

## **Optimizing High-Performance Weighted Extreme Learning Machine with DAPM for Imbalance Credit Card Fraud Detection**

**Sonam Mittal<sup>1</sup>, Prof. (Dr.) N.K. Joshi<sup>2</sup>**

<sup>1</sup>Department of Computer Science, MIMT, Kota, India  
E-mail: sonammittalkota@gmail.com

<sup>2</sup> Department of Computer Science, MIMT, Kota, India  
E-mail: nkjoshi@modiedukota.org

**Article History:** Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

---

**Abstract**—Fraudulent activities associated with financial transactions are observed in the present scenario, especially with the use of credit cards, at a fast rate. As banking services are rising in digitalization and mobile banking is on the increase in structured written requests, credit card payments rates rise all year, with billions of transactions detected as unfair. For financial institutions to maintain the goodwill of their customers a fraud detection system requiring different detection strategies is therefore extremely important. Researchers and practitioners, using different algorithms, have proposed many methods for fraud detection to find the pattern of fraud. Data mining (DM) algorithms were influential to detect fraudulent transactions by combating fraudsters' attacks on the classic frameworks for preventing fraud. The paper aims to classify fraudulent transactions by Weighted Extreme Learning Machine (WELM) classifiers of 2 Artificial Neural networks (ANN) and 3 separate data sets of Credit Card Fraud (CCF). We use a high-performance Weighted Extreme Learning Machine (HPWELM). The efficiency of the classifiers is calculated based on accuracy, precision, recall & G-mean. The research work has been implemented in Python 3.6. The results are represented in form of tables & snapshots. Results demonstrate that accuracy of the HPWELM classifier has achieved a remarkable improvement in the training and testing phases of the algorithm.

---

**Keywords**—Imbalance data classification, Credit card fraud, Fraud detection, Dandelion Algorithm, Weighted ELM.

---

### **I. INTRODUCTION**

Anomaly detection is an important problem that has been researched within diverse research areas and application domains. Anomaly detection is an important problem that has been researched within diverse research areas and application domains. Anomaly detection is an important problem that has been researched within diverse research areas and application domains. Anomaly detection is an important problem that has been researched within diverse research areas and application domains.

In the last few years, numerous companies have been using data mining to extract valuable and interesting information or trends from their data size, including industry, medical, finances, marketing & health. The estimation of the model, though, depends on the training dataset. The more data are gathered, the more accurate is the model's classification efficiency. In a few years, many methods for addressing the class imbalance were suggested by several scholars. There are 3 major approaches, namely data sampling, selection of features, and ensemble [1], to manage the imbalance classification challenge. Data classification in data mining is called a process of predicting proper intention class for every container in data. The scoring of balanced data sets is very simple and easy, and when the data also isn't balanced it then becomes complicated. Class Imbalance is a problem with machine learning in which the cumulative amount of a (positive) data class is much less than that of a different class of data (negative) [2]. Online fraud detection has a highly imbalanced and large dataset. For instance, there were only 5 cases of fraud within a broad data set of further than 300000 transactions per day, resulting in the task of perceiving very rare fraud spread across a vast amount of real transactions.[3]

In the last decades, there has been an increasing reliance on e-commerce and online payments. With information technology developing every day better over time, there has been a growing number of unlawful attempts at internet purchases around the world causing major financial losses to most organizations and people [4]. The increasing world has financing transactions mainly by transferring the amount through the internet through cashless payments. The rise of transactions has contributed to the production of vast volumes of data. The day-to-day transactions continue to grow, pursued beyond the limits of transactions and variations as big data at high speed. The systematic working of the current FDS (fraud detection system) could also be affected by some fraudsters [4]. Thus, the challenge is to develop the present FDS to satisfy the needs of FDS with optimum accuracy. When the payment is made using a credit card, the fraudsters can misuse their credit cards. The mechanism to identify fraudulent transactions is now important as a real-world task for the FDS and report it to the appropriate people and organizations to decrease the fraudulent rate [6].

Credit card fraud (CCF) has been one of the increasing issues in recent years. A major financial loss impacted the credit card customer, the dealers, and the banks very badly. The primary issue of CCF is the illegal

source of funds in transactions with a payment card such as a credit card. Fraud is an illegitimate way for goods and funds to be obtained. The aim of such illegal transactions may be to obtain goods from an account without paying and to collect unauthorized funds. Fraud is termed as a practice that includes deliberately representing a falsehood to deceive the other party [7]. The identification of such fraud is problematic and may pose a risk to businesses and corporations. Researchers cannot track all transactions in the real world FDS [8]. Hence most of the time it is difficult to identify CCF.

One of the applications for the predictive examination is CCFD. Internet transactions can happen, the user needs essential CC data, like validity, cardholder's name, CC no., & CVV no. The forecast of fraud transactions in CCFD is focused on historical credit card transaction information. Training data on the prediction of fraud is the historical information for card transactions [9]. To avoid CCF by carefully protecting our cards, but if specifics of the card are nevertheless compromised, then fraud should be detected as soon as probable, that fraud is being executed. The methods used for the detection of credit card fraud are categorized into two key categories: fraud analysis or customer action study (anomaly detection).

Machine learning is recognized as one of the best tools for detecting fraud. The method used to recognize fraud on a credit card is grouping and regression. The algorithms are split into two kinds of learning algorithms, supervised and unsupervised [10]. Researchers have provided a great deal to improve the precision of Machine Learning (ML) algorithms & a great deal of analysis is done quickly to improve machine intelligence. Learning[11] is a natural process in the conduct of humans and often becomes a crucial part of machines. Machine learning methods [12] play a major role in the detection of fraud, as they are also used to extract the hidden truths of very large amounts of data.

This paper provides the main contributions:

- 1) For the first time, DAPML, DAPMB, & DAPME have been used effectively to overcome the optimization of HPWELM.
- 2) In 3 imbalanced credit card datasets, three optimized HPWELMs are related to the existing WELM method. Experimental findings indicate that the proposed 3 HPWELMs are more efficient.
- 3) Proposed HPWELMs are implemented to CCFD, also can accomplish better classification performance than previous WELMs.

The following section is structured accordingly: In Section 2, we give thorough literature work about imbalanced CCF. The proposed HPWELMs are given in detail followed by problems that exist in prior work in Section 3. The experiments & results of research for proposed HPWELMs are discussed in section 4. Lastly, in Section 5 conclusion is summarized.

## II. LITERATURE SURVEY

From this literature, it is evident that very little attention is given to the detection of fraud-based transactions activities that identify fraudulent transactions during an online transaction for imbalanced data classification. There are numerous learning techniques to solve the class imbalanced problem.

**Bilal Mirza and Zhiping Lin (2016)** In this work, To class imbalance & concept-float learning, Meta-Cognitive Online Sequential ELM (MOS-ELM) is expected. MOS is applied to self-regulate learning in selecting appropriate learning options for class imbalances & drift issues in MOS-ELM. MOS-ELM efficiency is analyzed and compared with methods under a variety of conditions [13].

**Sanyam Shukla and Bharat Singh Raghuvanshi(2018)** Suggested a new RKWELM (Reduced Kernelized WELM) variant, that is the kernelized WELM variant to extra professionally manage a class imbalance problem. Owing to the arbitrary selection of kernel centroid, RKWELM efficiency differs. This thesis uses the ensemble approach to minimize this variation. Based on the degree of class inequity, this work produces a variety of balanced kernel subsets. This suggested algorithm is evaluated using the imbalanced datasets that the benchmark is downloaded from the repository of KEEL data sets. The experimental findings show that the study suggested is superior to the other classifiers for imbalanced classification problems[14].

**Lean Yu et al. (2018)** Proposed an SVM ensemble knowledge paradigm based on DBN-based resampling. The proposed paradigm is used in the credit classification to address the imbalanced data problem. To provide a more reliable method both for 'small class' performance, the DBN Ensemble strategy has been adopted. The results have been more reasonable across a revenue-sensitive-based revenue matrix.[15].

**S. K. Rath & D. Prusti (2019)** presented an application of usually applied classification methods like ELM, DT, MLP, K-NN & SVM to find accuracy for FD. They also suggested a model of hybridizing DT, SVM & K-NN models that greatly improved forecast exactness. Furthermore, two SOAP & REST services were used to efficiently share data through a heterogeneous platform in this research [16].

**F. Z. El hlouli et al. (2020)** MLP and ELM implemented on the CCF data set, have attempted to identify fraudulent transactions by using two ANN classifiers. These classifiers are measured for their results depends upon precision, accuracy, recall & time of classification. Findings indicate that accuracy of ELM and MLP classifiers is 97.84 percent & 95.46 percent respectively. If not, ELM will foresee new fraudulent transactions very easily[17].

**Sulin Pang et al. (2020)** Built selection algorithm for credit quality rating index of the borrower, & credit quality rating algorithm of a borrower. This study gathers survey data from 7706 internet borrowers. Credit scores, default probability & default loss are measured and the rate of payment of the creditor is evaluated. By computing a confusion matrix, they separated borrowers into seven & five grades. The experimental findings indicate that the total precision of the credit scoring model is 98.5%, with a non-default sample accuracy of 98.9% and a default sample accuracy of 88.3% [18].

**Honghao Zhu et al. (2020)** applied WELM to grip unwarranted categorization troubles. Its 2 parameters are found to affect its high efficiency. The intend of this project is to use different methods of intelligent optimization to optimize the WELM & evaluate its presentation in imbalanced classification. Test outcomes illustration that WELM by the DA will perform thru an improved PSO, bat algorithm, GA, DA & self-learning DA better than WELM by probability-based mutation. The proposed algorithm would also be used for the CCFD. The findings indicate that high detection efficiency can be achieved [19].

**Wen-hui Hou et al. (2020)** To handle data sets, SMOTE is first utilized to balance training set previous than creating applicant classifier pool; then, weighting instrument of DES-MI (multi-class imbalance) was utilized inefficiently to highlight the implication of minority instance when estimating classifier competences. Meta-learning method of META-DES is utilized to an explanation for several criteria, & 2-step selection strategy of DES-KNN is utilized to achieve trade-off b/w competence & variety of classifiers. Fifteen imbalanced data sets in a KEEL repository illustrate that in a region under seven known typical DES algorithms proposed model improves efficiency. Also, the type I error rate of the suggested technique in a real P2P loan dataset showing performance of future credit risk assessment process is less than that for XGboost& LightGBM[20].

**Damodar Reddy Edlaand Diwakar Tripathi (2020)** Presented a new activation function, also an evolutionary method, using the Bat optimization algorithm to obtain optimal weights and biases. Four benchmarked credit scoring datasets with different activation functions are used for the simulations. Simulated results show that EELM (Evolutionary ELM) suggested is more appropriate for credit risk assessment[21].

### III. DA WITH PROBABILITY-BASED MUTATION

In the following section, we present DAPM which is focused on a simple dandelion algorithm (DA). DA is a smart optimization algorithm that was recently proposed and is excellent for resolving problems of function optimization. novel SI algorithm, called DA, is suggested to optimize complex functions globally, stimulated by the behavior of dandelion sowing. In the DA [22], populations of dandelions are broken down into 2 sub-populations, which are appropriate for the seeds and are not ideal for sowing ways for various subpopulations. In the meanwhile, another method of sowing is to perform subpopulation that is appropriate for sowing, to avoid falling in the local optimum. Nevertheless, it slowly converges and simply falls into local optima like other intelligent algorithms. A probability-based DAPM algorithm (DAPM) [23] is proposed to resolve these two flaws. DAPM allows the interchangeable use of both Gaussian & levy mutations based on a particular probabilistic model. DAPM can balance exploitation and exploration. In the three probability models, linear, binomial & exponential mutations are chosen for Gaussian mutations. DAPM can be categorized into 4 main parts, much as other evolutionary algorithms[24]:

#### A. Initialization

DA produces N dandelions randomly as a first-generation dandelion population in the search range.

#### B. Normal Sowing

within a certain sowing radius per dandelion grows dandelion seeds. For minimization difficulty, smaller fitness assessment, more seeds are fashioned no. of kernels is determined by fitness value & sowing radius is modified dynamically. Moreover, methods of calculation for dandelion sowing radius with minimum fitness value also other dandelions will be different later.

#### C. Mutation Sowing

DAPM is former uses Levy mutation and Gaussian mutation. Levy mutation is applied to jump out of maximum locally, so this mutation operation is for minimum fitness dandelion only and is considered the best dandelion.

##### 1) DAPM based on Linear Model

To pick Gaussian or Levy mutation, a linear probability model is implemented. The new mutation strategy is described as follows for best dandelion:

$$X'_{BD} = \begin{cases} X_{BD} * (1 + \text{Gaussian}()), & \text{if } E < p \\ X_{BD} * (1 + \text{Levy}()), & \text{otherwise} \end{cases} \quad (1)$$

Where Levy () & Gaussian () are 2 random numbers produced by following Levy & Gaussian distributions, p is a super parameter that shows the probability of selecting Levy() or Gaussian(). E is dynamically calculated:

$$E = \frac{T_c}{T_{max}} \quad (2)$$

where current & maximum no. of function evaluations are  $T_c$  and  $T_{max}$ . As (2) is linear, we call this type of DAPM linear model also refer to this as DAPML.

**2) DAPM based on Binomial Model**

The following binomial model can also be used to calculate E:

$$E = r_1 * (\frac{T_c}{T_{max}}) + r_2 * (\frac{T_c}{T_{max}})^2(3)$$

Where both  $r_1$  and  $r_2$  are random no. from 0 to 1. We, therefore, call this type of DAPM binomial model, also indicate it as DAPMB.

**3) DAPM based on Exponential Model**

Likewise, the following exponential model can be calculated for E:

$$E = e^{-T_c^2/T_{max}^2} (4)$$

Where E value exponentially changes. We call this type of DAPM an exponential replica, represent as DAPME.

**D. Selection Strategy**

In the next generation, the best dandelion is still kept. Other  $N - 1$  dandelion from the others are selected according to a disruptive selection operator.

**E. Weighted Extreme Learning Machine (WELM)**

ELM [25] is a simple, random algorithm that is efficient to designed for the formation of Single Hidden Layer Feedforward Neural Networks (SLFNs). WELM is allocated among input layer & hidden layer & distortions of hidden nodes, while weights are analytically calculated between the hidden layer & output layer. SLFN architecture may be represented in triple  $(d, m, k)$ , Here  $d$  is no. of input layer nodes, that is a dimension of data input,  $m$  is no. of hidden nodes &  $k$  is no. of output layer nodes, that is no. of classes of input data. Specified preparation set  $D = \{(x_i, y_i) | x_i \in R^d, y_i \in R^k\}, 1 \leq i \leq n$ , SLFN with structure  $(d, m, k)$  may be demonstrated in subsequent eq. (5).

$$f(x_i) = \sum_{j=1}^m \beta_j g(w_j * x_i + b_j) (5)$$

Here weight vector is  $\beta_j$  which is connected to  $j$ th hidden node by output nodes, the activation function is  $g(*)$ , the weight vector is  $w_j$ , which connecting  $j$ th hidden node by input nodes,  $b_j$  is a bias of  $\beta_j$  hidden node. In Equation (5),  $w_j$  &  $b_j$  are randomly created,  $\beta_j$  can be found by resolving subsequent linear systems (6).

$$\sum_{j=1}^m \beta_j g(w_j * x_i + b_j) = y_i (6)$$

In matrix format, equation (6) may be written as

$$H\beta = Y (7)$$

$$H = \begin{bmatrix} g(w_1 * x_1 + b_1) & \dots & g(w_m * x_1 + b_m) \\ \vdots & & \vdots \\ g(w_1 * x_n + b_1) & \dots & g(w_m * x_n + b_m) \end{bmatrix}$$

$$\beta = [\beta_1^T, \beta_2^T, \dots, \beta_m^T]^T$$

$$Y = [y_1^T, y_2^T, \dots, y_m^T]^T$$

H is the SLFN input layer output matrix and is a regularly non-rectangle matrix. Approximate solution (7) may be found by addressing the optimization problem.

$$\min_{\beta} ||H\beta - Y|| (8)$$

Eqn. (8) is given by an approximate solution

$$\hat{\beta} = H^\dagger Y (9)$$

$H^\dagger$  is Moore-Penrose's generalized inverse of matrix H.

**IV. OPTIMIZING HPWELM WITH DAPM PROPOSED METHODOLOGY**

**A. Problem Statement**

Online financial operations are becoming complex and unrestricted, with substantial financial losses for all sides, customers, or organizations. Machine learning research models can learn normal behavior from patterns. These models can identify suspicious customers even if a chargeback is not yet available. However, besides the identification and control of fraudulent online transactions, all of these techniques have some limitations that do not make them very effective. As a result of critical analysis of existing work (based on transactions fraud, fraudulent transaction detection, machine learning-based fraud detection in the online transaction) following problems have been identified that need to be resolved:

- 1) It worked on highly imbalanced datasets that still exist classification problems.
- 2) The hidden truth behind huge amounts of data is urgently necessary to extracted and uncover.
- 3) Existing machine learning-based online fraud detection is not efficient to early and accurate adaptive methods to fraud detection.

- 4) A Weighted tremendous Learning Machine is established not most suitable for imbalanced classification harms.
- 5) The classification performance of WELM is not improved.

**B. Proposed Methodology**

To overcome this problem, an extended version of ELM has been proposed in this research i.e. High-Performance Weighted Extreme Learning Machine (HPWELM). HPWELM has 1 replica structure selection function which events a validation set of different no. of concealed neurons. It takes pre-computed  $\Omega_h, \Omega_t$  as input, & generates solutions  $\beta_k$  for various  $k \in [3, L]$  spaced in the same way on a logarithmic scale. Corroboration data are then iteratively predicted & errors are calculated on the same projected data on all k values. This function only projects the data once most time-consumingly (see V-E section). A minimal validation error selects the optimal no. of hidden neurons.

Both  $L^1$  &  $L^2$  regularization is obtainable in ELM. The MRSR, a multi-output version of LARS, allows the regularization of  $L^1$ . It classifies the neurons from which the problem is the most significant. With such ranked neurons, all model structure selection approaches work improved with compensation of additional operating time.

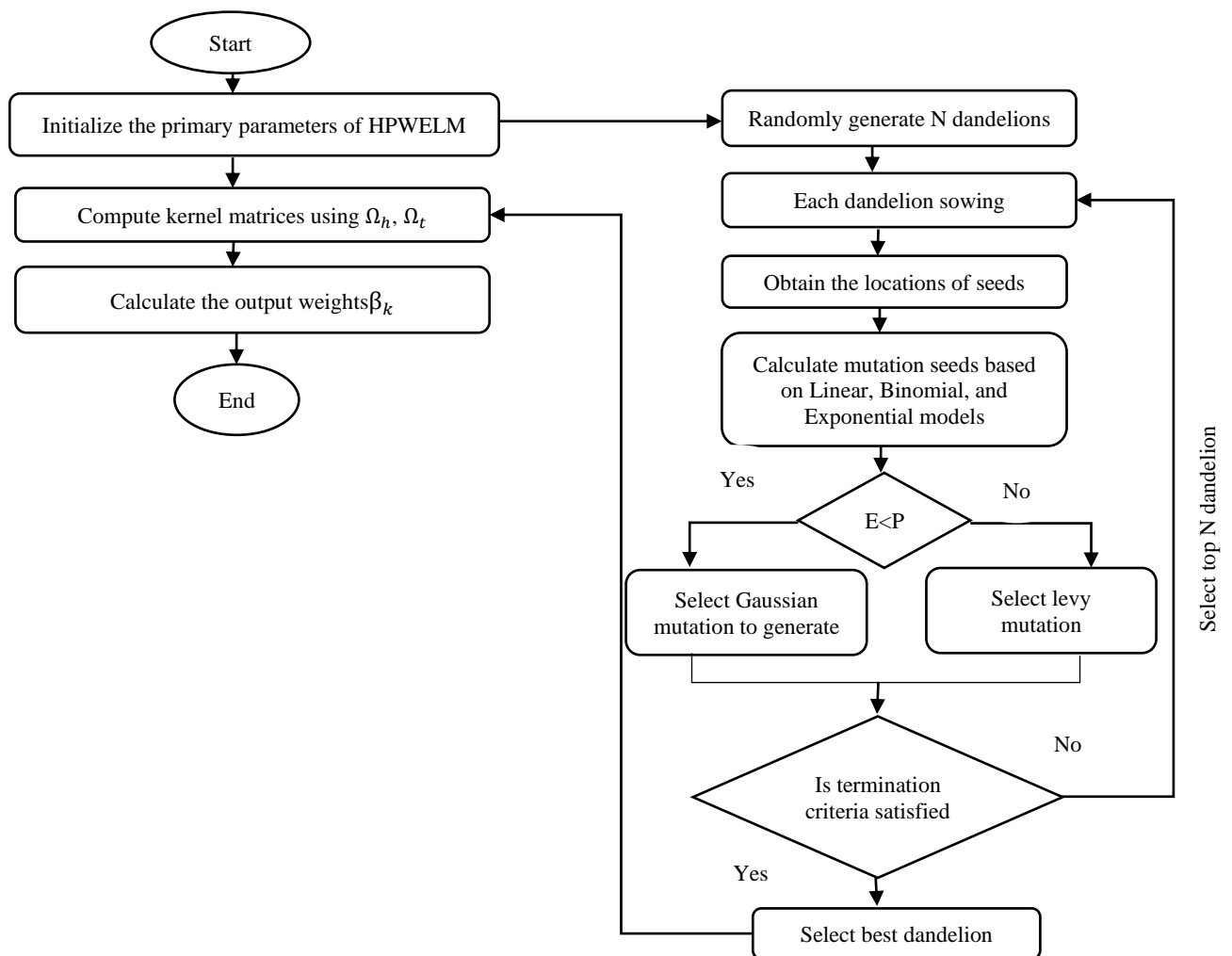


Fig. 1. Flow chart of optimizing HPWELM with DAPM.

The toolbox comprises a modified MRSR algorithm based on another approach for the performance of  $L^1$  regularisation. Initial MRSR includes a fraction with  $O(2c)$  difficulty with approbation tonno. of outputs  $c$ . It takes remarkable runtime with ten outputs in addition to renders the process with over fifteen outputs impractically slow. The convolution with an enhanced version is linear to the number of outputs.  $L^2$  can be regularised for a wider range of issues, as well as AE for ELMs in Image Processing.

$L^2$  regularization is a class restriction of ELM named alpha that may be distorted easily. A noteworthy advantage by  $L^2$  regularization is making ill-conditioned ELM answerable. One can utilize a single-variable optimization technique to discover the optimal value of  $L^2$  parameter.

Toolbox HP-ELM supports 3 classification types: multi-class (one accurate class for every sample), multi-label (arbitrary no. precise classes per model), weighted multi-class (per class has weight, it will be not controlled by another person of no. of samples in class). In every case, true sample classes are set to one & inappropriate classes are set to 0. ELM targets are to limit a feature of each class, (binary classes are multi classes). For the correct working of classification error and model selection structures, this Convention is required.

**V. EXPERIMENT RESULTS AND ANALYSIS**

In this section, the optimized high-performance weighted ELM(HPWELM) is applied to credit card fraud detection. These experiments have done using Jupyter Notebook in Python programming. To verify the accuracy of our proposed algorithms, we select three publicly accessible datasets. It is Loan Prediction [26], Creditcardsvpresent [27] and the Credit Card Client Default [28]. There are 614 samples of that 192 are positive samples in the loan calculation dataset. An overall of 3075 samples, 448 of which were positive samples, are included in the Creditcardsvpresent data set. There are a total of 30,000 samples in the Credit Card client dataset, of which 6636 are positive samples.

**A. Performance Measurements**

The results of these models are assessed by subsequent metrics: G-mean, Accuracy (Acc), Precision & Recall:

$$Acc = \frac{T_p + T_N}{T_p + F_p + T_N + F_N} \quad (10)$$

$$Precision = \frac{T_p}{T_p + F_p} \quad (11)$$

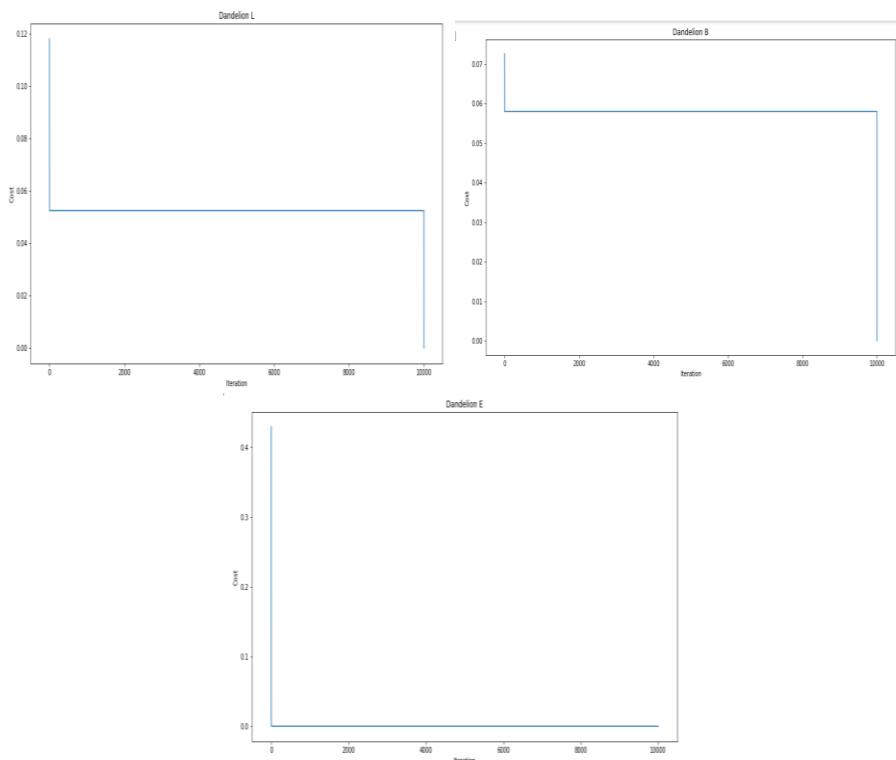
$$Recall = \frac{T_p}{T_p + F_N} \quad (12)$$

$$G\text{-mean} = \sqrt{Precision \times Recall} \quad (13)$$

Here,  $TP$  is no. of true positive cases,  $TN$  is no. of true negative cases,  $FP$  is no. of false-positive cases, also  $FN$  is no. of false-negative cases.

**B. Result Analysis**

**1) HPWELM with DAPM for default of credit card clients.xls dataset**



(a)(b)(c)

Fig.2. DAPM results for default of credit card clients: (a) linear (b) binomial (c) exponential  
 Figures 2 (a), (b), (c) are a representation of credit card clients for DAPML, DAPMB, and DAPME respectively proposed research work. The graphs in the above figures show the variation in the cost concerning the iterations. The training and test dataset resultsof credit card clients for both approaches in different parameters have represented in tabular form in table I and II.

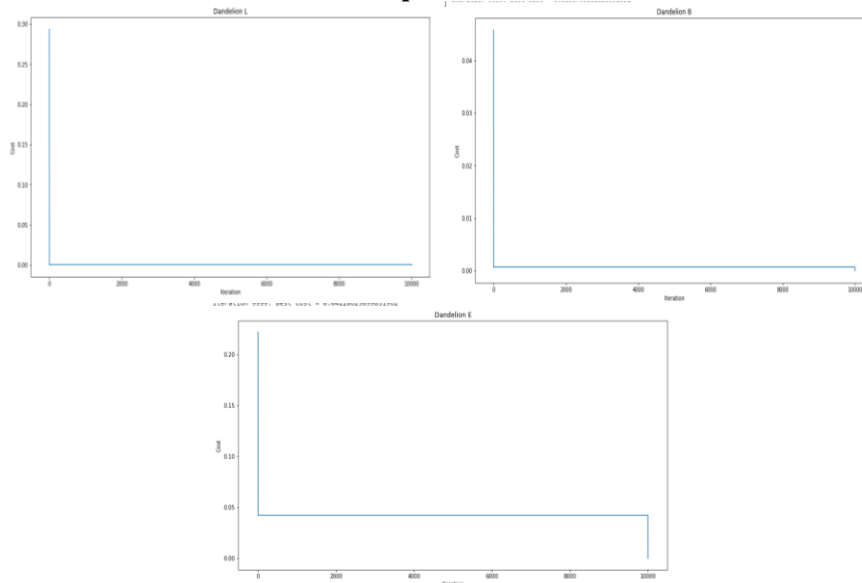
Table I. Comparing the performance parametersofWELM and HPWELM on *credit card clients* training dataset

Algorithms/ Parameters	Training									
	WELM					HPWELM				
	Acc	Prec	Recall	F1- measure	Gmean	Acc	Prec	recall	F1- measure	Gmean
DAPML	77.56	47.10	8.71	14.69	0.60	79.72	77.20	55.75	55.06	0.77
DAPMB	78.01	50.59	16.82	25.24	0.63	79.37	76.72	55.56	54.68	0.76
DAPME	79.30	63.27	16.32	25.95	71.27	79.5	76.51	55.17	54.08	0.76

Table II. Comparing the performance parametersofWELM and HPWELM on *credit card clientstest* dataset

Algorithms/ Parameters	Testing									
	WELM					HPWELM				
	Acc	Prec	recall	F1- measure	Gmean	Acc	Prec	recall	F1- measure	Gmean
DAPML	77.6	44.3	9.63	15.83	0.59	79.33	74.22	55.18	54.21	0.74
DAPMB	77.7	50.1	16.87	25.25	0.63	79.37	76.72	55.02	54.2	0.76
DAPME	79.5	61.86	15.79	25.15	0.70	79.35	74.83	55.16	54.13	0.74

2) HPWELM with DAPMfor creditcardsvpresentdataset



(a)(b)(c)

Fig.3. DAPM results for creditcardsvpresent: (a) linear (b) binomial (c) exponential  
 Figures 3 (a), (b), (c) are a representation of credit card clients for DAPML, DAPMB, and DAPME respectively of the HPWELM. The graphs in the above figures show the variation in the cost concerning the iterations. The training and test dataset resultsof creditcardsvpresent for both WELM and HPWELM approaches with different performance parameters have represented in tabular form in table III and IV.

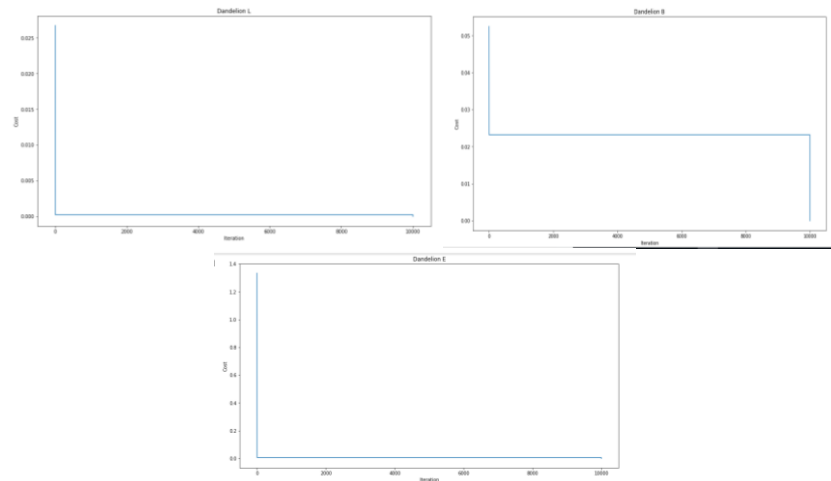
Table III. Comparing the performance parameters of WELM and HPWELM on *creditcardsvpresent* training dataset

Algorithms/ Parameters	Training									
	WELM					HPWELM				
	Acc	Prec	recall	F1- measure	Gmean	Acc	Prec	recall	F1- measure	Gmean
DAPML	96.8	95.3	82.05	88.21	0.96	96.95	96.73	90.9	93.53	0.96
DAPMB	85.4	100	5.02	9.5	0.92	97.03	97.01	90.75	93.56	0.97
DAPME	85.4	100	0.55	1.1	0.92	97.11	96.95	91.27	93.84	0.96

Table IV. Comparing the performance parameters of WELM and HPWELM on *creditcardsvpresenttest* dataset

Algorithms/ Parameters	Testing									
	WELM					HPWELM				
	Acc	Prec	recall	F1-measure	Gmean	Acc	Prec	recall	F1-measure	Gmean
DAPML	96.4	97.4	79.38	87.5	0.96	98.04	98.32	93.51	95.73	0.98
DAPMB	89.9	100	11.42	20.51	0.94	97.72	96.77	94.3	95.49	0.96
DAPME	80.16	68.4	12.9	21.71	0.28	97.23	96.77	91.85	94.11	0.96

3) HPWELM with DAPM for loan predictiondataset



(a)(b)(c)

Fig. 4. DAPM results for loan prediction: (a) linear (b) binomial (c) exponential

Figures 4 (a), (b), (c) are a representation of credit card clients for DAPML, DAPMB, and DAPME respectively of the proposed methodology. The graphs in the above figures show the variation in the cost concerning the iterations. The training and test dataset resultsof loan prediction for both WELM and HPWELM approaches with different performance parameters have represented in tabular form in table V and VI.

Table V. Comparing the performance parameters of WELM and HPWELM on for *Loan Prediction* training dataset

Algorithms / Parameter s	Training									
	WELM					HPWELM				
	Acc	Precisio n	recal l	F1-measur e	Gmea n	Acc	Precisio n	recal l	F1-measur e	Gmea n
DAPML	68.2	68.1	100	81.07	0.82	81.05	81.79	74.28	76.28	0.81
DAPMB	78.4	78.0	94.91	85.68	0.79	80.85	83.35	73.05	75.34	0.82
DAPME	73.31	72.3	99.11	83.65	0.80	80.65	83.24	73.54	75.76	0.81

Table VI. Comparing the performance parameters of WELM and HPWELM on for *Loan Prediction* test dataset

Algorithms/ Parameters	Testing									
	WELM					HPWELM				
	Acc	Prec	recall	F1-measure	Gmean	Acc	Prec	recall	F1-measure	Gmean
DAPML	72.3	72.13	100	83.81	0.84	74.79	72.9	66.39	67.24	0.72
DAPMB	79.6	79.4	96.59	97.18	80.3	82.11	82.05	74.89	77.29	0.8
DAPME	70.7	70.68	97.62	82	0.71	71.54	68	64.47	64.74	0.68

C. Comparative Representation



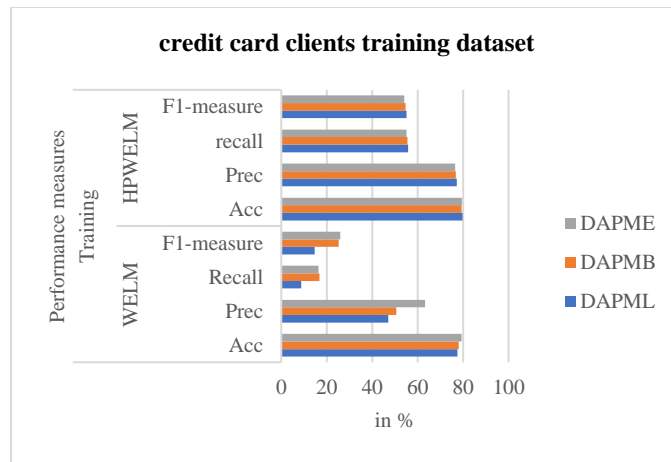


Fig. 5. Performance measurements for credit card client dataset for training set for both approaches

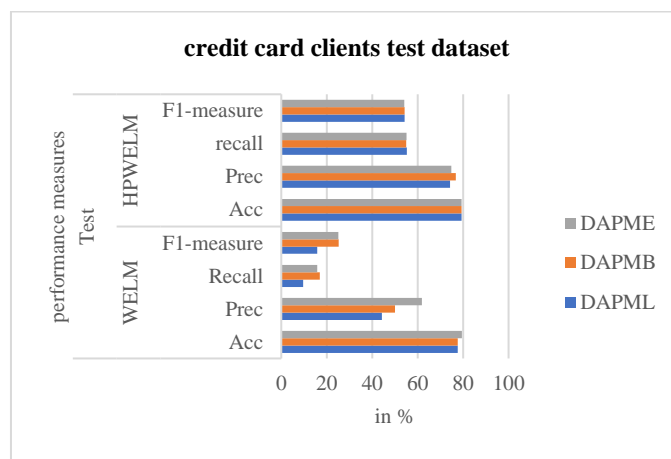


Fig. 6. Performance measurements for credit card client dataset for test set for both approaches

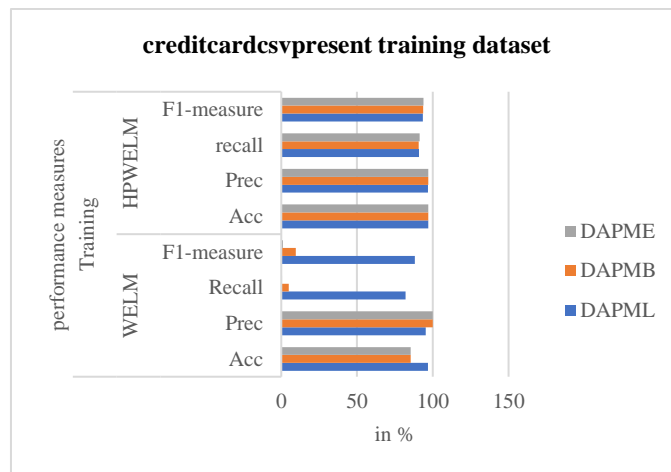


Fig. 7. Performance measurements for creditcardsvpresent dataset for training set for both approaches

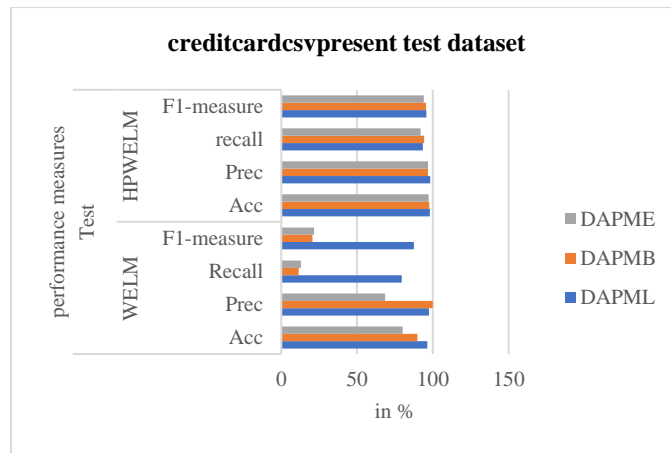


Fig. 8. Performance measurements for creditcardsvpresent dataset for test set for both approaches

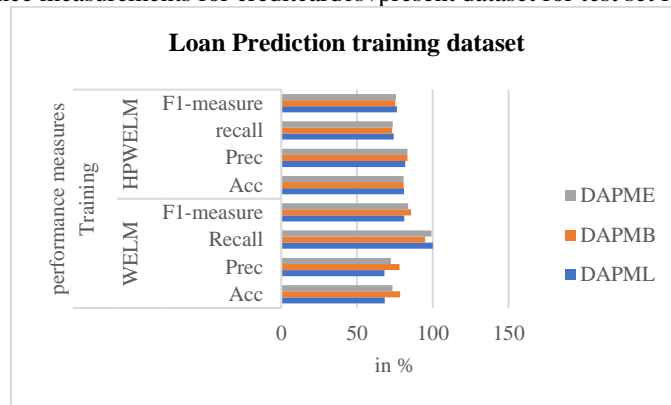


Fig. 9. Performance measurements for loan prediction dataset for training set for both approaches

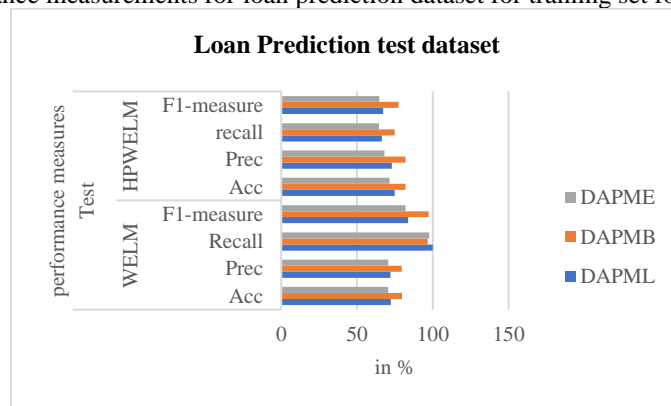


Fig. 10. Performance measurements for loan prediction dataset for test set for both approaches

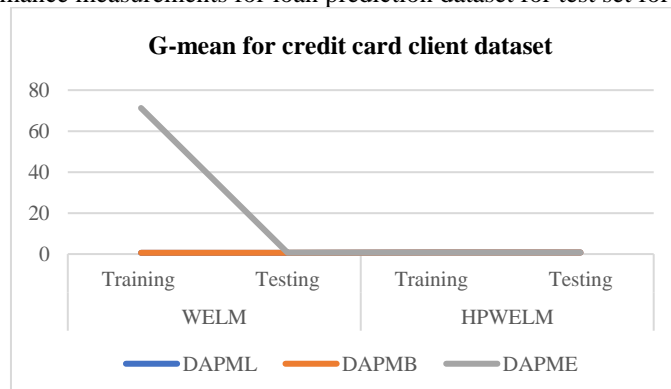


Fig. 11. G-mean for credit card client dataset for training and test set for both approaches

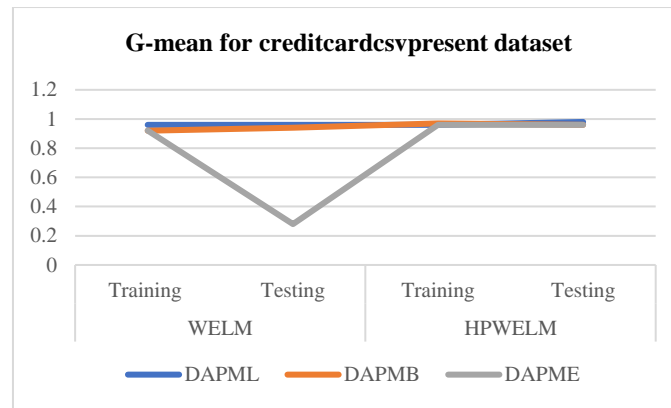


Fig. 12. G-mean for creditcardscvpresent dataset for training and test set for both approaches

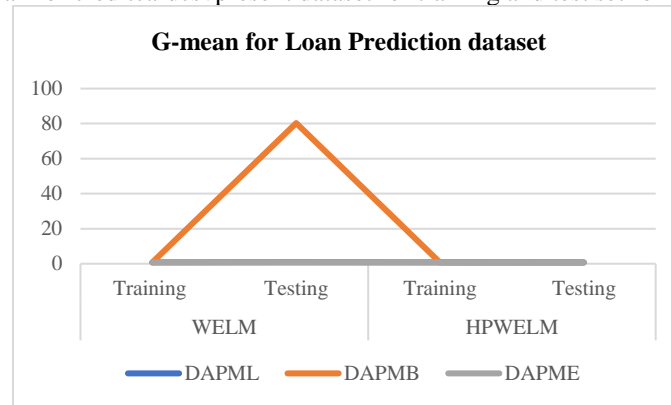


Fig. 13. G-mean for loan prediction dataset for training and test set for both approaches

## VI. CONCLUSION

For banks & card issuers, a credit card designed to provide the perpetrator with an unlawful gain is a major difficulty. Via way of buying requirements or to examine and not paying for them awaiting later card fraud, billions of dollars are lost per year. There is a lack of study into the evaluation of real-world transaction data because of confidentiality issues. Therefore, whether a transaction is fraudulent is incredibly relevant. As an annual rise in credit card fraud increases the need for advanced fraud detection technologies, a major barrier is to be removed in credit card transaction data sets by researchers in the quest to find innovative solutions. The main problems, techniques, and challenges in fraud detection have been discussed well. In this article, we create use of three to improve dandelion algorithms by a probability-based modification to make the best parameters of HPWELM also give 3 optimized HPWELMs for problems of imbalanced classification. Experimental conclusions reveal that 3 HPWELMs optimized to complete greater classification accuracy than the three imbalanced credit card datasets compared to the algorithms. The findings show its effectivity and the work also extends to CCFD by suggesting HPWELMs.

## References

- Hanskunatai, A. (2018). A New Hybrid Sampling Approach for Classification of Imbalanced Datasets. 2018 3rd International Conference on Computer and Communication Systems, ICCCS 2018. <https://doi.org/10.1109/CCOMS.2018.8463228>
- Shukla, P., & Bhowmick, K. (2018). To improve classification of imbalanced datasets. Proceedings of 2017 International Conference on Innovations in Information, Embedded and Communication Systems, ICIIECS 2017. <https://doi.org/10.1109/ICIIECS.2017.8276044>
- Elena-Adriana Minastireanu and Gabriela Mesnita, "An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection", *Informatica Economică* vol. 23, no. 1, 2019, pp. 5-16. DOI: 10.12948/issn14531305/23.1.2019.01
- Rudresha C.E, "Cashless Transaction in India: A Study," *International Journal of Scientific Development and Research (IJS DR)*, Volume 4, Issue 2, 2019, pp. 62-67.
- Isabella, s & Srinivasan, Sujatha & G., Suseendran. (2020). An Efficient Study of Fraud Detection System Using MI Techniques. 10.1007/978-981-15-3284-9\_8.
- Kuldeep Joshi, "Cashless Transaction Challenges and Remedies," *International Journal of Creative Research Thoughts (IJCRT)*, 2018, pp. 167-172.

7. Lina Fernandes, "Fraud In Electronic Payment Transactions: Threats and Countermeasures," *Asia Pacific Journal of Marketing & Management Review*, Vol.2, Issue 3, March 2013, ISSN 2319-2836, pp. 23-32
8. Jalinus, N., Nabawi, R. A., & Mardin, A. (2017). The Seven Steps of Project Based Learning Model to Enhance Productive Competences of Vocational Students. <https://doi.org/10.2991/ictvt-17.2017.43>
9. Agrawal A.N Dermal N., "Credit card fraud detection using SVM and Reduction of false alarms," *International Journal of Innovations in Engineering and Technology (IJET)*, vol. 7, no. 2, pp. 176-182, 2016.
10. K. Anupriya, Mrs. C. Kanimozhi, 2016, Scam Detection for Online Shopping using Deep Learning, *International Journal of Engineering Research & Technology (IJERT)*, Volume 4, Issue 11, 2016, pp. 1-6.
11. Neha Sharma, Reecha Sharma, Neeru Jindal, "Machine Learning and Deep Learning Applications-A Vision," *Global Transitions Proceedings*, 2021, ISSN 2666-285X, <https://doi.org/10.1016/j.glt.2021.01.004>.
12. Sadgali, N. Sael, F. Benabbou, "Performance of machine learning techniques in the detection of financial frauds," *Procedia Computer Science*, Volume 148, 2019, pp. 45-54, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2019.01.007>.
13. Bilal Mirza, Zhiping Lin, "Meta-cognitive online sequential extreme learning machine for imbalanced and concept-drifting data classification," *Neural Networks*, Volume 80, 2016, pp. 79-94, <https://doi.org/10.1016/j.neunet.2016.04.008>.
14. Bhagat Singh Raghuwanshi, Sanyam Shukla, "UnderBagging based reduced Kernelized weighted extreme learning machine for class imbalance learning," *Engineering Applications of Artificial Intelligence*, Volume 74, 2018, Pages 252-270, <https://doi.org/10.1016/j.engappai.2018.07.002>.
15. Lean Yu, Rongtian Zhou, Ling Tang, Rongda Chen, "A DBN-based resampling SVM ensemble learning paradigm for credit classification with imbalanced data," *Applied Soft Computing*, Volume 69, 2018, pp. 192-202, <https://doi.org/10.1016/j.asoc.2018.04.049>.
- D. Prusti and S. K. Rath, "Web service based credit card fraud detection by applying machine learning techniques," *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, Kochi, India, 2019, pp. 492-497, doi: 10.1109/TENCON.2019.8929372.
- E. Z. El hlouli, J. Riffi, M. A. Mahraz, A. El Yahyaouy and H. Tairi, "Credit Card Fraud Detection Based on Multilayer Perceptron and Extreme Learning Machine Architectures," *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Fez, Morocco, 2020, pp. 1-5, doi: 10.1109/ISCV49265.2020.9204185.
16. Sulin Pang, Xianyan Hou, Lianhu Xia, "Borrowers' credit quality scoring model and applications, with default discriminant analysis based on the extreme learning machine," *Technological Forecasting and Social Change*, Volume 165, 2021, 120462, <https://doi.org/10.1016/j.techfore.2020.120462>.
17. Honghao Zhu, Guanjun Liu, Mengchu Zhou, Yu Xie, Abdullah Abusorrah, Qi Kang, "Optimizing Weighted Extreme Learning Machines for imbalanced classification and application to credit card fraud detection," *Neurocomputing*, Volume 407, 2020, pp. 50-62, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2020.04.078>.
18. Wen-hui Hou, Xiao-kang Wang, Hong-yu Zhang, Jian-qiang Wang, Lin Li, "A novel dynamic ensemble selection classifier for an imbalanced data set: An application for credit risk assessment," *Knowledge-Based Systems*, Volume 208, 2020, 106462, <https://doi.org/10.1016/j.knosys.2020.106462>.
19. Diwakar Tripathi, Damodar Reddy Edla, Venkatanareshbabu Kuppili, Annushree Bablani, "Evolutionary Extreme Learning Machine with novel activation function for credit scoring," *Engineering Applications of Artificial Intelligence*, Volume 96, 2020, 103980, <https://doi.org/10.1016/j.engappai.2020.103980>.
20. Xiguang Li, Shoufei Han, Liang Zhao, Changqing Gong, Xiaojing Liu, "New Dandelion Algorithm Optimizes Extreme Learning Machine for Biomedical Classification Problems", *Computational Intelligence and Neuroscience*, vol. 2017.
21. H. Zhu, G. Liu, M. Zhou, Y. Xie and Q. Kang, "Dandelion Algorithm With Probability-Based Mutation," in *IEEE Access*, vol. 7, pp. 97974-97985, 2019, doi: 10.1109/ACCESS.2019.2927846.
22. X. Li, S. Han, L. Zhao, C. Gong, and X. Liu, "New dandelion algorithm optimizes extreme learning machine for biomedical classification problems," *Comput. Intell. Neurosci.*, vol. 2017, Sep. 2017, Art. no. 4523754.
23. 25 new learning scheme of feedforward neural networks. *Proceedings of International Joint Conference on Neural Networks (IJCNN2004)*, 2004, vol. 2, pp. 985-990.
24. [https://github.com/Paliking/ML\\_examples/blob/master/LoanPrediction/train\\_u6ljuX\\_CVtuZ9i.csv](https://github.com/Paliking/ML_examples/blob/master/LoanPrediction/train_u6ljuX_CVtuZ9i.csv)
25. <https://github.com/gksj7/creditcardsvpresent>.
26. <http://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients>.