European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 07, Issue 05, 2020

TIME VARYING INERTIA WEIGHT DRAGONFLY ALGORITHM WITH WEIGHTED FEATURE BASED SUPPORT VECTOR MACHINE FOR CREDIT CARD FRAUD DETECTION

G.K. Arun[1], Dr. K. Venkatachalapathy[2]

[1]*Research Scholar, Department of Computer and Information Science, Annamalai University, Chidambaram, India.*

[2]*Professor, Department of Computer and Information Science, Annamalai University, Chidambaram, India.*

[1]*arunnura2370@gmail.com*

[2]*omsumeetha@rediffmail.com*

*Abstract**: Credit card frauds have become significant owing to the rise of latest technologies and the global superhighway of communication. Credit card fraud identification models are necessary for any bank or financial institution to reduce the loss. Several methods are developed for the identification of credit card frauds. To solve this problem, the proposed system designed a Time-Varying Inertia Weight based Dragonfly Algorithm (TVIWDA) with Weighted Feature based Support Vector Machines (WFSVM) for classifying the normal/fraud behavior features. In this proposed work, initially, credit card dataset is taken as an input and feature selection is performed based on Time-Varying Inertia Weight based Dragonfly Algorithm (TVIWDA). According to the selected features, Weighted Feature based Support Vector Machines (WFSVM) approach is utilized for credit card fraud detection. A brief set of experiments were performed to highlight the betterment of the proposed model and the obtained simulation values ensured the better outcome of the proposed model over the compared methods.***

Keywords: Credit card fraud, Dragonfly Algorithm (DA), Weighted Feature based Support Vector Machines (WFSVM) and kernel function

1. INTRODUCTION:

Nowadays, usage of Credit Card has become a common habit globally [1-3]. It is employed for bill payments, online transactions, and so on. Even though the credit card application is useful and beneficial, it still suffers from insecurity and fraud activities [4]. Credit Card based illegal actions result in drastic financial drop. Credit Card fraud is performed in various domains like stealing the cards, generating forged cards, making a duplicate card as same as original, by changing the magnetic strip in the card which has user's details, stealing information from merchant's side. In order to overcome these issues, an effective credit card fraud prediction model has to be developed and it is categorized as 2 common features namely,

Fraud analysis (misuse prediction) and user behavior analysis (anomalous prediction). Initially, the first set of models deals with supervised classification process at transaction level.

Here, transactions are modeled as fraudulent or normal which depends upon the classical information. Then, the dataset is applied for developing classifiers for detecting the condition of newly developed records. Followed by, second module is operated in unsupervised fashion that depends upon the account behavior. In this scheme, a transaction is identifying as fraud activity when there is weird change in user's natural behavior. It is accomplished by monitoring the nature of fraudsters as nature may vary from authenticated users [5]. Thus, there is a requirement of extracting legitimate user behavioral mechanism (user profile) for an account and predict the illegal activities. Hence, the user profile is enclosed with details like account status like merchant types, amount, position, and transaction time. It is named as anomalous prediction.

Recently, Data Mining (DM) methods are well-known and effective models used in predicting fraud activities. The major aim of knowledge discovery and DM models is to identify the unknown patterns from large scale data and interpret only the useful data. Some of the prominently used fraud prediction schemes are Support Vector Machines (SVM), Hidden Markov Model (HMM) [6], Back Propagation Neural Networks (BPNN), and Artificial Neural Network (ANN) [7]. But, SVM based models are not completely automatic as it is user dependent model. While in case of HMM, when the recent transaction is not approved by trained HMM with maximum possible, then it can be referred as fraudulent. Then, BPNN requires prolonged training duration, wider testing, maintenance of parameters namely count of hidden neurons, rate of learning. Finally, NN based models are usually robust; however inaccurate simulation outcome is attained.

[8] projected a model by using Machine Learning (ML) for predicting the credit card fraud activities. Basically, reputed methods are employed and gradually hybrid models are employed where AdaBoost as well as majority voting frameworks are utilized. Followed by, commonly available data set is applied for estimating the system performance and alternate data set utilized from financial institution and examined the fraud actions. Next, noise is included in the data sample by where the model efficiency is calculated. Thus, the performance based on theoretical results implies that majority of voting models gain optimal accuracy rates for predicting credit card fraud. Therefore, it is finalized that the voting model has implied a stable performance even under the existence of noise.

[9] projected Deep autoencoder (Deep AE) which has been applied for extracting optimal features of data from credit card transactions. As a result, softmax software has been included to overcome the class labels problems. The over complete AE is applied for mapping the data to highly dimensional space as well as a sparse approach is applied extensively and provides numerous advantages for classifying different kinds of fraud activities. Then, Deep Learning (DL) method is a remarkable and well-known approach developed and applied for credit card fraud prediction. Also, Deep AE is utilized in certain stages for extracting optimal features of data and classification is performed. Moreover, maximum accuracy and minimum variance are accomplished in the system.

[10] established a real-time modeling and new learning principle for credit card fraud prediction. It is comprised of 3 major objectives. Firstly, using the industrial partner, development of fraud-prediction problem defines the operating state of Fraud-Detection System (FDS) which helps in examining numerous streams of credit card transactions. Moreover, most significant performance metrics are employed in fraud-prediction tasks. Secondly, develop and use new learning procedure which reports the class imbalance, concept

drift, and verification delay. Thirdly, depict the effect of class irregularity and concept drift in real-time data stream with excess transactions, authenticated over time window of 3 years.

[11] established aggregation and feedback models for credit card fraud prediction. The new fraud prediction approach is comprised of 4 states. To supplement the owner's behavioral patterns, apply the cardholders' transaction data for dividing the cardholders as diverse groups like transaction behaviors of users. Followed by, a window-sliding principle is used for collecting the transactions in all groups. Next, filter set of specific behavioral patterns for a cardholder according to the collected transactions and cardholder's historical data. Afterward, training the classifiers for all groups on the basis of behavioral patterns. Consequently, apply the classifier set for predicting the fraud online and if the new transaction is illegal where feedback is regarded in prediction task for resolving the issues of concept drift.

[12] developed advanced Data Mining (DM) model which is composed of Feature Selection (FS) and decision cost for enhancing the accuracy of credit card fraud prediction. Once the appropriate features are selected, extended wrapper mechanism has been applied and ensemble classification is carried out. Therefore, extended FS models are prior feature extraction as well as wrapper scheme under the application of C4.5 Decision Tree (DT). Ensemble classification by cost sensitive DT is processed in decision forest approach. Hence, locally collected fraud prediction dataset is applied for estimating the newly developed framework by means of accuracy, recall, and F-measure as estimation measures and related to previous classifiers like ID3, J48, Naïve Bayes (NB), Bayesian Network and NB tree.

[13] applied a 2-stage neuro-fuzzy expert model for credit card fraud prediction. The incoming transaction is computed using pattern-matching mechanism initially. This component is composed of fuzzy clustering and address-matching methodologies and allocates a value for transaction which depends upon the deviation. Hence, Fuzzy Inference System (FIS) estimates a malicious value by integrating the values and classify the transaction as normal and fraudulent. When the malicious transaction is predicted, Neural Network (NN) equipped with historical transactions are used for validating whether it is correct or incorrect by legitimate user. The proficiency of newly developed approach is validated by performing numerous processes and compared the function with alternate systems.

In this work, initially, credit card dataset is taken as an input and feature selection is performed based on Time-Varying Inertia Weight based Dragonfly Algorithm (TVIWDA). According to the selected features, Weighted Feature based Support Vector Machines (WFSVM) approach is utilized for credit card fraud detection. A brief set of experiments were performed to highlight the betterment of the proposed model and the obtained simulation values ensured the better outcome of the presented method over the compared methods.

2. RESEARCH ELABORATIONS

Actually, Credit card fraud activity is a serious issue in economic services. Also, it results in huge financial losses per annum. Unfortunately, only limited number of works has been developed for credit card fraud detection in real-time domains and further leads to confidentiality problems. The newly developed approach mechanism is TVIWDA with WFSVM classification models used for predicting the illegal actions in credit cards with better accuracy. The working process of newly developed work is depicted in Fig. 1.

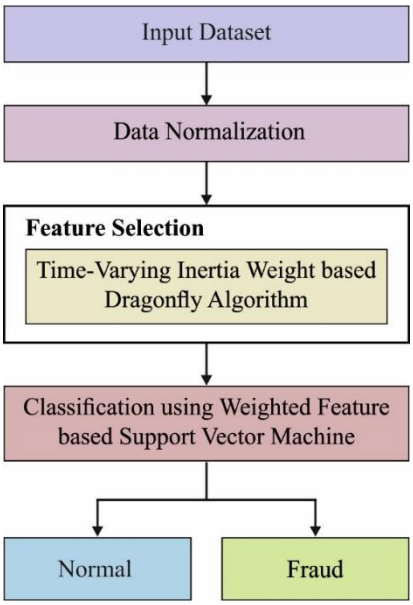

Fig. 1 Flow diagram of the proposed research work

An easy way to comply with EJMCM paper formatting requirements is to use this document as a template and simply type your text into it.

A. *Input*

Here, credit card transaction details are accumulated from https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data) and it is assumed to the input. This dataset is used for classifying users by collective parameters that good or bad risk.

B. *Min -max normalization*

The normalization is defined as the function of normalizing data. Here, data normalization and linear conversion are computed on actual input data. Minimum and maximum values are derived and interchanged with the given expression.

$$v' = \frac{v - \min(A)}{\max(A) - \min(A)} (new\_max(A) - new\_min(A)) + new\_min(A) \quad (1)$$

Where A implies Attribute data, Min(A), Max(A) indicates minimum and maximum value of A correspondingly. v' defines novel measure of entry in data, new_max(A), new_min(A) - max as well as min value of range correspondingly.

C. *Feature selection using Time-Varying Inertia Weight based Dragonfly Algorithm (TVIWDA)*

Feature selection (FS) is processed by applying TVIWDA. Here, DA is a novel metaheuristic optimization approach, which depends upon the swarming hierarchy of dragonflies [14-16]. Similarly, SI-based optimization models have DA where the optimization task is initialized by set of arbitrary solutions for applied optimization issues. Here, parameters in this dataset are referred as dragonflies. Generally, the count of initial dragonfly (M) affects the working function of DA. Measure the classification accuracy of event parameters. Once the positions of attributes are determined from lower and upper boundaries.

$$S_{(i,z)} = -\sum_{j=1}^{N} X_{(i,z)} - X_{(j,z)} \quad (2)$$

Where $X_{(i,z)}$ implies the location of i[th] parameter in the z[th] iteration; $X_{(j,z)}$ denotes the location of jth neighboring attribute in z[th] iteration; N depicts the count of adjacent parameters; and $S_{(i,z)}$ illustrates the separation action for i[th] attribute in z[th] iteration. Hence, alignment motion is estimated by,

$$A_{(i,z)} = \frac{\sum_{j=1}^{N} V_{(j,z)}}{N} \quad (3)$$

Where $V_{(j,z)}$ denotes the velocity of j[th] neighboring variable in z[th] iteration; and $A_{(i,z)}$ signifies the alignment motion for i[th] attribute. Hence, cohesion is quantified by,

$$C_{(i,z)} = \frac{\sum_{j=1}^{N} X_{(j,z)}}{N} - X_{(i,z)} \quad (4)$$

Where $C_{(i,z)}$ refers the cohesion for the i[th] attribute in z[th] iteration. Therefore, food attraction motion is measured by

$$F_{(i,z)} = X_{(foot,z)} - X_{(i,z)} \quad (5)$$

in which $X_{(foot,z)}$ resembles the location of the food source in z[th] iteration; and $F_{(i,z)}$ indicates the food attraction motion for i[th] attribute in z[th] iteration. Also, food is assumed to be a remarkable variable with optimal objective function. Thus, predator distraction is calculated by

$$E_{(i,z)} = X_{(enemy,z)} - X_{(i,z)} \quad (6)$$

Where $X_{(enemy,z)}$ defines the location of predator in the zth iteration; and $E_{(i,z)}$ depicts predator distraction motion. Finally, a step vector implies motion direction for every attribute and represented as,

$$\Delta X_{(i,z+1)} = (s \times S_{(i,z)} + a \times A_{(i,z)} + c \times C_{(i,z)} + f \times F_{(i,z)} + e \times E_{(i,z)}) + w \times \Delta X_{(i,z)} \quad (7)$$

After measuring the step vector, upgraded position vectors can be measured by:

$$X_{(i,z+1)} = X_{(i,z)} + \Delta X_{(i,z)} \quad (8)$$

Isolation of unwanted, cohesion, food attraction, as well as predator weights, diverse divergence and escalation hierarchies are accomplished by using optimization process. The position updating task is followed iteratively till reaching the termination condition. At this point, spots of dragonflies are upgraded:

$$X_{(i,z+1)} = X_{(i,z)} + Levy\ (d) \times X_{(i,z)} \quad (9)$$

Where d refers the count of decision parameters; and Lévy(d) indicates the Lévy flight motion. Also, position updating task is followed repeatedly till reaching the termination criteria.

Time-varying inertia weight

To enhance DA performance, the time-varying inertia weight has been projected. The inertia weight is reduced by concern of time. In general, the search process is composed of maximum inertia weight to improvise the global exploration and the inertia weight is mitigated for local exploration. Therefore, numerical expression for the similar function is represented as shown below:

$$Inertia\ weight\ \omega = \omega_{max} - (\omega_{max} - \omega_{min}) \frac{gen}{GENERATION} \quad (10)$$

Where, ω_1 defines the initial value, ω_2 refers the final values, iter signifies the current iteration, max iter denotes the maximum value for all iterations. Thus, improved step vector for an attribute can be illustrated as,

$$\Delta X_{(i,z+1)} = (s \times S_{(i,z)} + a \times A_{(i,z)} + c \times C_{(i,z)} + f \times F_{(i,z)} + e \times E_{(i,z)}) + \left(\omega_{max} - (\omega_{max} - \omega_{min}) \frac{gen}{GENERATION} \right) \times \Delta X_{(i,z)} \quad (11)$$

It is monitored that the DA and presented inertia weight provides optimal simulation outcome by means of best solution as well as rapid convergence.

D. *Classification using weighted feature based support vector machines*

In this proposed research work, based on the selected attributes the classification is performed by using WFSVM classifier. SVM is one of the supervised ML algorithms that imprinted its best performance in classification approaches [17-21]. It is far more sophisticated and it may be utilized as a discriminative classifier which is demonstrated to be more appropriate and accurate than any other classification model. The Structural Risk Minimization (SRM) principle is good for controlling the generalization ability and is so used in SVM that involves finding optimal separating hyper-plane. This makes it more accurate classifier which is adopted by most of the applications. The finding of a discriminant function f(x) such that y_i = f(x_i) has been the two-class classification problem in general for the N data samples (attributes) (x_1, y_1). . .(x_i, y_i).. . .(x_N, y_N). Moreover f(x) = sgn(w·x − b) is given as linear discriminant function were w·x− b = 0 is assigned for separating hyperplane in the considered data space. The linear discriminant formulation is given as

$$f(x) = sgn(\sum_{i=1}^{l} \alpha_i y_i (x_i \cdot x - b) \quad (12)$$

where training records are counted as l, the training data associated label is given as $y_i \in \{-1, +1\}$ with the range $0 \leq \alpha_i \leq$ C (constant C > 0), and x_i represents the support vectors.
When two classes separated by the surface goes not linear, then the data points are brought into linear separable by transforming data points to higher dimensional space which has been depicted as:

$$f(x) = sgn(\sum_{i=1}^{l} \alpha_i y_i \; K(x_i.x) + b) \quad (13)$$

Where $K(x_i.x)$ is used in transformation of data points which is also called as the kernel function.

All the attributes extracted from the training data sets will be treated alike and it can be found From the Kernel function format K($x_i.x$) of SVM which has an impact on accuracy. The importance of diverse attributes can be considered by adding weights to the Kernel function. In generic, new kernel function can be formulated as K($\omega x_i.\omega x$), here ω represents the weights of attributes of training dataset. The formulation of nonlinear discriminant function is given as,

$$f(x) = sgn(\sum_{i=1}^{l} \alpha_i y_i \; K(\omega x_i.\omega x) + b) \quad (14)$$

This is the enhanced kernel function which is influenced by the weights of the dataset and it is liberated from particular kernel functions. The suitable kernel functions from training data sets are calculated by the system through rough set theory and it varies for different applications. The proposed system works on the basis of the algorithm mentioned below. This algorithm adopts rough set theory for feature ranking and their weight calculation respectively. After the completion of ranking process, the attribute possessing null will be considered as of less importance and is neglected. The attribute weights are ranged from [0, 100]. Thus, WFSVM classifier is used to classify fraudulent and non-fraudulent.

3. RESULTS OR FINDING

The experiments are evaluated utilizing Matlab. The performance of the presented TVIWDA with WFSVM is related to the existing Random-tree-based RF and CART-based RF methods are evaluated using german credit card dataset. The credit card transaction information is collected from https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data). Table 1 illustrates the results analysis of the presented model. The performance of the presented TVIWDA with WFSVM is related to the existing Random-tree-based RF and CART-based RF methods with respect to accuracy and precision which are illustrated in Fig. 2. In this proposed research work, optimal attributes are selected by using Time-Varying Inertia Weight based Dragonfly Algorithm (TVIWDA). It improves detection accuracy. From the results, it is determined that the presented system achieves 97.82% of accuracy whereas existing Random-tree-based RF and CART-based RF attain 96.77% and 97.82% respectively. And also, the precision of the proposed work is 92.62% whereas existing Random-tree-based RF and CART-based RF approach attain 90.27% and 89.46% respectively.

Table 1 Result Analysis of Existing with Proposed TVIWDA with WFSVM Method in terms of Different Measures

Performance Analysis (%)				
Methods	**Accuracy**	**Precision**	**Recall**	**F-Measure**
Random Tree based Random Forest	91.96	90.27	67.89	78.11
CART based Random Forest	96.77	89.46	95.27	96.01
TVIWDA with WFSVM	97.82	92.62	96.45	97.50

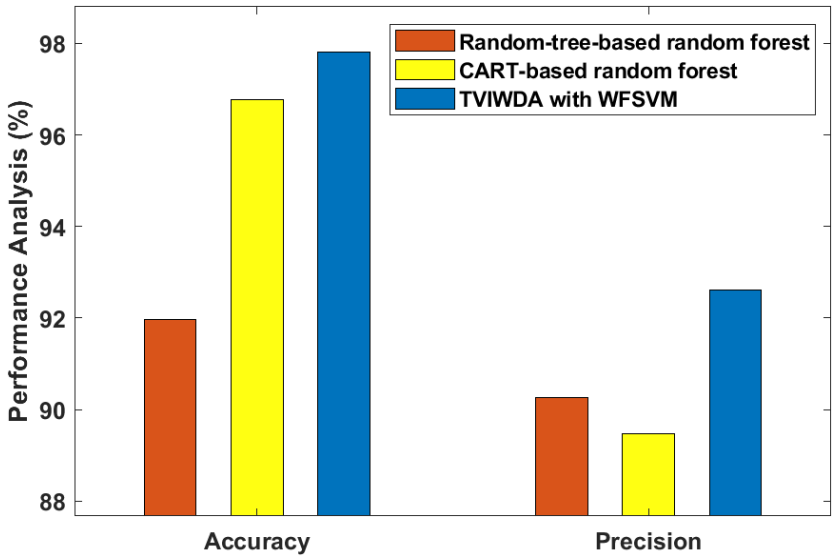

Fig. 2 Accuracy and precision comparison

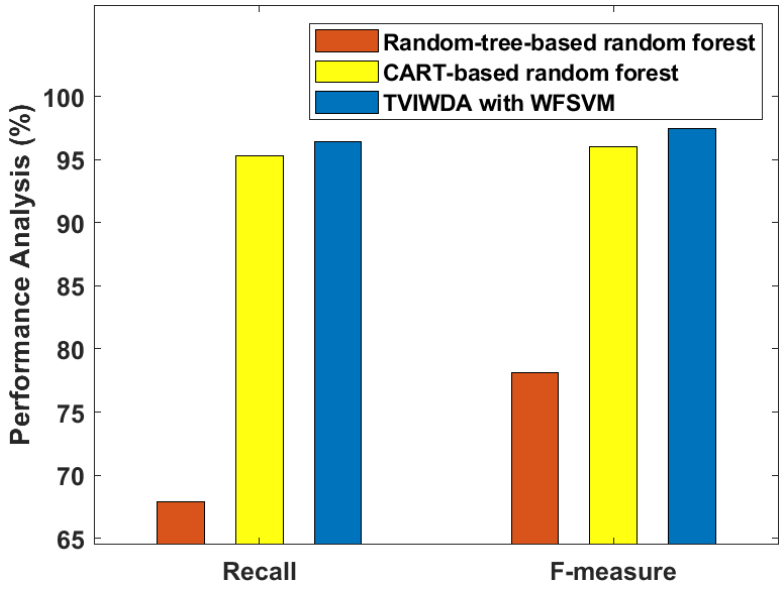

Fig. 3 Recall and F-measure comparison

Recall and f-measure of the proposed TVIWDA with WFSVM approach are compared with the existing Random-tree-based RF and CART-based RF methods which is shown in Fig. 3. In this proposed research work, fraud classification is done by using Weighted Feature based Support Vector Machines (WFSVM) scheme. In WFSVM, enhanced kernel function is influenced by the weights of the dataset to achieve better recall and f-measure. The experimental outcomes demonstrate that the presented model attains 96.45% of recall whereas existing Random-tree-based RF and CART-based RF attains 67.89% and 95.27% respectively. And also, f-measure of the proposed work is 97.5% whereas existing Random-tree-based RF and CART-based RF approaches attain 78.11% and 96.01% respectively.

4. CONCLUSIONS

In this proposed research work, TVIWDA with WFSVM is designed for credit card fraud detection. To improve the detection accuracy, an optimal attribute is selected by using TVIWDA. Then based on the selected attributes, the classification is carried out by using WFSVM classifier. The experimental outcomes show that the presented model attains better performance is related to the existing Random-tree-based RF and CART-based RF methods with respect to accuracy, precision, recall, and F-measure. Although WFSVM gets better outcomes on small set data, there are until a few the problems like imbalanced data. In the future work is an effort on solving these problems.

5. REFERENCES

[1] Zareapoor, M., & Shamsolmoali, P. (2015). Application of credit card fraud detection: Based on bagging ensemble classifier. Procedia computer science, 48(2015), 679-685.
[2] KhyatiChaudhary, JyotiYadav, BhawnaMallick, “A review of Fraud Detection Techniques: Credit Card”, International Journal of Computer Applications Volume 45– No.1 2012.

[3] Michael Edward Edge, Pedro R, Falcone Sampaio, “A survey of signature based methods for financial fraud detection”, journal of computers and security, Vol. 28, pp 3 8 1 – 3 9 4, 2009.
[4] Zhang, X., Han, Y., Xu, W., & Wang, Q. (2019). HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture. Information Sciences.
[5] Lebichot, B., Le Borgne, Y. A., He-Guelton, L., Oblé, F., & Bontempi, G. (2019, April). Deep-learning domain adaptation techniques for credit cards fraud detection. In INNS Big Data and Deep Learning conference (pp. 78-88). Springer, Cham.
[6] Srivastava, A., Kundu, A., Sural, S., and Majumdar, A. (2008). Credit card fraud detection using hidden markov model. IEEE Transactions on Dependable and Secure Computing, 5(1), 37-48.
[7] Maes S. Tuyls K. Vanschoenwinkel B. and Manderick B.; "Credit Card Fraud Detection Using Bayesian and Neural Networks"; Vrije University Brussel – Belgium; 2002.
[8] Kuldeep Randhawa, Chu Kiong Loo, Manjeevan Seera, Chee Peng Lim and Asoke K. Nandi, "Credit card fraud detection using AdaBoost and majority voting," IEEE Access, vol. 6, pp. 14277-14284, 2018.
[9] Zarrabi, H. Kazemi, "Using deep networks for fraud detection in the credit card transaction," IEEE 4th International Conference In Knowledge-Based Engineering and Innovation (KBEI), pp. 0630-0633, 2017.
[10] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2017). Credit card fraud detection: a realistic modeling and a novel learning strategy. IEEE transactions on neural networks and learning systems, 29(8),pp.3784-3797.
[11] Jiang, C., Song, J., Liu, G., Zheng, L., & Luan, W. (2018). Credit card fraud detection: A novel approach using aggregation strategy and feedback mechanism. IEEE Internet of Things Journal, 5(5), pp.3637-3647.
[12] Fadaei Noghani, F., & Moattar, M. (2017). Ensemble classification and extended feature selection for credit card fraud detection. Journal of AI and Data Mining, 5(2), 235-243.
[13] Behera, T. K., & Panigrahi, S. (2017). Credit card fraud detection using a neuro-fuzzy expert system. In Computational intelligence in data mining (pp. 835-843). Springer, Singapore.
[14] Mafarja, M. M., Eleyan, D., Jaber, I., Hammouri, A., & Mirjalili, S. (2017, October). Binary dragonfly algorithm for feature selection. In 2017 International Conference on New Trends in Computing Sciences (ICTCS) (pp. 12-17). IEEE.
[15] Mafarja, M., Heidari, A. A., Faris, H., Mirjalili, S., & Aljarah, I. (2020). Dragonfly algorithm: theory, literature review, and application in feature selection. In Nature-Inspired Optimizers (pp. 47-67). Springer, Cham.
[16] Tharwat, A., Gabel, T., & Hassanien, A. E. (2017, September). Parameter optimization of support vector machine using dragonfly algorithm. In International Conference on Advanced Intelligent Systems and Informatics (pp. 309-319). Springer, Cham.
[17] Goh, R. Y., & Lee, L. S. (2019). Credit scoring: a review on support vector machines and metaheuristic approaches. Advances in Operations Research, 2019.
[18] Aghila Rajagopal, A. Ramachandran, K. Shankar, Manju Khari, Sudan Jha, Yongju Lee, Gyanendra Prasad Joshi, “Fine-tuned Residual Network-based Features with Latent Variable Support Vector Machine-based Optimal Scene Classification Model for Unmanned Aerial Vehicles”, IEEE Access, Volume. 8, Page(s): 118396-118404, June 2020

[19] M. Sivaram, E. Laxmi Lydia, Irina V. Pustokhina, Denis A. Pustokhin, Mohamed Elhoseny, Gyanendra Prasad Joshi, K. Shankar, "An Optimal Least Square Support Vector Machine Based Earnings Prediction of Blockchain Financial Products", IEEE Access, Volume. 8, Page(s): 120321-120330, June 2020.

[20] K. Shankar, Lakshmanaprabu S.K, Deepak Gupta, Andino Maseleno, Victor Hugo C. de Albuquerque, "Optimal Features Based Multi Kernel SVM Approach for Thyroid Disease Classification", The Journal of Supercomputing - Springer, June 2018. In press: https://doi.org/10.1007/s11227-018-2469-4

[21] Lakshmanaprabu SK, K. Shankar, Ashish Khanna, Deepak Gupta, Joel J. P. C. Rodrigues, Plácido R. Pinheiro, Victor Hugo C. de Albuquerque, "Effective Features to Classify Big Data using Social Internet of Things", IEEE Access, Volume.6, page(s):24196-24204, April 2018.