

# A Robust Deep Learning Model for Financial Distress Prediction

Magdi El-Bannany<sup>1</sup>

Department of Accounting  
College of Business Administration  
University of Sharjah 27272, UAE

Meenu Sreedharan<sup>2</sup>, Ahmed M. Khedr<sup>3</sup>

Department of Computer Science  
College of Computing and Informatics  
University of Sharjah 27272, UAE

**Abstract**—This paper investigates the ability of deep learning networks on financial distress prediction. This study uses three different deep learning models, namely, Multi-layer Perceptron (MLP), Long Short-term Memory (LSTM) and Convolutional Neural Networks (CNN). In the first phase of the study, different Optimization techniques are applied to each model creating different model structures, to generate the best model for prediction. The top results are presented and analyzed with various optimization parameters. In the second phase, MLP, the best classifier identified in the first phase is further optimized through variations in architectural configurations. This study investigates the robust deep neural network model for financial distress prediction with the best optimization parameters. The prediction performance is evaluated using different real-time datasets, one containing samples from Kuwait companies and another with samples of companies from GCC countries. We have used the technique of resampling for all experiments in this study to get the most accurate and unbiased results. The simulation results show that the proposed deep network model far exceeds classical machine learning models in terms of predictive accuracy. Based on the experiments, guidelines are provided to the practitioners to generate a robust model for financial distress prediction.

**Keywords**—Financial distress prediction; multi-layer perceptron; long short-term memory; convolutional neural network; deep neural network; optimized deep learning model

## I. INTRODUCTION

Financial distress is a condition where a company faces financial difficulties, which is also referred to as Business or Corporate Failure. Financial distress can induce a great impact on any company, stakeholders, and the economy of a country. The investors rely on financial statements disclosed by any company, which can be forged by the company executives, leaving them with very little chance of getting the original financial information. Hence, a reliable distress prediction model is necessary for investors to adjust their investment strategies, so as to reduce the loss of investments. Also, it will help the company managers to take corrective measures to prevent the crisis before it happens. Many researchers have used primitive statistical techniques to generate relevant models, however, machine learning algorithms were found to create a more robust model for distress prediction.

Building a robust model with acceptable prediction performance can help the managers and investors to manage risks and take actions on time, to prevent bankruptcy before it

happens. Deep learning is a field of machine learning, containing multiple layers of nonlinear processing units, to learn features from real-time data. With low cost, high computational ability and availability of different optimization techniques, deep learning has been an area of interest for many types of research. So, the question is: Can we create a robust prediction model using deep learning techniques? Many research works are available on distress prediction using classical machine learning classifiers and ensemble techniques like Decision Tree [1,4], Neural Networks [5-9], Support Vector Machines [13], etc. But to the best of our knowledge, there is no existing research work analyzing different deep learning models on financial distress prediction and how to optimize these models. Hence, this paper provides an insight into the deep learning models for financial distress prediction which will be of great significance to generate a more robust model.

This paper focuses on building deep neural networks including multi-layer perceptron, LSTM, and Convolutional Neural Network and optimization of the same using different optimization techniques. We plan to train the networks using different transfer functions and training algorithms to generate a robust model for distress prediction. In the second phase of the study, we try to further optimize the models using different architectural configurations by varying the number of deep layers and neurons at each level. In this paper, we focus on two different datasets, one from the companies in Kuwait and other from companies in GCC, to analyze the performance with real-time and varying structures. The experiments in this paper mainly focus on providing proper guidelines for any researcher, to build a robust deep learning model for financial distress prediction in the future. The results of the study clearly indicate that the proposed model has significantly higher predictive accuracy compared to classical machine learning models.

This paper is organized as follows: The next section reviews the studies related to financial distress prediction. Research methodology: data collection and data modeling are described in Section 3. The prediction performance of all the selected deep models is described, optimized, analyzed and compared in Section 4. Finally, in Section 5, we infer our conclusion, with a set of guidelines to generate a robust model for financial distress prediction using deep neural networks.

## II. LITERATURE REVIEW

The major aim of a Financial Distress Prediction Model is to determine whether a company has a chance of experiencing financial distress in the future. Bankruptcy, Insolvency, etc. are the formal signs of financial distress in a company. Discriminant analysis [3] and logit model [12, 14] are the initial and traditional statistical models used in the field of distress prediction. These traditional linear techniques are simple but unrealistic, and therefore cannot be used to generate a robust model for making real-time predictions. In 2014, a simple hazard model for the distress prediction of banks in the Gulf Cooperation Council countries was built [2]. Machine learning using data mining techniques like Logistic Regression, Support Vector Machines [11, 13] and Neural Networks [15, 17, 19, 20] were introduced as alternatives in later researches. In 2015, Ruibin Geng, Indranil Bose, and Xi Chen evaluated the performance of machine learning techniques for the distress prediction of listed Chinese companies [21]. The paper compared the three highly used data mining classifiers and evaluated the performance by combining the results using Majority Voting. Researches on datasets, collected from different countries, using data mining algorithms are also available [10]. In 2019, an analysis of a two-stage model for distress prediction is studied in [16]. It basically focuses on feature selection as a critical step, through data envelopment analysis. Distress prediction using deep learning is presented in [18], which uses unstructured textual data in statements for prediction. This problem can be solved in case the data sets are distributed among a number of sites cross the networks [22-30].

To the best of our knowledge, no researches are available in the literature, investigating the performance of various deep learning network models on financial distress prediction and how to optimize them. In this paper, we evaluate the deep neural network models for financial distress prediction using two different datasets. We also apply the various optimization techniques on deep models to generate a robust model. As the last phase of this study, the results of the proposed model are compared with that of classical machine learning classifiers like support vector machine and decision tree. This paper helps researchers to develop a robust model for financial distress prediction using conclusions derived at the end of this paper.

## III. RESEARCH METHODOLOGY

This section explains data modeling, the algorithm, how the model is evaluated and the datasets used for simulation. The data modeling section is divided into two phases. In the first phase of the study, we analyze the performance of deep neural networks on financial distress prediction. We also investigate the predictive performance of the models by optimizing the models using various optimization techniques. In this phase, we select the best performing model obtained from the combination of optimization and activation functions for further analysis. Phase 2, focuses on optimizing the outstanding model from phase 1 by restructuring the architectural configurations i.e., Network depth and Network width. All the experiments in this study apply the technique of resampling using k-fold evaluation metrics, to get unbiased

and most accurate results. A schematic representation of the steps involved in this study is shown in Fig. 1.

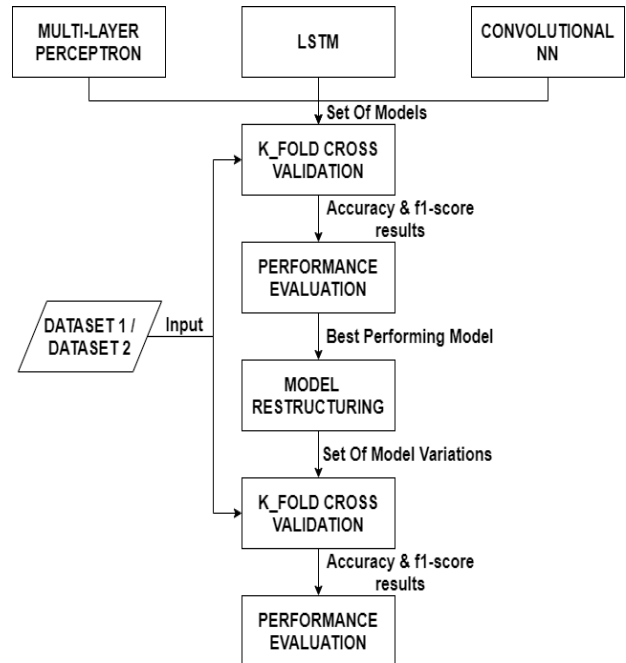


Fig. 1. Schematic Diagram Showing the Steps of Research.

### A. Data Collection and Financial Indicators

1) *Dataset1*: The primary dataset used for prediction performance analysis contains sample data collected from the companies in Kuwait. The dataset, referred to as dataset1, contains 64 sample companies with balanced data, 32 financially healthy and 32 financially distressed companies during the period of 2010 to 2017. Dataset1 contains 24 financial indicators or attributes extracted from financial statements and balance sheets of respective companies.

2) *Dataset2*: A second dataset, with sample data collected across different countries, is used for modeling in order to verify whether the models can identify the common dynamics across datasets. The second dataset, referred to as dataset2, contains 120 samples from six GCC countries including UAE, Bahrain, Kuwait, Qatar, Saudi Arabia, and Oman. In order to create a balanced data sample, an equal number of financially healthy and financially distressed companies are included in dataset2 as in dataset1. Hence, dataset2 contains 60 healthy and 60 distressed companies' data collected from the balance sheets and financial statements of respective companies during the period of 2010 to 2017. Since the data contains samples collected across different countries, the number of financial indicators is only 19, less when compared to dataset1, due to missing common attributes among countries.

### B. Data Interpretation and Preprocessing

The problem addressed in this paper is a binary classification problem, to determine whether a company can be labeled as financially distressed or not. The output/target attribute in the financial dataset belongs to two classes: one for financially distressed and the other for financially healthy

companies. Except for this target attribute, which is binary, all the other attributes in the dataset are continuous values. The initial datasets contain incomplete samples and those with missing data and null values, which are removed during preprocessing.

### C. Data Modeling

1) *Artificial neural network*: A neural network is a machine learning classifier based on an artificial representation of the human brain. Neural Network Architecture consists of an input layer, a layer of output nodes, and one or more intermediate layers.

The corresponding weights are multiplied with the input to calculate  $Y_k$  as,

$$Y_k = w_0 + \sum_{j=1}^n w_{jk}x_j$$

where  $Y_k$  is a weighted sum of input signals at node  $k$ ;  $w_0$  is bias value;  $w_{jk}$  is, the weight associated with the connection between node  $k$  and the input node  $j$ ;  $x_j$  is a value of input node  $j$ ;  $n$  is the number of input nodes. The weighted sum output is then served as input to an activation function.

$$(Y) = \frac{1}{1 + e^{-Y}}$$

The value after applying the activation function is the output value from node  $k$ , which is considered as the input to the next layer in the architecture.

Phase 1: In this phase, three highly preferred deep learning network models namely MLP, LSTM, and CNN are developed, trained and tested for financial distress prediction. All three models are trained with different activation and optimization algorithms with a fixed number of neurons at each level, to evaluate the prediction performance. We have generated different variations of each deep learning model through different combinations of transfer functions and optimization algorithms. The different transfer functions used in this study include sigmoid or logistic function, reLu or Rectified linear units and tanh or hyperbolic tangent and the optimization functions are stochastic gradient descent, Adam and Adagrad optimizers. The generated model variations were then trained and tested using two datasets. This phase aims to investigate the best deep neural network model for financial distress prediction, which is further optimized in phase 2.

Phase 2: A variation of predictive performance was observed with different layers of MLP in Phase 2. Hence the second phase of our study focuses on further evaluation of Multi-Layer Perceptron, the outstanding model from phase 1 in terms of accuracy and f1-score for different architectural configurations. There exists a myriad of hyperparameters that can be tuned to improve the predictive performance of a deep neural network. We focus on tuning the main two hyperparameters namely Network depth and Network width, which can make a difference in the algorithm exploding or converging. MLP models are designed by varying the number of hidden layers and the number of neurons at each level. The aim of this phase is to investigate the optimum parameters for Network depth and Network width, to generate a robust model for financial distress prediction. Python's Scikit-learn and

Keras packages are used for training the models and to generate the results. We have used the resampling technique called Cross-validation for performance evaluation, which is further discussed in Section 3.4.

Finally, we have also compared the performance of the proposed model with that of classic machine learning models. The simulation results indicated that the proposed model has significantly higher predictive accuracy when compared to support vector machine and decision tree classifier models.

### D. Evaluation

In this paper, we have used Keras, the most powerful deep learning library in python, to build and evaluate deep learning models. The models are evaluated using k-fold cross-validation in python's scikit-learn. Thus, we use the technique of resampling to estimate the performance of models. In this technique, the data is split into  $k$ -parts, and the model is trained using all parts except 1, which is kept aside as test data for evaluating the performance of the model. In this paper, we have chosen to repeat this process 10 times and the average value across all the built models is used as the robust prediction performance estimation. This process is stratified because it attempts to balance the number of samples belonging to each class in the  $k$ -splits.

In this study, the predictive performance of the machine learning classifiers is measured in terms of accuracy and f1-score based on the common evaluation metrics of machine learning. Training and testing accuracy measures are used for performance evaluation. Since we use a  $k$ -fold cross-validation score, the mean and standard deviation across the 10 models are calculated for the metrics training accuracy, testing accuracy, and f1-score. Training accuracy is the ratio of correct predictions on the training dataset while testing accuracy is the same calculated on the testing dataset. F1 Score is a function of precision and recall, where precision-recall and F1 score are defined as follows:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} = \frac{\text{True Positive}}{\text{Total Predicted Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} = \frac{\text{True Positive}}{\text{Total Actual Positive}}$$

$$\text{F1 Score} = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

where the true positives, true negatives, false positives, and true negatives are defined by the confusion matrix:

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

In our study negative indicates financially distressed companies and positive indicates financially healthy companies. A balance between precision and recall can be obtained if f1 score is used as a performance measure.

## IV. RESULT ANALYSIS

In this paper, we have performed the analysis of three types of deep neural networks namely MLP, LSTM, and CNN

on two financial datasets. All these networks were trained and tested using samples from both the datasets. The results of phase1 i.e. accuracy and F1 score of deep learning models – MLP, LSTM and CNN on financial distress prediction using dataset1 and dataset2 are shown in Fig. 2, 3, 4 and 5, respectively. In phase 1 experiments, all three models were trained with different optimization technique, to evaluate the prediction performance. In the experiments, we have used combinations of different optimization functions with activation functions for optimizing each deep learning model. The above steps were repeated for dataset1 and dataset2. It was found that all the three deep learning models with sigmoid activation function and Adam optimizer outperformed any other combinations. The accuracy and f1 score of models with the best optimization techniques (Sigmoid + Adam optimizer) are shown in Fig. 2, 3, 4, and 5. The combinations of other optimization algorithms could not generate a robust predictive model (not shown).

In short, the phase 1 experiments concluded that the deep learning models designed with sigmoid activation function and Adam optimizer yielded the best predictive performance for financial distress prediction. Fig. 2 and 3 indicates the prediction performance of models for dataset1, while that of dataset2 is depicted in Fig. 4 and 5. The two graphical representations clearly show that Multi-layer Perceptron (MLP) has the highest performance in terms of accuracy, precision, and recall.

The predictive performance of dataset1 is higher compared to dataset2 because the former has more financial attributes compared to the latter, which helps to build a better classification model during training. However, the predictive performance of MLP outperforms the performance of LSTM and CNN models with dataset1 and dataset2. Hence, we can conclude that MLP is the best suited deep learning classifier for financial distress prediction. Accordingly, in the next phase of this study, we have selected Multi-Layer Perceptron model with sigmoid transfer function and Adam optimizer.

The results above represent the mean and standard deviation (in brackets) obtained using 10 times repeated random sub-sampling.

The results above represent the mean and standard deviation (in brackets) obtained using 10 times repeated random sub-sampling.

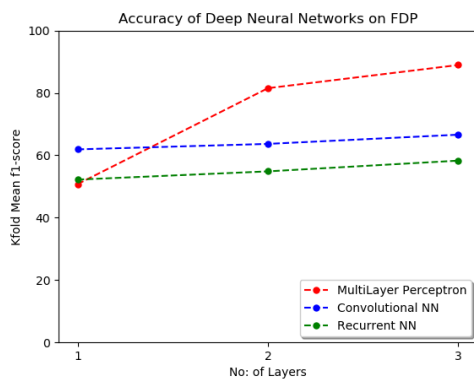


Fig. 2. Mean Accuracy (from 10 Repeated Random Sub-Sampling) of MLP, CNN and LSTM using DataSet1.

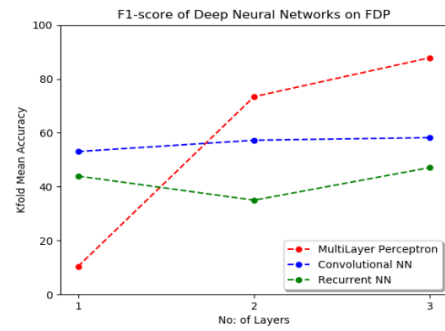


Fig. 3. Mean F1 Score (from 10 Repeated Random Sub-Sampling) of MLP, CNN and LSTM using DataSet1.

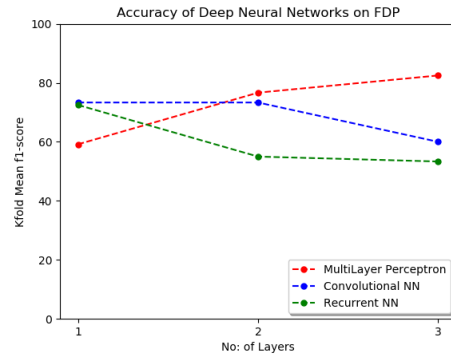


Fig. 4. Mean Accuracy (from 10 Repeated Random Sub-sampling) of MLP, CNN and LSTM using DataSet2.

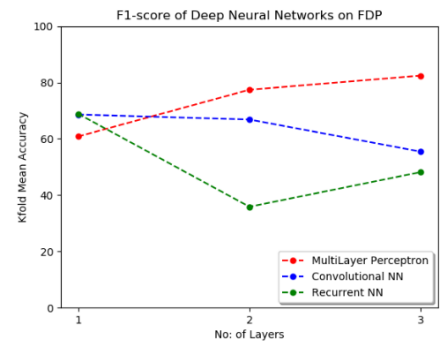


Fig. 5. Mean F1 Score (from 10 Repeated Random Sub-Sampling) of MLP, CNN and LSTM using DataSet2.

In phase 2, we further analyze the deep learning model – MLP for financial distress prediction. The results of phase 2 with different configurations of MLP are shown in Tables I and II. Based on the experiments from the preliminary study, the training method and the activation function were Adam Optimizer and sigmoid function respectively. The number of hidden layers and the number of neurons at each layer are varied in this phase, for further optimization of MLP architecture. The results of MLP on Financial distress prediction, with 16 different architectures developed, trained and tested are listed in Tables I and II. It can be noticed that any changes in the number of layers or the number of neurons in each layer affect the proficiency of the model. For example, as shown in Table I, MLP with configuration 10-10-10-10-10 had an acceptable accuracy value of 90.76% but the network with 50-50-50-50-50 configuration had a poor prediction

accuracy of 70.91%. An optimized architecture for dataset1 is 10-20-10-20-10(93.79%) and that of dataset2 is 20-20-10-10-10(84.17%). The variation in the performance of models on two datasets is due to the change in the number of financial indicators in each dataset. A dataset with a higher number of attributes can be trained better and can generate a more robust model than a dataset with a smaller number of attributes.

A reduction in the accuracy (< 80% for dataset1 and < 70% for dataset2) was observed with networks containing more than 5 layers and hence were not able to generate a robust model (networks are not shown in the results table). Maximum prediction performance is obtained with a 4-layer architecture for experiments with dataset1 and dataset2. Also, the accuracy value started decreasing when the number of neurons at each level was approaching twice the number of attributes in the input dataset. A robust model was not generated after the number of nodes was set equal to and greater than 40(72.22%) and 30(68.33%) for dataset 1 and dataset 2 respectively. Hence this study indicates that higher prediction performance is obtained when the number of neurons at each level is less than twice the number of input attributes in the dataset. The prediction performance is maximum with an architecture containing a combination of 10 and 20 neuron units at hidden layers for both dataset1 and dataset2.

In the final phase of the study, we have compared our optimized MLP performance with the classic machine learning algorithms including support vector machine and decision tree. The prediction results in terms of accuracy are shown in Table III. The statistical results indicated that the prediction accuracy of the proposed optimized model was significantly higher than that of base machine learning models using both datasets.

TABLE I. PREDICTIVE ACCURACY OF MLP USING DATASET1

Structure	Training Accuracy: Mean (Standard Deviation)	Testing Accuracy: Mean (Standard Deviation)
5-5	72.44(21.97)	71.06(19.61)
5-5-5	95.57(2.81)	90.76(7.43)
10-10-10	87.87(6.25)	86.06(8.52)
5-5-5-5	95.89(2.75)	92.27(6.29)
10-10-10-10	97.16(3.22)	89.09(6.19)
20-20-20-20	96.49(3.22)	90.61(5.27)
20-10-20-10	96.83(2.08)	91.97(6.81)
10-20-10-20	97.15(2.38)	90.76(7.43)
5-5-5-5-5	96.20(1.91)	92.27(6.29)
10-10-10-10-10	97.46(2.60)	90.76(7.43)
20-20-20-20-20	96.50(2.55)	89.90(6.67)
30-30-30-30-30	98.10(1.09)	87.42(8.76)
10-20-10-20-10	97.77(2.04)	93.79(4.40)
20-20-10-10-10	98.72(0.90)	90.76(7.43)
50-50-50-50-50	72.77(20.43)	70.91(20.29)
100-50-100-50-100	55.33(3.12)	52.27(2.27)

TABLE II. PREDICTIVE ACCURACY OF MLP USING DATASET2

Structure	Training Accuracy: Mean (Stand Deviation)	Testing Accuracy: Mean (Stand Deviation)
5-5	73.33(15.82)	70.00(14.43)
5-5-5	75.83(14.62)	73.33(13.74)
10-10-10	78.61(9.25)	76.67(10.23)
5-5-5-5	80.56(4.87)	75.83(9.82)
10-10-10-10	87.83(3.53)	82.50(3.82)
20-20-20-20	84.67(2.60)	80.83(9.75)
20-10-20-10	78.33(10.67)	78.33(5.00)
10-20-10-20	81.33(8.47)	78.83(9.43)
5-5-5-5-5	87.22(4.61)	81.67(3.73)
10-10-10-10-10	86.11(8.03)	82.50(8.62)
20-20-20-20-20	81.17(7.99)	75.17(9.75)
30-30-30-30-30	70.28(15.00)	68.33(14.04)
20-20-10-10-10	87.67(5.73)	84.17(5.34)
10-20-10-20-10	83.83(6.64)	81.67(2.89)
50-50-50-50-50	53.33(7.45)	52.50(5.59)
100-50-100-50-100	55.33(8.30)	53.33(5.53)

TABLE III. PREDICTION RESULTS (ACCURACY) OF MLP, SVM, AND DT

	DataSet1	DataSet2
Deep Neural Network (MLP)	0.93	0.84
Support Vector Machine (SVM)	0.85	0.65
Decision Tree Classifier (DT)	0.80	0.67

## V. CONCLUSION

In this paper, we have investigated the performance of deep learning neural networks namely MLP, LSTM, and CNN, on financial distress prediction. We have found that MLP networks are the best-performing distress prediction model. In the last phase of the paper, we have applied different architectural variations to MLP for further optimization. It was found that an accepted predictive performance rate can be achieved if the model is designed with 3 or 4 hidden layers with the neuron count at each level not exceeding twice the number of input attributes in the dataset. We have trained and tested the models using two different datasets with a varying number of input attributes and found that more the number of financial indicators, a better robust model can be generated. The simulation results also indicate that the proposed model has higher performance when compared to classic machine learning models like support vector machine and decision tree.

## REFERENCES

- [1] Adrian Gepp, Kuldeep Kumar, "Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques", *Procedia Computer Science*, Volume 54, 2015.
- [2] Aktham I. Maghyereha, Basel Awartani, Bank distress prediction: Empirical evidence from the Gulf Cooperation Council countries. United Arab Emirates University, the United Arab Emirates and Plymouth University, U, 2014.

- [3] Altman, E. I. , Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609, 1968.
- [4] Chang, C. L., & Chen, C. H. (2009). Applying the decision tree and neural network to increase the quality of dermatologic diagnosis. *Expert Systems with Applications*, 36, 4035–4041
- [5] Chia-Pang Chan, Ching-Hsue Cheng, An Attribute Selection Based Classifier to Predict Financial Distress, 2012.
- [6] D. Wu, L. Liang, Z. Yang, Analyzing the financial distress of Chinese public companies using probabilistic neural networks and multivariate discriminate analysis, *Socio-Econ. Plan. Sci.* 42 (3) 206–220, 2008.
- [7] Han, J., & Kamber, M. (2001). *Data mining: Concepts and techniques*. San Francisco, CA, USA: Morgan Kaufmann.
- [8] Huang, M. J., Chen, M. Y., & Lee, S. C. Integrating data mining with case-based reasoning for chronic disease prognosis and diagnosis. *Expert Systems with Applications*, 32(3), 856–867, 2007.
- [9] I. Guyon, and A. Elisseeff, “An introduction to variable and feature selection,” *The Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.
- [10] Jae Kwon Bae, Predicting financial distress of the South Korean manufacturing industries, *Expert Systems with Applications*, Volume 39, Issue 10, 2012.
- [11] Jie Sun, Hamido Fujita, Peng Chen, Hui Li. "Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble", *Knowledge-Based Systems*, 2017
- [12] Jie Sun, Hui Li, Qing-Hua Huang, Kai-Yu He, Predicting financial distress and corporate failure: A review from the state-of-art definitions, modeling, sampling and featuring approaches, *Knowledge-based systems*, 57:41-56, 2014.
- [13] J.H. Min, Y.-C. Lee, Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters, *Expert Syst. Appl.* 28 (2005) 128–13.
- [14] J. Ohlson, Financial ratio and the probabilistic prediction of bankruptcy, *J. Account. Res.* 18, 109–131, 1980.
- [15] L. Cleofas-Sánchez, V. García, A.I. Marqués, J.S. Sánchez, Financial distress prediction using the hybrid associative memory with translation, *Applied Soft Computing*, Volume 44, 2016.
- [16] Mohammad Mahdi Mousavi, Jamal Ouenniche, Kaoru Tone, “A comparative analysis of two-stage distress prediction models”, *Expert Systems with Applications*, Volume 119, 2019.
- [17] M. A. Hall, and G. Holmes, “Benchmarking attribute selection techniques for discrete class data mining,” *Knowledge and Data Engineering, IEEE Transactions on*, vol. 15, no. 6, pp. 1437-1447, 2003.
- [18] Rastin Matin, Casper Hansen, Christian Hansen, Pia Mlgaard, "Predicting distresses using deep learning of text segments in annual reports", *Expert Systems with Applications*, Volume 132, 2019.
- [19] P. Ravisankar, V. Ravi a, I. Bose, Failure prediction of dotcom companies using neural network-genetic programming hybrids, *Expert systems with applications*, 36(3):4830-4837, 2009.
- [20] P. Ravisankar, V. Ravi a, Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP, *Knowledge-based systems*, 23(8):823-831, 2011.
- [21] Ruibin Geng, Indranil Bose, Xi Chen, Prediction of financial distress: An empirical study of listed Chinese companies using data mining, 2015.
- [22] Ahmed M. Khedr, and Bhatnagar R., New Algorithm for Clustering Distributed Data using k-means, *Computing and Informatics*, Vol. 33, pp. 1001-1022, 2014.
- [23] Ahmed M. Khedr, Decomposable Naive Bayes Classifier for Partitioned Data, *Computing and Informatics*, Vol. 31, pp. 1511-1531, 2012.
- [24] Ahmed M. Khedr, Nearest Neighbor Clustering over Partitioned Data, *Computing and Informatics*, Vol. 30, pp. 1001-1026, 2011.
- [25] Ahmed M. Khedr and Salim A., Decomposable Algorithms for Finding the Nearest Pair, *J. Parallel Distrib. Comput.*, Vol. 68, pp. 902-912, 2008.
- [26] Ahmed M. Khedr, Learning k-Classifer from Distributed Databases, *Computing and Informatics Journal*, Vol. 27, pp. 355-376, 2008.
- [27] Ahmed M. Khedr and Bhatnagar, R., Agents for Integrating Distributed Data for Complex Computations, *Computing and Informatics Journal*, vol. 26, No.2, pp. 149-170, 2007.
- [28] Ahmed M. Khedr and Mahmoud, Rania. Agents for integrating distributed data for function computations. *Computing and Informatics*. 31. 1101-1125, 2012.
- [29] Ahmed M. Khedr, Decomposable Algorithm for Computing k-Nearest Neighbors across Partitioned Data, in: *International Journal of Parallel, Emergent and Distributed Systems*, vol. 31, no. 4, pp. 334-353, 2016.
- [30] Ahmed M. Khedr, Walid Osamy, Ahmed Salim and Abdel-Aziz Salem, Privacy Preserving Data Mining Approach for IoT based WSN in Smart City, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 8, 2019.