsistency which often arise in the real time dataset when there are transactions which cannot be finitely defined either as fraudulent or normal transaction. To overcome this problem the proposed model works in two stages, in order to reduce the redundancy among features involved in fraudulent detection and increase the relevancy among feature and class. This is achieved by adapting intuitionistic fuzzy mutual information as feature subset selection, whose ultimate goal is to choose the most significant attributes involved in process of prediction. The pattern of credit card transactions is analyzed in depth by using fuzzy Deep belief network which is fine tuned by introducing sea turtle foraging algorithm which optimizes the assignment of weights in FDBN.

A. Dataset Description

This work used two different credit card datasets for fraudulent transaction detection. The first dataset comprised of 284,807 transactions [15]. The input variables are in PCA transformation and they are denoted as V1, V2, ... V28 vector values. Other variables are time, amount is not transformed and class is a feature which is considered as a target variable. The second credit card dataset comprised of 30,000 transactions with 25 features including the class variable [16]. The features of this dataset are limit value, gender, education, marital status, age, payment and billing details. This proposed work uses these two credit card datasets to discover the fraudulent transaction.

B. Normalized Intuitionistic Fuzzy Mutual Information

A searching procedure which selects a subset of features which greatly influence the classification process is known as feature subset generation. This method applies a subset evaluation function to assess the current subset, if the present subset performs better than the previous subset, the current subset is replaced with the previous set [17]. This is a cyclic process which repeats the subset generation and evaluation until termination condition is met. The termination condition relies on both evaluation and generation function, the former

case the iteration terminates when the insertion or deletion of an attribute doesn't produce a better subset. In later case until a predefined number of attributes are selected or specified number of iterations are done.

This research work develops a normalized intuitionistic fuzzy feature subset selection scheme, which starts with an empty feature subset E. Consecutively, each feature is selection in such a way that it maximized the criteria of evaluation and add the relevant feature to E. The selected feature subset is evaluated based on the minimum redundancy and maximum relevancy principle. Each features relevancy is measure using Intuitionistic Fuzzy mutual information (IFMI) between feature fr_i and the class variable fr_{cl} . The feature's redundancy is evaluated by finding IFMI between fr_i and the subset of previously selected features which are in the E list is computed. To overcome the biased nature of multivalued features this work uses Normalized Intuitionistic Fuzzy mutual information (NIFMI). The NIFMI among two features fr_s and fr_t is computed by finding the ratio between the intuitionistic fuzzy mutual information IFMI(fr_s ; fr_t) of the two attributes and the minimum entropies of those two attributes ($H(fr_s):H(fr_t)$). Likewise, Normalized Intuitionistic Fuzzy mutual information is defined as.

$$NIFMI(fr_s, fr_t) = \frac{IFMI(fr_s, fr_t)}{\text{Min}\{IFMI(fr_s, fr_t)\}} \quad (1)$$

Let F be the list of attributes of a dataset, which is represented as $F = \{fr_1, fr_2, fr_3, fr_4, \dots, fr_n\}$ where n denotes number of attributes. Let us defined that C and D are two Intuitionistic fuzzy sets defined on the fuzzy sets Y. The Intuitionistic fuzzy membership value of R_{th} feature for i^{th} class represented as $\mu_{i,r}$, degree of non-membership value is $\vartheta_{i,r}$ and its degree of hesitation is represented as $\pi_{i,r}$.

The membership value $\mu_{i,r}$ is computed as

$$\mu_{i,r} = \left(\frac{\|\bar{fr}_i - fr_r\| \sigma}{d + \epsilon} \right)^{\frac{-2}{q-1}} \quad (2)$$

The non-membership value $\vartheta_{i,r}$ is computed as

$$\vartheta_{i,r} = \frac{1 - \mu_{i,r}}{1 + \tau \mu_{i,r}} \quad (3)$$

The indeterminacy value $\pi_{i,r}$

$$\pi_{i,r} = 1 - \mu_{i,r} - \vartheta_{i,r} \quad (4)$$

Where q is the intuitionistic fuzzification coefficient, $\epsilon > 0$ is used to evade distinctiveness, σ denotes standard deviation while performing distance calculation [18]. \bar{fr}_i signifies attributes mean value which belongs the class i. The radius of data d is denoted as $d = \max(\|\bar{fr}_i - fr_r\| \sigma)$. The intuitionistic fuzzy entropy (IFE) of the fuzzy sets C and D is computed as follows:

$$IFE(C) = -\frac{1}{n} \sum_{fr \in F} [\mu_C(fr) \log \mu_C(fr) + \vartheta_C(fr) \log \vartheta_C(fr) - (1 - \pi_C(fr)) \log(1 - \pi_C(fr))] \quad (5)$$

$$IFE(D) = -\frac{1}{n} \sum_{fr \in F} [\mu_D(fr) \log \mu_D(fr) + \vartheta_D(fr) \log \vartheta_D(fr) - (1 - \pi_D(fr)) \log(1 - \pi_D(fr))] - \pi_D(fr) \quad (6)$$

$$IFE(C \cup D) = -\frac{1}{n} \sum_{fr \in F} [\mu_C(fr) \vee \mu_D(fr)] \log [\mu_C(fr) \vee \mu_D(fr)] + [\vartheta_C(fr) \vee \vartheta_D(fr)] \log [\vartheta_C(fr) \vee \vartheta_D(fr)] - [1 - \pi_C(fr) \vee \pi_D(fr)] \log [1 - \pi_C(fr) \vee \pi_D(fr)] - [\pi_C(fr) \vee \pi_D(fr)] \quad (7)$$

$$IFMI(C, D) = IFE(C) + IFE(D) - IFE(C, D) \quad (8)$$

Normalized Intuitionistic Fuzzy Mutual Information based feature selection with maximized relevancy and minimized redundancy of an Intuitionistic Fuzzy dataset C with the class feature cl is computed as in "equation (9)".

$$NIFMI(C, f_{cl}) = \frac{IFMI(C, f_{cl})}{\text{Min}\{IFMI(C, f_{cl})\}} \quad (9)$$

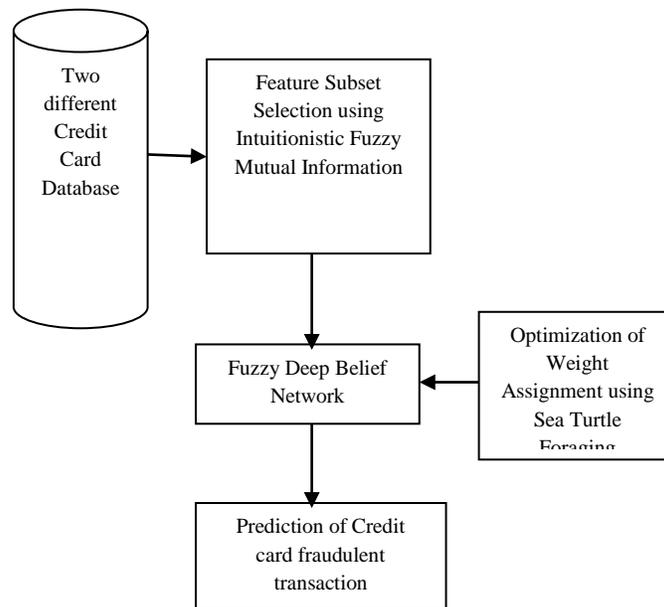


Fig. 1. Overall Workflow of the Proposed Population based Optimized Fuzzy Deep Belief Network for Credit Card Fraudulent Detection.

C. Fuzzy Deep Belief Network

Prior to assigning the input to the DBN network the domain value has to be converted to the fuzzy representation using the formula:

$$\mu_{CDS}(x) = \left\{ \begin{array}{ll} 0 & (x < a) \text{ or } (x > d) \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & b \leq x \leq c \\ \frac{d-x}{d-c} & c \leq x \leq d \end{array} \right\} \quad (10)$$

where $\mu_{CDS}(x)$ refers to the membership of each instances of credit card dataset towards normal transaction as in “Fig. 2”.

A sort of Deep neural network which comprised of multiple layers of belief network known as deep belief network. In this model each layer is a Restricted Boltzmann Machine (RBM) which are stacked to each other and constructs deep belief network. DBN consist of two various types of networks they are belief network and restricted Boltzmann machine [19]. A Belief network is comprised of layers of stochastic binary units whose connections are weighted. This network is acyclic graph which permits to observe the kind of data the belief network believes. It adjusts the weights of the states between these units so that the network can produce appropriate result. The binary units in belief networks have either the state 0 or 1.

The initial process of DBN is to learn a layer of features of the visible units with contrastive divergence method. Next, to treat the activations of previously trained features as visible units and learn features of features in a second layer. At last, the entire DBN is trained when the final hidden layer finished its learning process. The greedy learning approach is used for training the DBN, because while training RBM with CD for each layer it falls under local optimum and the next stacked RBM layers takes those trained optimal values and look more local optimum. Finally, all the layers are consistently involving for local optima it gets its global optimum.

As shown in the “Fig. 3”, Restricted Boltzmann machine is a recurrent neural network which consist of binary units and undirected edges among units. The probability distributions on visible and hidden units are termed with its function of energy. The functions are formulated as follows:

$$\text{Prob}(Vs, Hd) = \frac{1}{Z} \exp(-Eg(Vs, Hd)) \quad (11)$$

$$Z = \sum_{Vs, Hd} \exp(-Eg(Vs, Hd)) \quad (12)$$

$$Eg(Vs, Hd) = -\sum_i va_i Vs_i - \sum_j hb_j Hd_j - \sum_i \sum_j Vs_i wt_{i,j} Hd_j \quad (13)$$

where Vs refers to the visible node and Hd refers to Hidden nodes, va denotes visible bias, hb signifies the hidden bias of the final layer and finally wt refers to the weight value between the previous layer and the present layer.

Like logistic regression, the conditional probabilities $\text{Prob}(Vs_i = 1 | Hd)$ and $\text{Prob}(Hd_j = 1 | Vs)$ and when a hidden vector $Hd(Hd_1, \dots, Hd_j, \dots, Hd_m)$ is known, the activation probability of the ith visible unit can be computed as follows:

$$\text{Prob}(Vs_i = 1 | Hd) = \sigma(va_i + \sum_{j=1}^n wt_{i,j} Hd_j) \quad (14)$$

Similarly, Activation probability of jth hidden unit can be formulated when a set of visible vector $vs(vs_1, \dots, vs_i, \dots, vs_n)$ is represented as:

$$\text{Prob}(Hd_j = 1 | Vs) = \sigma(hb_j + \sum_{i=1}^m wt_{i,j} Vs_i) \quad (15)$$

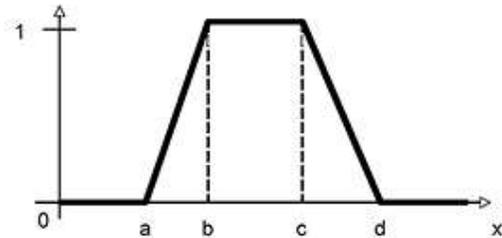


Fig. 2. Representation of Membership Function.

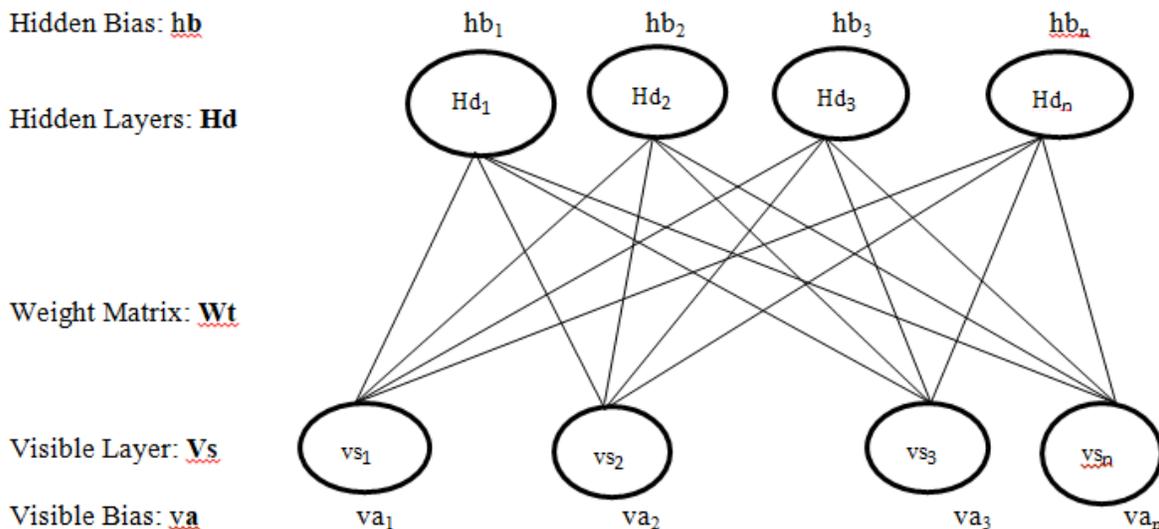


Fig. 3. Deep Belief Network RBM.

where, $\sigma(\cdot)$ is a sigmoid function, $w_{i,j}$ is the link weight among the i^{th} visible unit and the j^{th} hidden unit and the hb_j is the hidden bias of the j^{th} hidden unit.

To maximize the joint probability of the bunch of training inputs then it is signified as in “equation (16)”.

$$\arg \max_{wt} \prod_{v \in V_S} Prob(V_S) \quad (16)$$

where, V_S is set of all training inputs of dataset.

D. Sea Turtle Foraging Algorithm

This algorithm is inspired by the sea turtle’s food searching behaviour. The sea turtle senses a kind of odor smell known as dimethyl sulfide came from their sources of food and they move towards the food source which gives out the strongest odor [20]. Ocean current also helps the turtle for their movement. The artificial foraging process of sea turtle is discussed in the subsequent steps

Algorithm: Sea Turtle Foraging Algorithm.

Steps involved:

1. Initialize N population of turtles.
2. Initialize the position of each turtle in a random manner

$$Pos_i(0) = [pos_1^i, pos_2^i, pos_3^i \dots pos_D^i] \quad (17)$$

where, $i = 1$ to N number of turtles and D refers to D dimensional searching space which is contentious

3. Generate the initial velocities of turtles in a random manner $Vel_i(0) = [vel_1^i, vel_2^i, vel_3^i \dots vel_D^i]$ and the velocity is controlled with in the predefined boundaries as follows:

$$Vel_{\max} = \alpha[TUB-TLB] \quad (18)$$

$$Vel_{\min} = -Vel_{\max} \quad (19)$$

where, TUB and TUL are the upper bound and the lower bound of the D dimensional search space of the turtle. α is the constant variable with the value ranges from 0 to 1.

4. Initial position of M food sources generated randomly

$$Fd_j(0) = [fd_1^j, fd_2^j, fd_3^j \dots, fd_D^j] \quad (20)$$

where, $j = 1$ to M food sources which has D dimensional space for searching it.

5. Initial position of each source of food is given as the input into the objective function and estimate it to get the fitness value of that source of food.

6. Assign the position of each turtle as objective function and estimate it to get the fitness value of the concern turtle, turtle with the highest fitness value is recorded as I

$$I = \arg \max_i (fit_{pos_i(t)}) \quad (21)$$

where, $f_{pos_i(t)}$ is the turtle i ’s fitness value at time t.

7. Velocity of each turtle is updated as shown:

$$vel_i(t) = vel_i(t-1) + \left(\frac{fit_{pos_i(t)} - fit_{pos_i(t-1)}}{fit_{pos_i(t-1)}} \right) (Pos_i(t) - Pos_i(t-1)) \quad (22)$$

where $Pos_i(t)$ refers to the turtle position at time t and $Pos_i(t-1)$ signifies the turtle position at time t-1.

8. Compute the ocean currents velocities by

$$VOC_i(t) = [voc_1^i, voc_2^i, voc_3^i \dots, voc_D^i] \quad (23)$$

$$VOC_i(t) = \alpha (Pos_i(t) - Pos_i(t-1))$$

9. Sum the turtle velocity to that of ocean current velocity to get the united velocity

$$UVOC = vel_i(t) + VOC_i(t) \quad (24)$$

10. If the turtle’s fitness value is less than that of the food source, then its contribution of food source (CFS) is represented as

$$CFS_j = \frac{fit_{FS_j}}{\sum_{q=1}^M fit_{FS_q}} \quad (25)$$

where, fit_{FS_j} refers to the food source j^{th} fitness value.

11. Determine the distance among the turtle and the food source using the formula

$$Dist_{ij} = ||Pos_i - FS_j|| \quad (26)$$

12. Compute the level of odor of the food source j seeming by the turtle i

$$C_{ij}(t) = (CFS_j * \exp\left(-\frac{Dist_{ij}^2}{2\sigma^2(t)}\right)) \quad (27)$$

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{T}\right) \quad (28)$$

where, $\sigma(t)$ indicates the level of fading of odor with short-lived time, σ_0 is a persistent set to be equivalent to 1, and T is referring to iterations denoting the longest time the odor completely disappears.

13. Discover the turtle i ’s best food source, which has the highest value of C_{ij} compared to all other food sources available.

$$J = \arg \max_j (C_{ij}) \quad (29)$$

14. Update each turtles position as follows:

$$Pos_i(t+1) = Pos_i(t) + UVOC_i(t) + C_{ij}(t)(FS_j - Pos_i) \quad (30)$$

15. Stop the process if the maximum no of iteration is reached or else go to the step 6.

E. Proposed Algorithm: Fuzzy Deep Belief Network enriched with Sea Turtle Foraging Algorithm (EFDBN-STFA)

```

Input: Credit Card Dataset: CCDS
N : number of instances in the CCDS
Procedure
Stage 1: Feature Subset Selection
    • Apply preprocessing on CCDS
    • For I = 1 to N
        ◦ Calculate the Intuitionistic fuzzy Mutual Information by applying the equations (1) to (9)
        ◦ Generate Reduced Feature subset of CCDS to RCCDS
Stage 2: Convert the input to Fuzzy domain representation by applying equation (10)
Stage 3: Classifying the credit card transaction using EFDBN-STFA
Training Data: Tr-RCCDS; Test dataset Ts-RCCDS
Number of Layers NLY, Number of Epochs NE;
Number of hidden layers in FDBN is HDL1,...,HDLNL
Number. of units in each hidden layer HD1...HDN;
Weight Wt = { wt1, . . . , wtn}; biases Va, Hb
For G = 1 to NLY- 1
    For L = 1 to NE
        For Q = 1 to (TR + TS)
            ◦ Apply the equation (16) for supervised learning done by FDBN
            ◦ Compute the parameter values using the equations (14) and (15)
            ◦ Improve the EFDBNI hidden nodes HDL,N
            ◦ Call Sea turtle Algorithm for Weight optimization in EFDBN
            ◦ Classify the transaction based on the trained EFDBN using the equation (16)
End
End
    
```

IV. CONCEPTUAL FRAMEWORK

The proposed EFDBN-STFA credit card fraudulent transaction detection is deployed using python code. The performance analysis is done on two different credit card datasets. The other detection models used for comparison are Aggrandized Random Forest (RF), Aggrandized Kernel based Support Vector Machine (SVM) and Artificial Neural Network (ANN). The evaluation metrics used for examining the performance of the detection models are done using accuracy, precision and recall measures.

Accuracy: It is defined as the ratio of transaction which are correctly predicted to the total credit card transactions. This measure instantly specifies how well a model is trained.

$$Acc = \frac{Tot.no.of\ correctly\ detected\ geunie\ and\ fraudulent\ transactions}{total\ no\ of\ transaction\ in\ credit\ card\ dataset}$$

Precision: This metric is defined as the ratio of correctly predicted fraudulent transactions to the total transactions predicted as fraud.

$$PrCs = \frac{Tot.no.of\ correctly\ detected\ fraudulent\ transactions}{total\ no\ of\ transaction\ predicted\ as\ fraud}$$

Recall: This is signified as ratio of correctly predicted fraudulent transaction to the actual number of fraudulent transactions in the credit card dataset.

$$Rcl = \frac{Tot.no.of\ correctly\ detected\ fraudulent\ transactions}{total\ no\ of\ acutal\ fraudulent\ transactions\ in\ credit\ card\ dataset}$$

V. RESULTS AND DISCUSSIONS

By applying the normalized intuitionistic fuzzy mutual information feature methods, the five important features of each dataset is selected which is then used in further steps of the proposed system extended fuzzy deep belief network enriched with sea turtle foraging algorithm (EFDBN-STFA). This work is an extension of my previous work aggrandized random forest(RF) and aggrandized kernel based SVM(AKSVM+FPSO) which detects the credit card fraud transactions, thus leading to prediction based on the behavior patterns of the user with important features.

The “Fig. 4(a)” and “Fig. 4(b)” portraits the importance of features which involves in maximizing the relevancy among features with class and minimizes redundancy within features. By applying intuitionistic fuzzy mutual information (IFMI), the figures displays top five best features of two different credit card datasets. IFMI well treats the problem of inconsistencies in determining potential features of this credit card fraudulent detection. The features with highest score are considered for prediction process, using reduced feature subset which influences the process of credit card fraudulent detection more accurately.

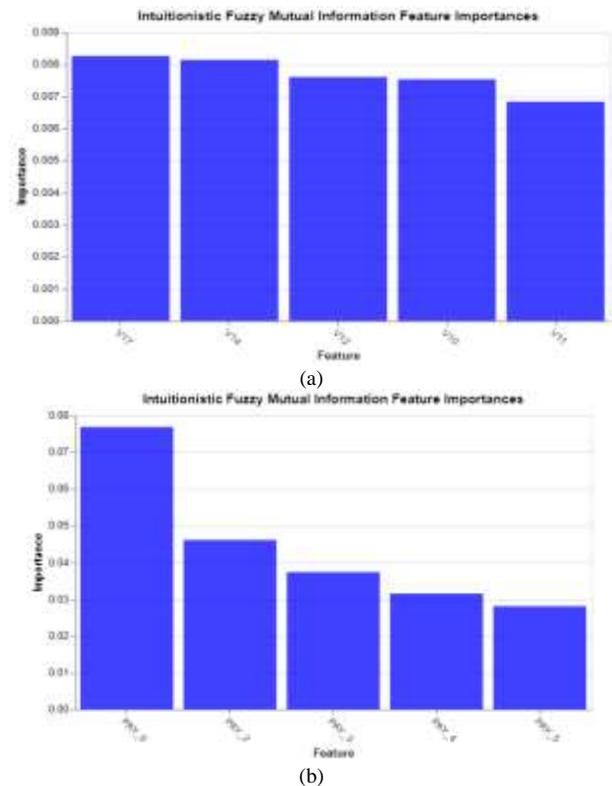


Fig. 4. (a) Feature Selection of Credit Card Dataset1, (b) Feature Selection of Credit Card Dataset2.

From the above “Fig. 5” it is observed that the performance of the proposed EFDBN-STFA produced highest accuracy rate of 96.2% for dataset1 and 95.3% for dataset 2. The other three classification models produce less accuracy because they failed to handle uncertainty in classifying credit card transactions as either genuine or fraudulent. With the knowledge of intuitionistic fuzzy mutual information is used to determine the relevancy among features with class, the redundancy is greatly reduced and the proposed EFDBN-STFA uses only the attributes with highest information score. Thus, it achieves to produce better accuracy while comparing other three models.

The “Fig. 6” illustrates the performance of the four different credit card fraud detection models based on the precision measure on two different credit card datasets. With the ability of the fine tuning their weights and biases assigned to the hidden nodes in fuzzy deep belief network, it surpasses the performance of other conventional detection models. With the enriched knowledge of sea turtle foraging algorithm this proposed model instead of assigning the weights in the random manner, they optimized the weight assignment of the deep belief network more precisely.

It is proved from the results shown in the “Fig. 7” which compares the performance of the four different credit card fraud detection models namely RF, SVM, ANN and EFDBN-STFA. The percentage of total relevant results correctly classified was done by the proposed model EFDBN-STFA. As the real-world credit card datasets cannot be handled in the crisp value to get appropriate interpretation, the fuzzy model which treats them in the form of linguistic terms using membership grade have greatly influence the process of fraud detection process. The IFMI induce the significant features which has to be used as the input is used by EFDBN for credit card fraud detection process. The merit of population-based metaheuristic searching is done by sea turtle foraging mechanism so that the weights assigned to the deep belief networks are fine-tuned based on the optimization and thus this enriched model achieves highest recall value compared to the other models.

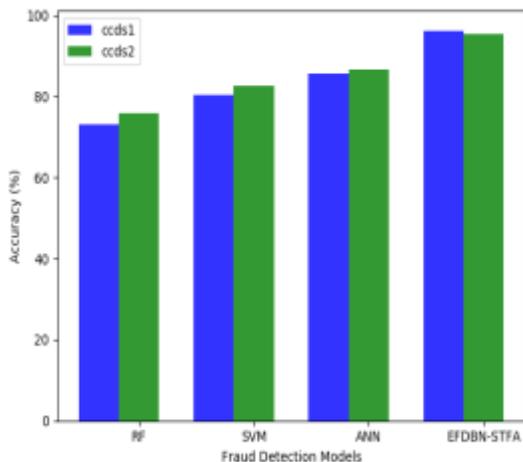


Fig. 5. Accuracy Comparison of Four different Credit Card Fraud Detection Models.

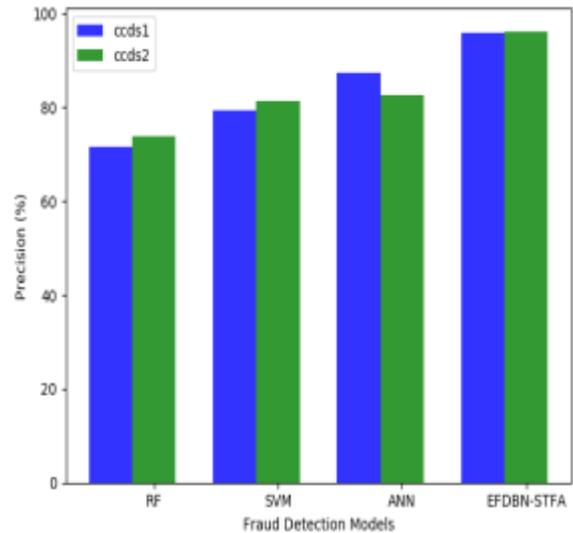


Fig. 6. Precision Measure Comparison of Four different Credit Card Fraud Detection Models.

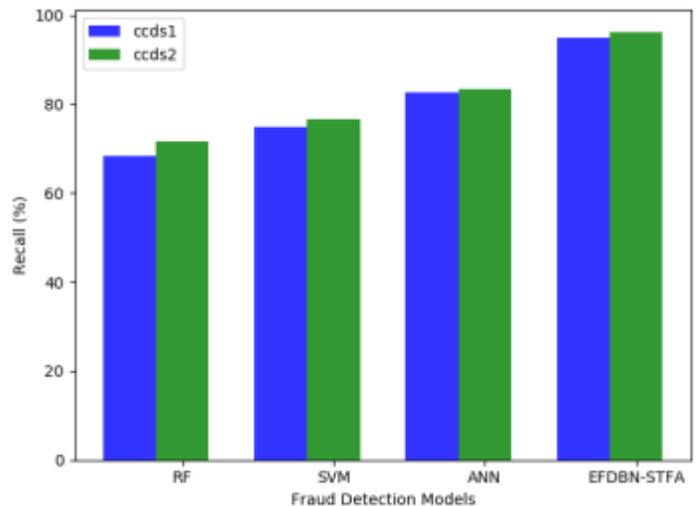


Fig. 7. Comparison of Recall Metric for Four different Credit Card Fraud Detection Models.

VI. CONCLUSION

The ultimate motive of this proposed research work is to detect the credit card fraudulent transaction as it is continuous in nature. This work developed an enriched fraud detection model which handles the presence of vagueness and complexity in determine the pattern of transaction by focusing on three different dimensions. As a primary factor, significant feature of the datasets is selected to handle voluminous credit card datasets using the intuitionistic fuzzy mutual information about the features. With the reduced feature set, the fraud detection process is greatly influenced by fuzzy deep belief network, which gains the deep knowledge of the datasets and the relationship among the features using the stacked restricted Boltzmann machine. Instead of assigning the weights in a chaotic manner, with the inspiration of sea turtle foraging based optimization the weights assigned to the hidden layers are fine-tuned and thus the expected and the observed results produced an accepted outcome. The performance of EFDBN-

STFA is done on two different credit card datasets which are represented entirely in a different domain value. The results proved the empowerment of EFDBN-STFA for credit card fraudulent detection in presence of uncertainty by consequence higher detection rate of fraudulent transaction compared to the existing models. The proposed model restricts the frauds to happen while transactions and promote prevention of fraud in the future.

ACKNOWLEDGMENT

We authors thank all the contributors of this journal for considering the article. I would like to thank my guide for giving his support and encouragement in my work. Also, thank the authors of the references giving me the privilege to cite their article and enhance my knowledge. With the responsibility as Ph.D. Scholar at Nehru Arts and Science College, I thank all the teachers and my friends in giving their valuable ideas and support. Wish this article will be beneficial for future scholars.

REFERENCES

- [1] Wang, S., "A Comprehensive Survey of Data Mining-Based Accounting-Fraud Detection Research." International Conference on Intelligent Computation Technology and Automation, **vol.1**, pp.50-53, 2010.
- [2] Ngai, E.W.T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature", Decision Support System 2010.
- [3] Askari.S and Hussain.A, "Credit Card Fraud Detection Using Fuzzy ID3", IEEE, Computing, Communication and Automation (ICCCA), 446-452, 2017.
- [4] Vaishnavi Nath Dornadula, Geetha S, "Credit Card Fraud Detection using Machine Learning Algorithms", International Conference On Recent Trends In Advanced Computing 2019, ICRTAC 2019 Computer Science 165 (2019) 631-641.
- [5] Imane sadgali, Nawal Sael, Faouzia Benabbou, "Fraud detection in credit card transaction using neural Networks", Association For Computing Machinery. ACM SCA2019, October 2-4.
- [6] Wen-Fang YU and Na Wang, "Research on Credit Card Fraud Detection Model Based on Distance" International Joint Conference On Artificial Intelligence 2009.
- [7] S P Maniraj, Aditya Saini, Swarna Deep Sarkar Shadab Ahmed, " Credit Card Fraud Detection using Machine Learning and Data Science", International Journal of Engineering Research & Technology (IJERT), **Vol. 8** Issue 09, September, 2019.
- [8] Andrea Dal Pozzolo, Giacomo Boracchi, Olivier Caelen, Cesare Alippi, and Gianluca Bontempi, (2017), "Credit Card Fraud Detection: a Realistic Modeling and a Novel Learning Strategy." IEEE Transactions on Neural Networks and Learning Systems · September, PP(99):1-14, 2017.
- [9] Navneet Kr. Kashyap, B.K. Pandey H.L. Mandoria, "Analysis of Pattern Identification Using Graph Database for Fraud Detection". Oriental Journal of Computer Science & Technology, August, **Vol. 9**, No. (2): Pgs. 81-91, 2016.
- [10] Salvatore J. Stolfo, David W. Fan, Wenke Lee and Andreas L., "Prodromidis Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results", In AAAI-97 Workshop on Fraud Detection and Risk Management, 1997.
- [11] V. Dheepa, R. Dhanapal, Behavior Based Credit Card Fraud Detection Using Support Vector Machines, ICTACT Journal on Soft Computing, Volume: 02, Issue: 04, 2012.
- [12] Chuang-Cheng Chiu and Chich-Yuan Tsai, "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection", Proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service, pp. 177-181, 2004.
- [13] Tao Guo and Gui-Yang Li, "Neural Data Mining for Credit Card Fraud Detection", International conference on Machine Learning and Cybernetics, **Vol. 7**, pp. 3630-3634, 2008.
- [14] Suvasini Panigrahi, Amlan Kundu, ShamikrSural and A.K.Majumadar, "Credit Card Fraud Detection: A Fusion Approach Using Dempster Shafer Theory And Bayesian Learning", Information Fusion, **Vol. 10**, No. 4, pp. 354-363, 2009.
- [15] Kaggle Dataset1: <https://www.kaggle.com/mlg-ulb/creditcardfraud>
- [16] Default of credit cards client dataset: I-Cheng Yen, Department of Information Management, Chung Hua University, Taiwan.
- [17] Vergara, J. R. and Estevez, P. A. "A review of feature selection methods based on mutual information." Neural Computing and Applications, 24(1):175-186, 2014.
- [18] L. Visintin, R. H. S. Reiser and B. R. C. Bedregal, "Interval-Valued Intuitionistic Fuzzy Implications," 2011 Workshop-School on Theoretical Computer Science, Pelotas, RS, 2011, pp. 46-52, doi: 10.1109/WEIT.2011.22.
- [19] Feng Shuang, C. L. Philip Chen, "Fuzzy Restricted Boltzmann Machine and Deep Belief Network: A Comparison on Image Reconstruction", IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp 1828-1833, 2017.
- [20] Daranat Tansui, Arit Thammano, "Sea Turtle Foraging Algorithm for Continuous Optimization Problems.", Proceedings of 2016 6th International Workshop on Computer Science and Engineering, (WCSE 2016), 17-19 , pp. 678 -681.