Tech Science Press

Optimized Stacked Autoencoder for IoT Enabled Financial Crisis Prediction Model

Mesfer Al Duhayyim[1], Hadeel Alsolai[2], Fahd N. Al-Wesabi[3,4], Nadhem Nemri[3], Hany Mahgoub[3], Anwer Mustafa Hilal[5], Manar Ahmed Hamza[5,*] and Mohammed Rizwanullah[5]

[1]Department of Natural and Applied Sciences, College of Community-Aflaj, Prince Sattam bin Abdulaziz University, Saudi Arabia
[2]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Saudi Arabia
[3]Department of Computer Science, King Khalid University, Muhayel Aseer, Saudi Arabia
[4]Faculty of Computer and IT, Sana'a University, Sana'a, Yemen
[5]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
*Corresponding Author: Manar Ahmed Hamza. Email: ma.hamza@psau.edu.sa
Received: 26 June 2021; Accepted: 30 July 2021

Abstract: Recently, Financial Technology (FinTech) has received more attention among financial sectors and researchers to derive effective solutions for any financial institution or firm. Financial crisis prediction (FCP) is an essential topic in business sector that finds it useful to identify the financial condition of a financial institution. At the same time, the development of the internet of things (IoT) has altered the mode of human interaction with the physical world. The IoT can be combined with the FCP model to examine the financial data from the users and perform decision making process. This paper presents a novel multi-objective squirrel search optimization algorithm with stacked autoencoder (MOSSA-SAE) model for FCP in IoT environment. The MOSSA-SAE model encompasses different subprocesses namely preprocessing, class imbalance handling, parameter tuning, and classification. Primarily, the MOSSA-SAE model allows the IoT devices such as smartphones, laptops, etc., to collect the financial details of the users which are then transmitted to the cloud for further analysis. In addition, SMOTE technique is employed to handle class imbalance problems. The goal of MOSSA in SMOTE is to determine the oversampling rate and area of nearest neighbors of SMOTE. Besides, SAE model is utilized as a classification technique to determine the class label of the financial data. At the same time, the MOSSA is applied to appropriately select the 'weights' and 'bias' values of the SAE. An extensive experimental validation process is performed on the benchmark financial dataset and the results are examined under distinct aspects. The experimental values ensured the superior performance of the MOSSA-SAE model on the applied dataset.

Keywords: Financial data; financial crisis prediction; class imbalance problem; internet of things; stacked autoencoder

1 Introduction

Financial crisis over the globe has highlighted the responsibility of financial connectedness as a probable resource of systematic risks and macro-economic instabilities. The crisis has showcased the requirement of clearly understanding whether a rise in connectedness resulted in a maximum possibility of a financial crisis. Forbes et al. [1] defined that contagion as a noteworthy rise in market relations next to a considerable shock to all economic channels. Particularly, the contagion defines the status where the dispersion of financial disturbance can be observed among countries or from a particular economic channel to other. Besides, when two indices demonstrate a maximum level of co-movement at the time of tranquil period, and they endure to be extremely correlated next to a shock in the market, it does not comprise a financial contagion. Several other studies defined that contagion as an extreme rise in the relationship between the countries triggering the crisis and all other countries.

In previous years, the increase in the financial crisis of the companies takes place over the globe, where the companies have started to concentrate on the concept of financial crisis prediction (FCP) [2]. The process of FCP is highly needed for modelling a trustworthy, accurate, and early predictive technique for forecasting the significant risks of the company's economic status earlier. The FCP is generally considered as the binary classification process that can be resolved in rational ways. The results of the classifier technique undergo categorization into two kinds namely, the one indicating the failing condition of an organization and the second one representing the non-failing condition of the same organization. The input to the classifier technique is generally the statistical ratio attained from the financial statement of the original organization. Presently, several classifier techniques have been presented by the use of different domain knowledge for FCP.

Statistics-based and machine learning (ML) techniques are commonly utilized to find out the important factors of the FCP. On the domain of FCP, the ML models are applied in several manners [3]. It is used for the modal construction process for validating the techniques for the identification of financial crisis. The major assumption is that the financial parameters extracted from the open-access financial stamen like financial ratio comprises large amount of data relating the financial details of the company and is helpful for the FCP process [4]. The FCP is a tedious procedure to utilize the interconnected economic details and other information about the company's strategic affordability to active data for the construction of a novel technique. Along with the AI and database concepts, data mining techniques are widely used in several domains. In FCP, data mining approaches are commonly available in two manners namely early warning and decision-making models. It will be helpful for taking necessary action to eliminate the financial failure of the organization. Followed by, the FCP techniques help the decision-making experts of the company to evaluate and select the firm to collaborate or invest. In order to deeply understand a large amount of data, it is needed to afford high computation time and processing. At the same time, class imbalance problem poses a major issue in FCP which greatly affects the predictive outcome.

This paper presents a novel multi-objective squirrel search optimization algorithm with stacked autoencoder (MOSSA-SAE) model for FCP in internet of things (IoT) environment. The MOSSA-SAE model encompasses diverse subprocesses namely pre-processing, class imbalance handling, parameter tuning, and classification. In addition, SMOTE technique is employed to handle class imbalance problems. The goal of MOSSA in SMOTE is to compute the oversampling rate and area of nearest neighbors of SMOTE. Also, SAE model is exploited as a classification technique to determine the class label of the financial data. Simultaneously, the MOSSA is utilized for the appropriate election of the 'weights' and 'bias' values of the SAE. A comprehensive experimental analysis is carried out on the benchmark financial dataset and the results are inspected under different aspects.

2 Related Works

Several researchers have developed a large number of FCP models by the exploitation of the details of the financial company. Presently, ML based artificial neural network (ANN) model is employed as a popular method for the prediction procedure. A new support vector machine (SVM) based ensemble model is developed for FCP, which is mainly dependent on the particular SVM classifier technique [5]. The resultant outcome demonstrated that the SVM based ensemble model offers the particular SVM classifier. Another SVM technique designed in [6] with classical model revealed that the raking method varies with respect to accuracy and precision of Taiwan stock exchange. In Geng et al. [7], a new FCP model is solely developed for Chinese corporations. The experimental outcome denoted that the neural network (NN) is considered to be an effective model compared to traditional ML models like decision tree (DT) and SVM. On the other hand, metaheuristic algorithms are also started to be employed for resolving the binary classification problem.

A GA based feature selection (FS) with FCP technique has been developed in [8] to achieve minimal cost with higher accurateness. At the same time, a new fruitfly optimization based FCP model has been presented in [9] by the integration of general regression NN [10]. The resultant values are interconnected to general and multiple regression NN. The presented methodology is explored to be the effective method with respect to the convergence rate and the ability of the classification model. An enhanced boosting model terms as FS-Boosting is devised as feature selection (FS)-Boosting [11] model. By the application of FS in the boosting process, the presented technique is found to be effective and results in higher accurateness. Another type of FS technique is developed in Lin et al. [12], which integrates the popularity level of the experts and wrapper based FS model. Primarily, the financial standards are partitioned into a total of seven parts and afterward the wrapper model is applied for choosing the feature subsets. The experimental values demonstrated that the presented model outperforms the other FS algorithms to increase the classification performance.

An efficient Ant Colony Optimization (ACO) algorithm for the design of FCP model is developed in [13] and the effectiveness has been analyzed on qualitative as well as quantitative datasets. In Metawa et al. [14], other FCP technique is developed by the use of ACO algorithm, which makes use of ACO for FS and classification process. Another FCP model on the basis of clustering with classification model is developed in Uthayakumar et al. [15] by the use of improved K-means based clustering and a fitness scaling chaotic genetic ant colony algorithm (FSCGACA) classification technique. Firstly, the introduced clustering principle groups the data and removes the misclustered data instances. Subsequently, a ruling technique is elected to make it appropriate for the intended dataset. At last, the FSCGACA is utilized for the election of optimum values of the ruling technique. Sankhwar et al. [16] designed a novel FCP technique utilizing improved grey wolf optimization (IGWO) and fuzzy neural classifier (FNC) models. The IGWO technique solely depends upon the amalgamation of the GWO technique and tumbling effect. Although several FCP methodologies have existed in the previous works, it is still needed to concentrate on the development of FCP for advanced prediction outcomes.

3 The Proposed MOSSA-SAE Based Predictive Model

The working principle of the MOSSA-SAE model is demonstrated in Fig. 1. Mainly, the MOSSA-SAE model allows the IoT devices such as smartphones, laptops, etc to collect the financial details of the users which are then transmitted to the cloud for further analysis. At the cloud server, data pre-processing is initially performed to remove the unwanted data and improve the compatibility of the data for the subsequent processes. Followed by, the SMOTE with SSA is applied for handling the class imbalance problem. Finally, the SAE with SSA model is applied to carry out the classification process

and determine the appropriate class label of the financial data. The detailed working of the processes is discussed in the subsequent sections.
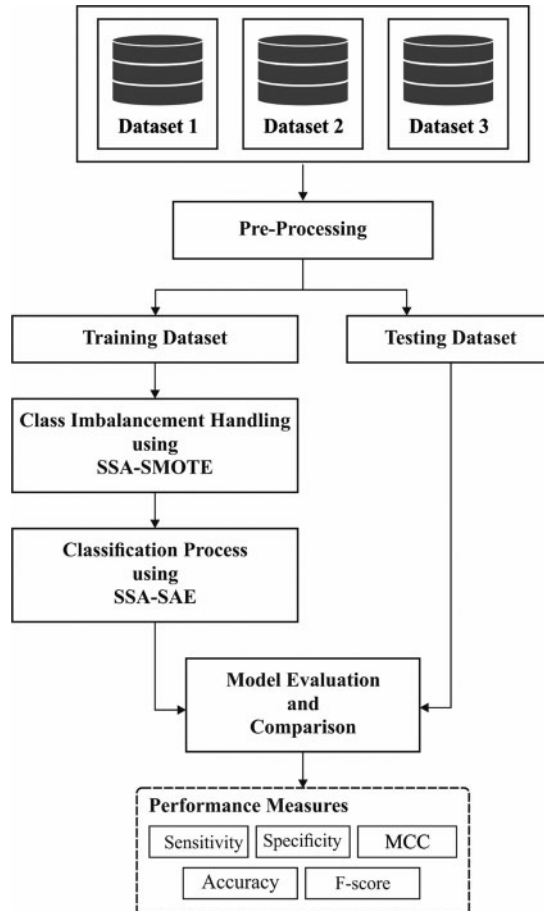
Figure 1: The working process of MOSSA-SAE model

3.1 Overview of SSA

The SSA model is based on the adaptive foraging characteristics of the southern flying squirrels by gliding, an efficiency model employed by the small mammals to pass over long distances. During summer, the squirrels frequently adapt the location to glide between trees and determine the source of food. At the same time, it has the ability to determine the acorn nuts for daily food. Besides, the hickory nuts are discovered and stored for winter season. During winter, they are not active and sustain with the energy that comes from the stored hickory nuts. When the season is altered, they are again come back into active model [17,18]. Based on the foraging nature of flying squirrels, optimization of SSA is derived as given below.

3.1.1 Initiate the Algorithm Parameters

Major parameters of SSA are maximal count of iterations $Iter_{max}$, the population size NP, the count of decision parameters n, predator presence possibility P_{dp}, scaling factor sf, gliding constant

G_c, upper and lower bounds for decision parameters FS_U and FS_L. Therefore, the parameters are fixed in the initial stage of SSA principle.

3.1.2 Initialize Flying Squirrels 'Locations and Sorting'

A location of squirrels is fixed arbitrarily in a search apace as given below:

$$FS_{i,j} = FS_L + rand(\ ) * (FS_U - FS_L) \quad i = 1, 2, \ldots, NP, \quad j = 1, 2, n \tag{1}$$

where rand ( ) refers the uniformly distributed random value from [0, 1].

The fitness measure $f = (f_1 f_2, \ldots, f_{NP})$ of flying squirrel's position is defined by replacing the measure of decision parameters into FF:

$$f_i = f_i \left(FS_{i,1}, FS_{i,2}, \ldots, FS_{i,n} \right), \quad i = 1, 2, \cdots, NP \tag{2}$$

Followed by, the supremacy of food sources is described by the fitness value of flying squirrels' position is arranged in increasing order:

$$\left[sorted_f, sorte_{index} \right] = sort(f) \tag{3}$$

Once the food sources of flying squirrel's position are organized, 3 classes of trees are classified namely, Hickory tree, Oak tree (acorn nuts food source), and Normal tree. A position of supreme food source (low fitness value) is considered to be the hickory nut tree (FS_{ht}), the position of 3 food sources is subjected to be acorn nuts trees (FS_{at}), and remaining are assumed to be the normal trees (FS_{nt}):

$$FS_{ht} = FS\left(sorte\_index\,(1)\right) \tag{4}$$

$$FS_{at}\,(1:3) = FS\left(sorte\_index\,(2:4)\right) \tag{5}$$

$$FS_{nt}\,(1:NP - 4) = FS(sorte\_index(5:NP)) \tag{6}$$

3.1.3 Produce New Locations by Gliding

Tree scenarios might be displayed after dynamic gliding procedure of squirrels.

Condition 1. The flying squirrels residing on Oak tree intends to travel to the hickory nut tree. Hence, novel position is generated in the following:

$$FS_{at}^{new} = \begin{cases} FS_{at}^{old} + d_g G_c \left(FS_{ht}^{old} - FS_{at}^{old} \right), & if\ R \geq P_{dp} \\ random\ location, & otherwise \end{cases} \tag{7}$$

where d_g implies random gliding distance, R_1 defines a function that returns a value from uniform distribution from [0, 1], and G_c refers the gliding constant.

Condition 2. A few squirrels placed on normal trees travel to the Oak tree and meet the regular energy demands. Hence, new position is generated by:

$$FS_{nT}^{new} = \begin{cases} FS_{nT}^{old} + d_g G_c \left(FS_{aT}^{oId} - FS_{nT}^{oId} \right), & if\ R \geq P_{dp} \\ random\ location, & otherwise \end{cases} \tag{8}$$

where R_2 implies a function that returns a value from uniform distribution within [0, 1].

Condition 3. A few squirrels living in normal trees go to hickory nut trees when the energy requirements are satisfied previously. Here, new position of squirrels is produced as shown in the following:

$$FFS_{nt}^{new} = \begin{cases} FS_{nt}^{old} + d_g G_c \left(FS_{ht}^{old} - FS_{nt}^{old} \right), & if \ R \geq P_{dp} \\ random \ location, & otherwise \end{cases} \tag{9}$$

where R_3 defines a function that returns a value from uniform distribution ranged from [0, 1].

From the above conditions, gliding distance d_g is defined to a greater extent. But this value is maximum and results in massive perturbations (7)–(9) that result in unsatisfactory function. For accomplishing a better performance, the scaling factor (sf) has been developed as a divisor of d_g and the values are selected 18. Fig. 2 depicts the flowchart of SSA algorithm.
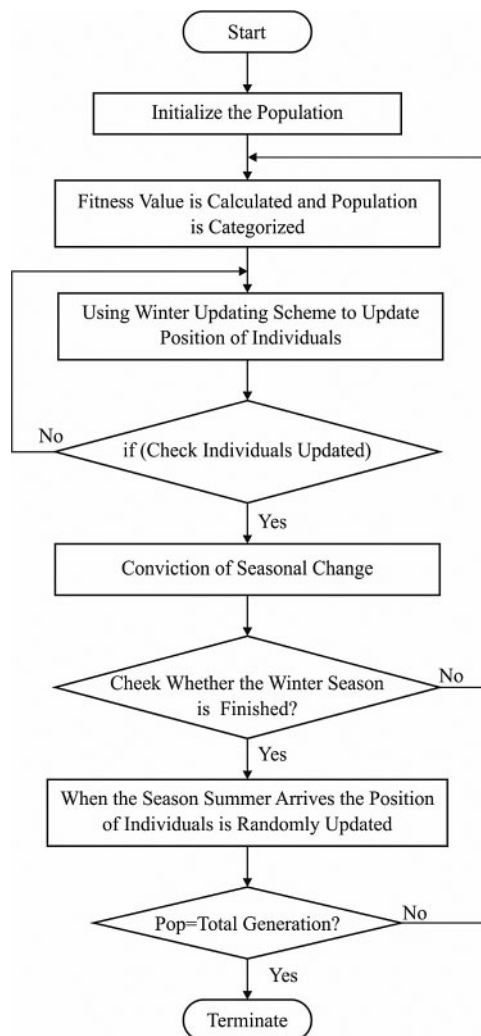
Figure 2: Flowchart of SSA

3.1.4 Check Seasonal Monitoring Condition

In foraging nature of flying squirrels are influenced by season modifications. Hence, seasonal monitoring is deployed to eliminate the trapping condition of local optimal results. Also, seasonal constant Sc and small values are estimated by:

$$S_c^t = \sqrt{\sum_{k=-1}^{n} \left(FS_{atk}^t - FS_{ht,k}\right)^2}, \quad t = 1, 2, 3 \tag{10}$$

$$S_{cmin} = \frac{10E-6}{365^{Iter/(Iter_{max})/2.5}} \tag{11}$$

Followed by, seasonal monitoring mechanism is validated. By considering $S_c^t < S_{cmin}$, the winter season is completed and flying squirrels that have lost the capabilities to find the forest is relocated arbitrarily to the food source:

$$FS_{nt}^{new} = FS_L + Levy(n) \times (FS_U - FS_L) \tag{12}$$

where Levy distribution defines a powerful mathematical device for enhancing the global exploration ability of optimization frameworks:

$$Levy(x) = 0.01 \times \frac{r_a \times \sigma}{|r_b|^{1/\beta}} \tag{13}$$

where r_a and r_b indicates 2 functions that provide a value from uniform distribution within the ranges of [0, 1], β denotes a constant ($\beta = 1.5$), and σ is measured by the given function:

$$\sigma = \left(\frac{\Gamma(1+\beta) \times \sin(\pi\beta/2)}{\Gamma((1+\beta)/2) \times \beta \times 2^{((\beta-1)/2)}} \right)^{1/\beta} \tag{14}$$

where $\Gamma(\beta) = (x-1)!$.

3.1.5 Termination Criterion

The processes get terminated when maximal count of iterations is reached. Else, the nature of producing novel position and validating seasonal monitoring condition is followed repeatedly.

3.2 Class Imbalance Data Handling Process Using SSA-SMOTE

For improving the importance of the smaller and particular regions belong to the positive class in the decision region, the SMOTE is employed for the generation of synthetic instances for the positive class for balancing the training dataset. In general, SMOTE is defined as an over-sampling mechanism presented by [19] and performs in feature space and not in data space. Under the application of this model, count of samples for minority class in actual data set has been maximized by developing novel synthetic samples and results in better decision making process for the minority class whereas naive over-sampling along with replacement results in the decision region of minority class which has to be specific. Here, novel synthetic instances are developed by special kind of parameters namely, over-sampling rate (%) as well as count of nearest neighbors (k). A novel synthetic sample for continuous attributes has been produced as defined in the following:

Step 1: Measure the distance from feature vector in minority class as well as the k nearest neighbors (kNN).

Step 2: Combine the distance attained in Step 1 by using random values from 0 and 1.

Step 3: Sum the value gained from Step 2 for the feature value of actual feature vector. Hence, a fresh feature vector has been developed by

$$x_n = x_o + \delta \cdot (x_{oi} - x_o) \qquad (15)$$

where x_n indicates a novel synthetic sample, x_o is presented by a feature vector of every sample in a minority class, x_{oi} defines the ith chosen NN of x_o, and δ indicates a random value from 0 and 1. For instance, it is applied that $\beta\% = 900\%$ and $k = 5$, so that new synthetic samples have been generated for actual samples. The above steps are repeated for several instances and new synthetic instance have been deployed where NN of x_o is selected randomly. Furthermore, synthetically produced samples are applied for nominal features processed from the given steps [20]:

Step 1: Gain the majority vote from the features by considering the kNN for nominal feature score. At this point, selected in random fashion.

Step 2: Allocate the given value for new synthetic minority sample.

For instance, it is applied with collective features from a sample as $\{A, B, C, D, E\}$ and 2 nearest neighbors are composed of collective features from $\{A, F, C, G, N\}$ as well as $\{H, B, C, D, N\}$, the novel synthetic sample is composed of massive features as depicted by $\{A, B, C, D, N\}$. On the other hand, the SSA is applied to determine the oversampling rate and area of nearest neighbors of the SMOTE technique.

3.3 SSA-SAE Based Predictive Model for FCP

Once the IoT devices transmitted the data, the classification process takes place using SAE. Generally, SAE is a type of unsupervised DL method, which has been stacked by using various AE. The AE is comprised of 2 portions namely, encoder and decoder. Initially, encoder layer has been employed for converting the input x into hidden illustration h, which is depicted as $h = f(wx + b)$, where f, w and b denotes the activation function, weight matrix, bias of current encoder layer whereas the decoder layer is employed to reform x from h, which we devised as $x' = g(w'h + b')$, where x', g, w' and b' refers the simulation outcome, activation function, weight matrix, and bias of present decoder layer, correspondingly. Fig. 3 demonstrates the structure of autoencoder. Moreover, complete training model of AE is pre-training stage and fine-tuning stage. First, AE managed to reduce the cost function and depicted as,
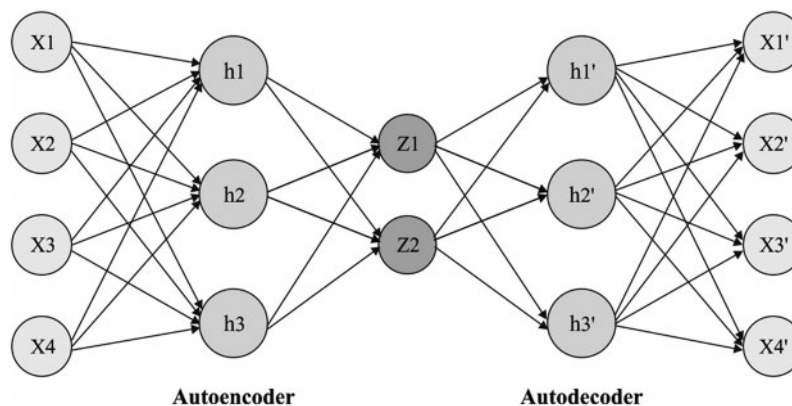

Figure 3: Structure of autoencoder

$$L(x) = \frac{1}{2m} \sum_{i=1}^{m} ||x_i - x'_i||^2 \tag{16}$$

where x_i implies the AE input that indicates the ith sample. x'_i defines the AE result, and reformation of ith sample and m illustrates the count of samples. SAE is developed by including hidden layers. Thus, hidden encoded implementation of existing AE is considered as input of upcoming AE. When the pre-training layers of SAE have been completed, the decoder for an AE has been dropped [21]. This is followed by, connect the encoders as well as fine-tune weights of SAE BY using softmax regression.

Furthermore, it is constrained term that is added to the loss function of AE. The extended loss function is depicted as,

$$L(x) = \frac{1}{2m} \sum_{i=1}^{m} ||x_i - x'_i||^2 + \lambda ||w||_1 \tag{17}$$

where w indicates the weight matrix, and λ signifies the balance factor. When the count of hidden neurons is maximum than input neurons from input layer, it generally adds KL divergence to loss function of AE. Thus, modified loss function is depicted by,

$$L(x) = \frac{1}{2m} \sum_{i=1}^{m} ||x_i - x'_i||^2 + \lambda \sum_{j=1}^{H} KL\left(\rho||\hat{\rho}_j\right) \tag{18}$$

where H defines the count of hidden neurons, and KL shows the KL divergence. ρ refers the sparsity parameter and $\hat{\rho}_j$ demonstrates the moderate activation of input samples of jth hidden neurons.

In this model, Sun refers that minimizing the loss functions (16)–(18) which does not make sure the AE learns "meaningful" features for specific operation. In case of image classification, "useful" features are represented for enhancing the classification rate. Thus, explicit guiding norms relevant to special tasks such as the cross-entropy loss function are recommended to be included in the loss function of AE and approved to be significant. Thus, the modified loss function is illustrated as,

$$L(x) = \frac{1}{2m} \sum_{i=1}^{m} ||x_i - x'_i||^2 + \frac{\lambda}{m} \sum_{i=1}^{m} \sum_{j=1}^{c} label\,(i,j) \cdot log\,pred\,(i,j) \tag{19}$$

where C denotes the count of classes. $label\,(i,j)$ implies the true probability from jth class of ith sample. $log\,pred\,(i,j)$ indicates the examined probability from jth class of ith sample. Followed by, a variant of SAE is applied in the newly generated model. In this case, the weights and bias values of the SAE are chosen by the use of SSA. Fig. 4 showcases the flow diagram of SAE.
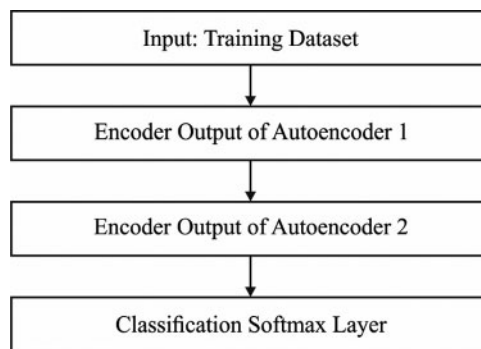
Figure 4: Flow diagram of stacked autoencoder

4 Performance Validation

This section investigates the predictive performance of the MOSSA-SAE algorithm against three benchmark datasets namely qualitative, Polish, and Weislaw datasets. The first qualitative dataset comprises 250 samples with a total of 6 attributes and 2 classes. Similarly, the second Polish dataset includes a set of 43405 samples with 64 attributes and 2 classes. Thirdly, the Weislaw dataset comprises a collection of 240 instances with 30 attributes and 2 classes. The information related to these datasets is given in Tab. 1. Besides, a set of rules generated by the MOSSA-SAE model to carry out the classification process is tabulated in Tab. 2.

Table 1: Dataset description

Dataset	Source	Instances	Attributes	Classes	FC/Non-FC
Qualitative	UCI	250	6	2	107/143
Polish	UCI	43405	64	2	2091/41314
Weislaw	Pietruszkiewicz	240	30	2	112/128

Table 2: Sample rules generated by MOSSA-SAE method

Rule	Antecedent	Consequent
Qualitative dataset		
1	IF FF = 'A'	Non_FC
2	IF CR = 'P'	Non_FC
3	IF CR = 'N' && CO = 'A'	FC
4	IF IR = 'P' AND OP = 'A'	FC
Polish dataset		
5	IF A14 > 145.890 && A20 <= 127.90 && A64 <= 26.245	Non_FC
6	IF A27 > −2.1780 && A21 > 0.6743	Non_FC
7	IF A35 <= 0.2567 && A34 <= 0.6978 && A9 > 0.680	FC
8	IF A13 <= 0.2742 && A63 > 2.894 && A12 <= 2.756	FC
Weislaw dataset		
9	IF X23 <= 1.0245 && X8 <= 9.7895	Non_FC
10	IF X9 > 0.0212 && X23 > 0.2784	Non_FC
11	IF X6 <= 0.1908 && X18 <= 6.1754 && X21 > 0.3134	FC

(Continued)

Table 2: Continued

Rule	Antecedent	Consequent
12	IF X2 <= 0.0875 && X6 <= 0.2032 && X18 <= 4.2753	FC

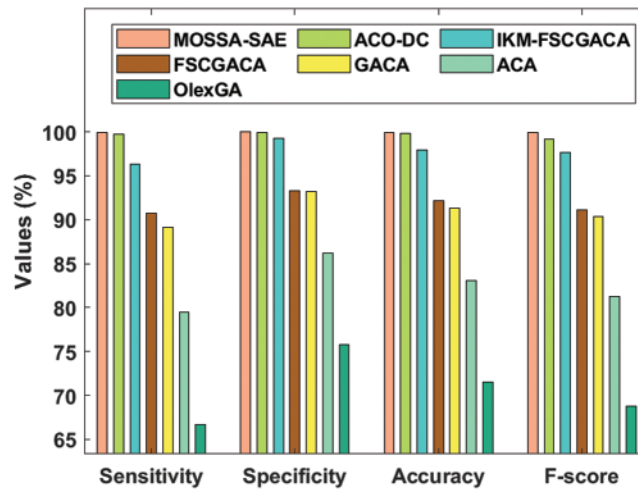### 4.1 Results Analysis on Qualitative Bankruptcy Dataset

Tab. 3 and Fig. 5 studied the comparative results analysis of MOSSA-SAE model on the applied qualitative bankruptcy dataset [14,15]. The resultant values denoted that the OlexGA model has showcased inferior results with the sens. of 66.66%, spec. of 75.78%, accu. of 71.48%, F-score of 68.77%, and MCC of 0.43. At the same time, it is noticed that the ACA model has offered slightly improved outcomes with the sens. of 79.46%, spec. of 86.15%, accu. of 83.05%, F-score of 81.27%, and MCC of 0.66. Eventually, the GACA model has showcased certainly enhanced performance over ACA and OlexGA models with the sens. of 89.09%, spec. of 93.18%, accu. of 91.32%, F-measure of 90.32%, and MCC of 0.82.

Table 3: Result analysis of MOSSA-SAE model with existing models on qualitative bankruptcy dataset

Techniques	Sens.	Spec.	Accu.	F-measure	MCC
MOSSA-SAE	99.95	99.98	99.96	99.95	0.99
ACO-DC	99.76	99.89	99.79	99.14	0.98
IKM-FSCGACA	96.36	99.24	97.93	97.69	0.95
FSCGACA	90.74	93.28	92.14	91.16	0.84
GACA	89.09	93.18	91.32	90.32	0.82
ACA	79.46	86.15	83.05	81.27	0.66
OlexGA	66.66	75.78	71.48	68.77	0.43

Additionally, the FSCGACA model has tried to show moderate results with the sens. of 90.71%, spec. of 93.28%, accu. of 92.14%, F-measure of 91.16%, and MCC of 0.84. Furthermore, the IKM-FSCGACA model has demonstrated moderate outcome with the sens. of 96.36%, spec. of 99.24%, accu. of 97.93%, F-measure of 97.69%, and MCC of 0.95. Moreover, the ACO-DC model has exhibited near optimum outcomes with the sens. of 99.76%, spec. of 99.89%, accu. of 99.79%, F-measure of 99.14%, and MCC of 0.98. But the MOSSA-SAE model has outperformed all the other methods by accomplishing a maximum sens. of 99.95%, spec. of 99.98%, accu. of 99.96%, F-measure of 99.95%, and MCC of 0.99.

Fig. 6 shows the ROC analysis of different FCP models on the applied qualitative bankruptcy dataset. The figure demonstrated that the MOSSA-SAE model has resulted to a maximum ROC of 99.13% whereas the ACO-DC, IKM-FSCGACA, FSCGACA, GACA, and ACA models have led to a minimum ROC of 99.21%, 97.39%, 92.03%, 91.15%, 82.85%, and 71.02% respectively. Therefore, the presented MOSSA-SAE model has showcased effective outcome on the applied qualitative bankruptcy dataset.

Figure 5: Result analysis of MOSSA-SAE model on qualitative bankruptcy dataset
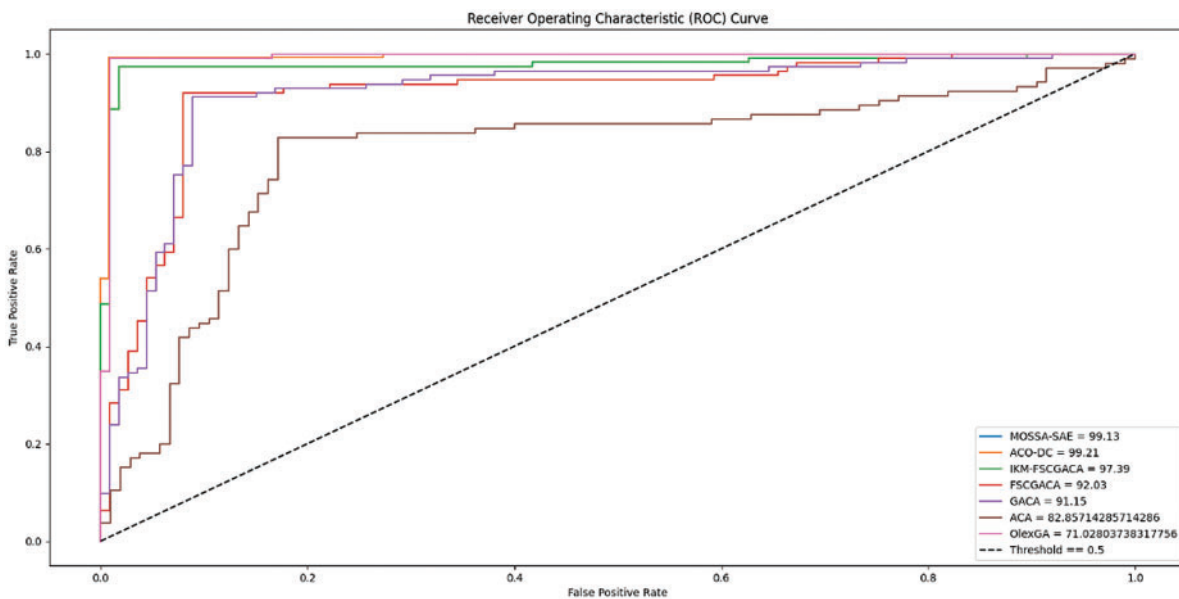
Figure 6: ROC analysis of different FCP models on qualitative bankruptcy dataset

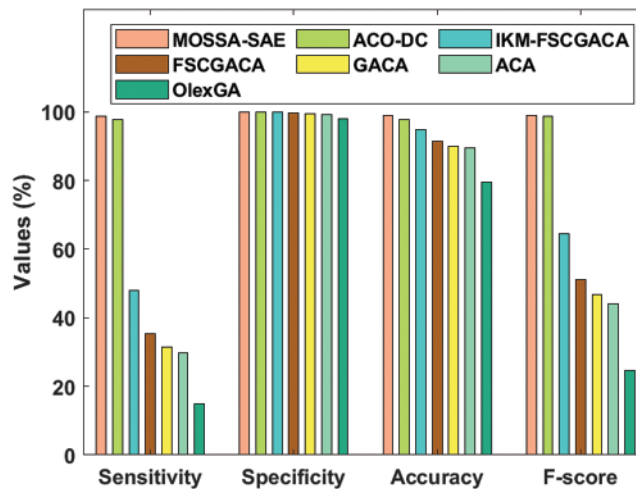### 4.2 Results Analysis on Polish Dataset

Tab. 4 and Fig. 7 presented the comparative results analysis of MOSSA-SAE method on the given polish bankruptcy dataset. The final measures have implied that the OlexGA scheme has demonstrated poor results with the sens. of 14.99%, spec. of 98.06%, accu. of 79.46%, F-measure of 24.64%, and MCC of 0.25. Simultaneously, it is clear that the ACA approach has provided moderate results with the sens. of 29.65%, spec. of 99.20%, accu. of 89.41%, F-measure of 44.09%, and MCC of 0.46.

Followed by, the GACA framework has depicted improved performance over ACA and OlexGA methodologies with the sens. of 31.55%, spec. of 99.48%, accu. of 89.98%, F-measure of 46.84%, and MCC of 0.50. In addition, the FSCGACA scheme has attempted to display considerable outcomes

with the sens. of 35.29%, spec. of 99.58%, accu. of 91.39%, F-measure of 51.12%, and MCC of 0.54. Moreover, the IKM-FSCGACA framework has illustrated acceptable results with the sens. of 47.91%, spec. of 99.89%, accu. of 94.72%, F-measure of 64.39%, and MCC of 0.66. Eventually, the ACO-DC technique has implied closer optimal results with the sens. of 97.63%, spec. of 99.90%, accu. of 97.69%, F-measure of 98.80%, and MCC of 0.71. However, the MOSSA-SAE approach has surpassed the compared models by gaining higher sens. of 98.76%, spec. of 99.93%, accu. of 98.99%, F-measure of 98.97%, and MCC of 0.96.

Table 4: Result analysis of MOSSA-SAE model with existing models on polish bankruptcy dataset

Techniques	Sens.	Spec.	Accu.	F-Measure	MCC
MOSSA-SAE	98.76	99.93	98.99	98.97	0.96
ACO-DC	97.63	99.90	97.69	98.80	0.71
IKM-FSCGACA	47.91	99.89	94.72	64.39	0.66
FSCGACA	35.29	99.58	91.39	51.12	0.54
GACA	31.55	99.48	89.98	46.84	0.50
ACA	29.65	99.20	89.41	44.09	0.46
OlexGA	14.99	98.06	79.46	24.64	0.25

Figure 7: Result analysis of MOSSA-SAE model on polish bankruptcy dataset

Fig. 8 displays the ROC analysis of dissimilar FCP models on the applied Polish bankruptcy dataset. The figure revealed that the MOSSA-SAE model has led to a higher ROC of 99.13% whereas the ACO-DC, IKM-FSCGACA, FSCGACA, GACA, and ACA models have accomplished to a lower ROC of 99.21%, 97.39%, 92.03%, 91.15%, 82.85%, and 71.02% respectively. So, the presented MOSSA-SAE model has offered superior performance on the applied dataset.
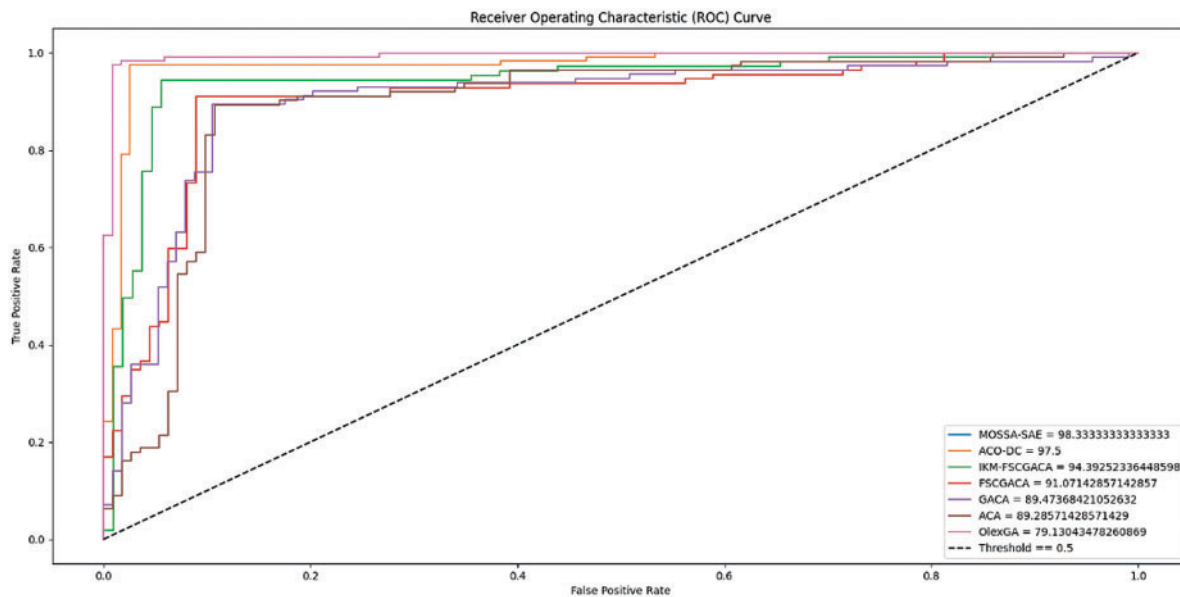
Figure 8: ROC analysis of different FCP models on polish bankruptcy dataset

5 Conclusion

This paper has developed an effective MOSSA-SAE model for FCP in IoT environments. The MOSSA-SAE model encompasses diverse subprocesses, *i.e.*, preprocessing, class imbalance handling, parameter tuning, and classification. Mainly, the MOSSA-SAE model allows the IoT devices such as smartphones, laptops, etc. to collect the financial details of the users which are then transmitted to the cloud for further analysis. At the cloud server, data preprocessing is firstly accomplished to remove the unwanted data and improve the compatibility of the data for the subsequent processes. After that, the SMOTE with SSA is applied for handling the class imbalance problem. Finally, the SAE with SSA model is applied to carry out the classification process and determine the appropriate class label of the financial data. The application of SSA helps to tune the parameters involved in SMOTE and SAE. A widespread experimental validation process is executed on the benchmark financial dataset and the results are studied under separate aspects. The experimental values guaranteed the superior performance of the MOSSA-SAE method on the applied dataset. As a part of future work, the predictive outcome of the MOSSA-SAE method can be improved using clustering and feature selection methodologies.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] K. J. Forbes and R. Rigobon, "No contagion, only interdependence: Measuring stock market comovements," *The Journal of Finance*, vol. 57, no. 5, pp. 2223–2261, 2002.

[2] M. Ala'raj and M. F. Abbod, "Classifiers consensus system approach for credit scoring," *Knowledge-Based Systems*, vol. 104, no. 8, pp. 89–105, 2016.

[3] J. Uthayakumar, N. Metawa, K. Shankar and S. K. Lakshmanaprabu, "Financial crisis prediction model using ant colony optimization," *International Journal of Information Management*, vol. 50, no. 5, pp. 538–556, 2020.

[4] Z. Huang, H. Chen, C.-J. Hsu, W.-H. Chen and S. Wu, "Credit rating analysis with support vector machines and neural networks: A market comparative study," *Decision Support Systems*, vol. 37, no. 4, pp. 543–558, 2004.

[5] J. Sun and H. Li, "Financial distress prediction using support vector machines: Ensemble *vs.* individual," *Applied Soft Computing*, vol. 12, no. 8, pp. 2254–2265, 2012.

[6] Y. Lin, H. Guo and J. Hu, "An SVM-based approach for stock market trend prediction," in *The 2013 Int. Joint Conf. on Neural Networks (IJCNN)*, Texas, USA, pp. 1–7, 2013.

[7] R. Geng, I. Bose and X. Chen, "Prediction of financial distress: An empirical study of listed chinese companies using data mining," *European Journal of Operational Research*, vol. 241, no. 1, pp. 236–247, 2015.

[8] Y. Li and X. Zeng, "Sequential multi-criteria feature selection algorithm based on agent genetic algorithm," *Applied Intelligence*, vol. 33, no. 2, pp. 117–131, 2010.

[9] J.-D. Shie and S.-M. Chen, "Feature subset selection based on fuzzy entropy measures for handling classification problems," *Applied Intelligence*, vol. 28, no. 1, pp. 69–82, 2008.

[10] W. TsaoPan, "A new fruit fly optimization algorithm: Taking the financial distress model as an example," *Knowledge-Based Systems*, vol. 26, no. 7, pp. 69–74, 2012.

[11] G. Wang, J. Ma and S. Yang, "An improved boosting based on feature selection for corporate bankruptcy prediction," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2353–2361, 2014.

[12] F. Lin, D. Liang, C.-C. Yeh and J.-C. Huang, "Novel feature selection methods to financial distress prediction," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2472–2483, 2014.

[13] J. Uthayakumar, T. Vengattaraman and P. Dhavachelvan, "Swarm intelligence based classification rule induction (CRI) framework for qualitative and quantitative approach: An application of bankruptcy prediction and credit risk analysis," *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 6, pp. 647–657, 2020.

[14] N. Metawa, I. V. Pustokhina, D. A. Pustokhin, K. Shankar and M. Elhoseny, "Computational intelligence-based financial crisis prediction model using feature subset selection with optimal deep belief network," *Big Data*, vol. 9, no. 2, pp. 100–115, 2021.

[15] J. Uthayakumar, N. Metawa, K. Shankar and S. K. Lakshmanaprabu, "Intelligent hybrid model for financial crisis prediction using machine learning techniques," *Information Systems and e-Business Management*, vol. 18, no. 4, pp. 617–645, 2020.

[16] S. Sankhwar, D. Gupta, K. C. Ramya, S. Sheeba Rani, K. Shankar *et al.,* "Improved grey wolf optimization-based feature subset selection with fuzzy neural classifier for financial crisis prediction," *Soft Computing*, vol. 24, no. 1, pp. 101–110, 2020.

[17] M. Jain, V. Singh and A. Rani, "A novel nature-inspired algorithm for optimization: Squirrel search algorithm," *Swarm and Evolutionary Computation*, vol. 44, no. 4, pp. 148–175, 2019.

[18] T. Zheng and W. Luo, "An improved squirrel search algorithm for optimization," *Complexity*, vol. 2019, pp. 1–31, 2019.

[19] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[20] K.-J. Wang, B. Makond, K.-H. Chen and K.-M. Wang, "A hybrid classifier combining SMOTE with PSO to estimate 5-year survivability of breast cancer patients," *Applied Soft Computing*, vol. 20, no. 3, pp. 15–24, 2014.

[21] G. Liu, Y. Li, L. Jiao, Y. Chen and R. Shang, "Multiobjective evolutionary algorithm assisted stacked autoencoder for PolSAR image classification," *Swarm and Evolutionary Computation*, vol. 60, no. 6, pp. 100794, 2021.