

# Machine Learning Based Classification Models for Financial Crisis Prediction



S. Anand Christy, R.Arunkumar

**Abstract:** Financial Crisis Prediction (FCP) being the most complicated and expected problem to be solved from the context of corporate organization, small scale to large scale industries, investors, bank organizations and government agencies, it is important to design a framework to determine a methodology that will reveal a solution for early prediction of the Financial Crisis Prediction (FCP). Earlier methods are reviewed through the various works in statistical techniques applied to solve the problem. However, it is not sufficient to predict the results with much more intelligence and automated manner. The major objective of this paper is to enhance the early prediction of Financial Crisis in any organization based on machine learning models like Multilayer Perceptron, Radial basis Function (RBF) Network, Logistic regression and Deep Learning methods and conduct a comparative analysis of them to determine the best methods for Financial Crisis Prediction (FDP). The testing is conducted with globalized benchmark datasets namely German dataset, Weislaw dataset and Polish Dataset. The testing is performed in both WEKA and Rapid Miner Framework design and obtained with accuracies and other performance measures like False Positive Rate (FPR), False Negative Rate (FNR), Precision, Recall, F-score and Kappa that would determine the best result from specific algorithm that will intelligently identify the financial crisis before it actually occurs in an organization. The results achieved the algorithms DL, MLP, LR and RBF Network with accuracies 96%, 72.10%, 75.20% and 74% on German Dataset, 91.25%, 85.83%, 83.75% and 73.75% on Weislaw dataset, 99.70%, 96.30%, 96.21% and 96.14 on Polish dataset respectively. It is evident from all the predictive results and the analytics in Rapid Miner that Deep Learning (DL) is the best classifier and performer among other machine learners and classifiers. This method will enhance the future predictions and would provide efficient solutions for financial crisis predictions.

**Keywords:** Financial Crisis Prediction; Machine learning; Artificial intelligence; Deep learning

## I. INTRODUCTION

Financial companies, corporate, borrowing firms as well as government agencies urge to design models to effectively investigate the possibility of counterparty default. Though default actions act in a stochastic manner, financial data can be employed to design financial crisis prediction (FCP) models.

For instance, [1], applied the multivariate statistic methodologies basically, discriminant analysis for classifying solvent and insolvent companies by exploiting financial data. The financial crisis happens not only because of bankruptcy and also due to the degrading of debt ratings of credit-related properties. Though default approaches have been used for the past years, the 2007/2008 financial crisis lead to the effective FCP models with utmost priority. But, [2] suggested that no standard theories or models exists for corporate FCP. The absence of theoretical model to investigate financial crisis for exploratory actions for the identification of discriminant features and prediction models using trial and error [3][4].

The academicians and professionals wanted to enhance the performance of FCP models by the use of diverse quantitative models. For example, [5] developed the earliest logistic regression (LR) approach for default computation.

Contrastingly, [1] provides a score to classify the observations as either good or bad customers; Ohlson's model computes the standard possibility of the significant. Assuming the relative ease of performing discriminant analysis and logistic regression, different works has been done to carry out identical tests. However, [6] disagreed that the famous Altman (1968) and Ohlson (1980) models are not precise and recommended the requirement of improvements in the modeling of default risks. Researchers discovered the artificial intelligence and ML approaches to measure credit risk using the recent technologies. As the investigation of financial crisis is identical to the pattern-recognition problems, methodologies can be employed for the classification of the creditworthiness, hence enhancing the conventional methods using earlier multivariate statistical methodologies like discriminant analysis and LR. Artificial neural networks (ANN) are also employed in various forms and the integration of ML algorithms in FCP is found to be interesting. Though numerous works has been investigated FCP by the use of recent techniques, [2] found that the results has not identifies the novel approach.

More number of FCP models are developed using the conventional statistical models and early artificial intelligence models. The key facts of this investigation are to examine the generous change in forecast exactness utilizing ML strategies compared to statistical models. This paper performs a comparative analysis of deep learning (DL), multilayer Perceptron (MLP), radial basis function (RBF) network and logistic regression (LR). For evaluation, three benchmark dataset namely German dataset, Weislaw dataset and Polish dataset. From experimentation, it is reported that the DL based classifier outperforms the other algorithms in terms of various performance measures.

Manuscript published on November 30, 2019.

\* Correspondence Author

**S. Anand Christy\***, Department of Computer and Information Science, Annamalai University, Chidambaram, Cuddalore, TamilNadu, India. Email: scholarchristy83@gmail.com

**Dr.R.Arunkumar**, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Cuddalore, TamilNadu, India. Email: arunkumar\_an@yahoo.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

# Machine Learning Based Classification Models for Financial Crisis Prediction

## II. RELATED WORKS

FCP using the past history of the financial data is an interesting topic. Several works has been done on the domain of FCP [31].Discriminant analysis and Logit

analysis are the widely used statistical models for FCP [32]. Altman Z-score [33] is most highly employed in this discriminant analysis.

A survey of diverse works on FCP is shown in Table 1.

**Table 1 A review of various studies on bankruptcy prediction and credit scoring**

Reference	Classifiers	Datasets	Evaluation methods				
			CRD	Accuracy	Type I/II error	F-Score	Kappa
[7]	MLP	Australian	No	Yes	No	No	No
[8]	MLP	Australian/German	No	Yes	No	No	No
[9]	MLP	US	No	Yes	No	No	No
[10]	MLP+LDA	Taiwan	No	Yes	Yes	No	No
[11]	MLP	Taiwan/US	No	Yes	No	No	No
[12]	MLP	Korea	No	Yes	No	No	No
[13]	MLP	Korea	No	Yes	No	No	No
[14]	MLP ensembles	Australian/German	No	Yes	No	No	No
[15]	MLP	Taiwan	No	Yes	Yes	No	No
[16]	MLP ensembles	Australian/ German	No	Yes	Yes	No	No
[17]	GA	Korea	No	Yes	No	No	No
[18]	GA	Korea	Yes	Yes	Yes	No	Yes
[15]	GA+SVM	Korea	No	Yes	No	No	No
[19]	LDA	Spanish	No	Yes	No	No	No
[20]	LDA	South African	No	Yes	Yes	No	No
[21]	PLS-DA	USA	No	Yes	Yes	Yes	No
[22]	LDA	USABD/JPNBD	No	Yes	No	No	No
[23]	B-CDT	Australian/ German/Japanese	No	Yes	No	No	No
[24]	WNN	Turkish/Spanish/US	No	Yes	No	No	No
[25]	SVM	Canada	No	Yes	No	No	No
[26]	DT, LR, SVM	US	No	Yes	No	No	No
[27]	MLP, DT, SVM, LR	Australian/German Japanese/Bankruptcy Data/Ucc	No	Yes	No	No	No
[28]	MDA, LR, CRT, and ANNs	Interfax SPARK database	No	Yes	Yes	Yes	No
[29]	DA, LR, RF, MLP,SVM, SOM	U.S. commercial banks	No	Yes	Yes	Yes	No
[30]	EXGB with other methods	Polish companies	No	Yes	No	No	No

Based on the above comparisons some issues are listed

- (1) Many existing approaches used only one dataset which reduces the reliability to find optimized solution. It is essential to use different datasets for properly evaluating the system performance.
- (2) Less number of approaches only uses type I and type II errors to analyze the average prediction performance of the models.

## III. ML BASED CLASSIFICATION ALGORITHMS

The Methodology is created as a framework for predicting the overall performance of machine learning algorithms and also to detect the best performance model among them. It is depicted in Fig.1.

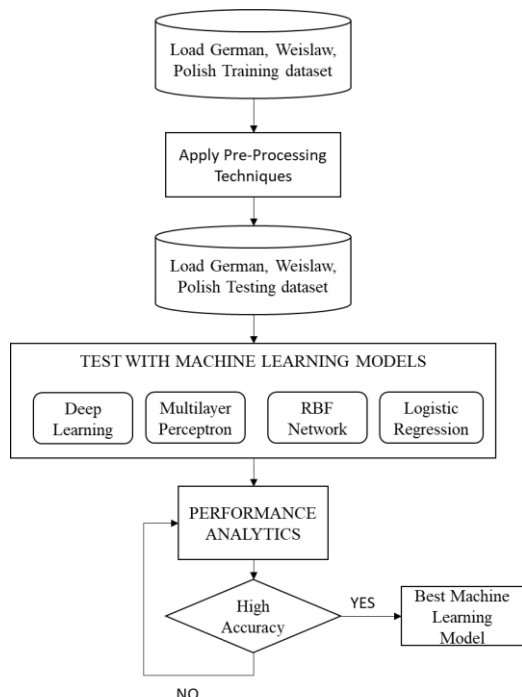


Fig. 1. Framework of Proposed Model

The initial training datasets of all three types of German, Weislaw and Polish datasets are loaded for pre-processing through WEKA to form the testing set. The testing dataset is experimented with machine learning models like Deep Learning (DL), Multi-Layer Perceptron (MLP), Linear Regression (LR) and Radial Basis function (RBF) Network respectively. The outcomes in terms of performances are analyzed and checked for accuracy levels. The model with best accuracy or high accuracy with all feature subsets are selected as the best model. The Machine learning models utilized for the research are analyzed and studied as given under.

**A. Deep Learning (DL)**

DL is a ML approach which allows the system to teach itself like the human beings, i.e. learn by experiences [34]. It is the fundamental logic behind the working of many fascinating applications like driverless car, voice automated mobile phones, TV and so on. Recently, DL is mainly employed in decision making situations to predict the occurrence at the earlier stages. In DL, the automated system performs the classification task by learning the straightly from images, text, or sound. It results to improved accuracy than the state of art methods and it might exceed manual results. DL performs training process using the labeled data and NN model contains more than one layer. The architecture of the DL model is given in Fig. 1. DL achieves better classification than existing methods and it helps users in various real-world applications.

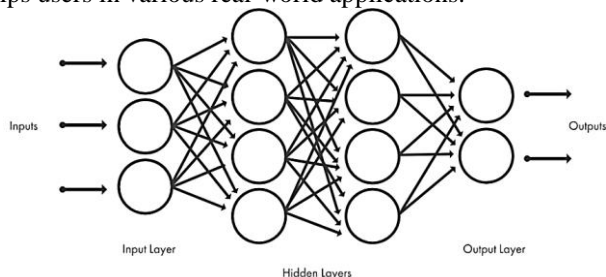


Fig. 2. NN structure

Since DL is developed in the year of 1980, it becomes familiar because of the following facts:

- It requires high quantity of labeled data
- It requires high computation power. The integration of DL with clusters or cloud, it allows the designer to significantly reduce the training time.

The usage of NN architectures in DL, it is sometimes called as deep neural networks (DNN). The word deep indicates the total of hidden layers in the NN. The traditional NN architecture contains 2-3 hidden layers and the DL has higher than 150 layers. During the training phase, large amount of labeled data is employed where the NN learns the features directly from the data rather than extracting features. This automatic nature of extracting features becomes highly beneficial for DL and enhances the classification performance. The overall process of the DL classifier is shown in Fig. 2.

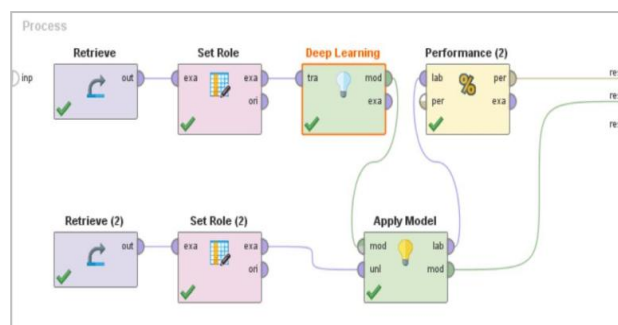


Fig. 3. Overall classification process of DL

**B. Logistic Regression (LR)**

LR [35] is a one of the conventional statistical approach employed for decision making processes. Different works ensured that these statistic models are found to be efficient than the recent ML approaches. The logistic distribution forms the basics of the logit model with its distribution function as indicated below.

$$F(X_i\beta) = \frac{\exp(X_i\beta)}{1+\exp(X_i\beta)} \tag{4}$$

and its density function is denoted in Eq. (5).

$$F(X_i\beta) = \frac{\exp(X_i\beta)}{[1+\exp(X_i\beta)]^2} \tag{5}$$

**C. Multilayer Perceptron (MLP)**

MLP is the NN based classification model which can be activated by loading the input layer with the input vector and then propagates the actions in a feed forward way using weighted connections in the entire network [36]. Let  $w_k$  be an input and the state of the  $i^{th}$  neuron ( $s_i$ ) can be calculated as

$$(w_{i,0} + \sum_{j \in P_i} w_{i,j} \times s_j) \tag{3}$$

where,  $f$  indicates the activation function,  $P_i$  is the collection of nodes which reaches node  $i$ ,  $w_{i,j}$  is the weight of the link between node  $i$  and  $j$ . MLP employs a repetitive function to learn data from the initialization of random weights. A training approach is utilized to manage the weights to a needed target value and the training will be ended when the error slop approaches to zero.



# Machine Learning Based Classification Models for Financial Crisis Prediction

Due to the advantages of MLP, it can provide approximate solutions for more complex problems like fitness approximation.

## D. Radial Basis Function (RBF)

Radial Basis Function (RBF) is a commonly used classifier because of its faster learning capability. It classifies the complex problems using a hidden layer whereas MLP contains more than one hidden layer. The training process in RBF network is very simple [37]. It employs Gaussian function as activation functions and least-square criterion as objective functions. It contains one input layer, one hidden layer and an output layer. The input layer computes the norm of the input from the neuron. The output of the  $j^{\text{th}}$  neuron can be represented in Eq. (5).

$$y_j = \sum_{k=1}^K c_k g_k(x) = \sum_{k=1}^K w_{kj} \exp\left(\frac{-||x-\mu_k||^2}{2\sigma_k^2}\right) \quad (5)$$

where, K indicates the number of Gaussian neurons,  $w_{kj}$  indicates the weight between the  $k^{\text{th}}$  Gaussian neuron and  $j^{\text{th}}$  output neuron,  $\mu_k$  is the centre location,  $\sigma_k$  is the width of the  $k^{\text{th}}$  Gaussian neuron and  $x$  is the input vector.

## III. PERFORMANCE ANALYSIS

For evaluating the individual performance of various classifiers on FCP, three benchmark dataset German dataset [38], Polish dataset [39] and Weislaw dataset [40] are used. The details of the dataset are tabulated in Table 2. The maximum number of instances in the dataset varies from 240 to 1000 whereas 20 to 64 attributes are present. The two class labels exist in the dataset are bankrupt and non-bankrupt.

Table 2 Dataset Description

Dataset	Source	# of instances	# of attributes	# of class	Bankrupt/Non-Bankrupt
German	UCI	1000	20	2	700/300
Weislaw	UCI	240	32	2	6756/271
Polish	UCI	7027	64	2	112/128

## A. Metrics

The measures used to analyze the classification results are false positive rate (FPR), false negative rate (FNR),

precision, recall, accuracy, F-score and kappa value. The formula of these measures is given in Table 3.

Table 3 Performance measures

Measure	Mathematical formula
FPR	$FP/FP + TN$
FNR	$FN/FN + TN$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$
F-score	$2TP / (2TP + FP + TN)$
Kappa	$\frac{Obs. Agreement - Exp. Agreement}{100 - Exp. Agreement}$

analysis states that the DL classifier attains better performance than other classifiers.

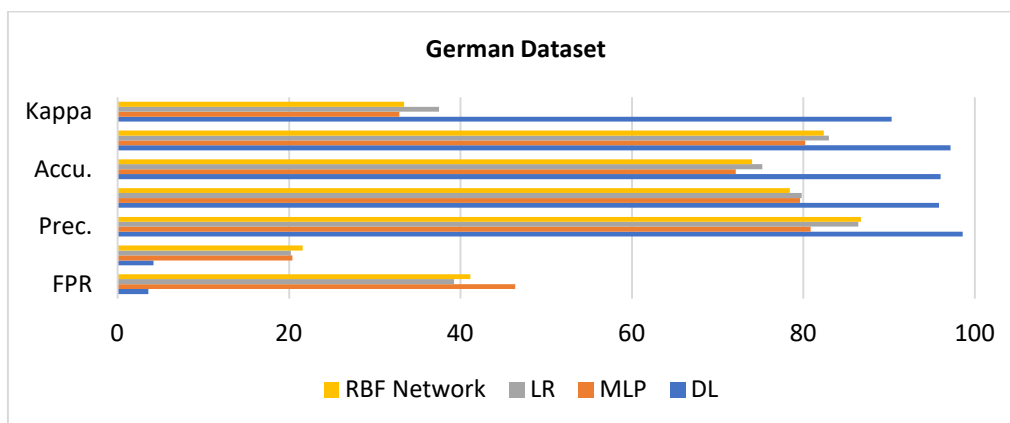
## IV. RESULT ANALYSIS

This subsection compares the results of four classifiers against three benchmark datasets. The experimental

Table 4 Performance Evaluation of German Dataset using various Classifiers

Classifier	FPR	FNR	Prec.	Rec.	Accu.	F-score	Kappa
DL	3.57	4.16	98.57	95.83	96.00	97.18	90.29
MLP	46.36	20.39	80.85	79.60	72.10	80.22	32.87
LR	39.25	20.18	86.42	79.81	75.20	82.99	37.50
RBF Network	41.15	21.57	86.71	78.42	74.00	82.36	33.40

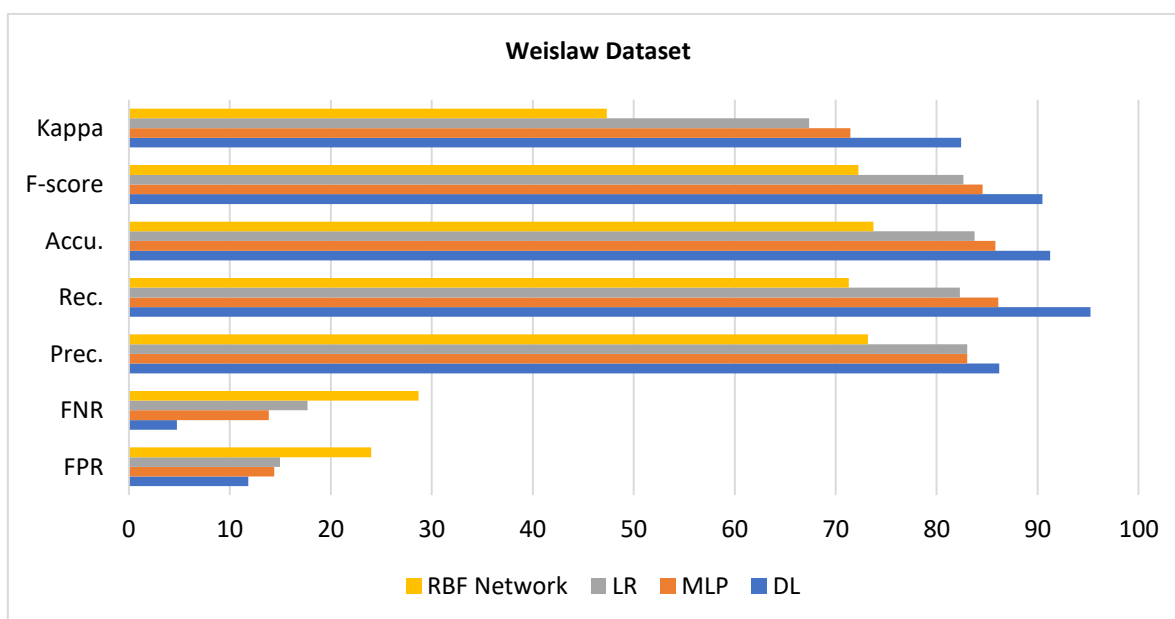




**Fig. 4 Performance Evaluation of German Dataset using various Classifiers**

From Table 4 and Fig. 3, it is clear that DL achieves better results than other classifiers and attains an accuracy of 96%. It is noted that MLP classifier produces poor results with a lower accuracy of 72.10. Next, RBF network achieves 75.20 accuracy which is better than the accuracy attained by MLP. Though LR produce accuracy of 75.20% which is higher than other classifiers, but it is not efficient than DL model. In terms of F-score, MLP achieves worse performance with a lowest value of 80.22 whereas the RBF classifier obtains 74 which are better than MLP. But, LR achieves an F-score of 82.99 whereas DL model attains a value of 97.18 respectively.

Classifier	FPR	FNR	Prec.	Rec.	Accu.	F-score	Kappa
<b>DL</b>	11.85	4.76	86.2	95.23	91.25	90.49	82.42
<b>MLP</b>	14.39	13.88	83.03	86.11	85.83	84.54	71.47
<b>LR</b>	14.96	17.69	83.03	82.3	83.75	82.66	67.37
<b>RBF Network</b>	24	28.69	73.21	71.3	73.75	72.24	47.35



**Fig. 5. Performance Evaluation of Weislaw Dataset using various Classifiers**

From Table 5 and Fig. 4, it is showed that DL attains better results than other classifiers for Weislaw dataset. DL attains an accuracy of 91.25. It is noted that RBF classifier produces poor results than other classifiers. Next, LR achieves 83.75% accuracy which is better one among RBF but lesser than MLP and DL classifiers. Though MLP produces an accuracy of 85.83 which is higher than other classifiers, but it is not efficient than DL. With respect to FPR and FNR, the DL attains lowest value of 11.85 and

4.76, which implies the effective performance of DL classifier. At the same time, the RBF attains highest value of FPR and FNR value of 24 and 28.69 which indicates the worse classification performance. In the same way, the order of effective classification performance in terms of precision and recall is DL, MLP, LR and RBF. From the above discussion, it is clear that DL is the better model for FCP on Weislaw dataset.

Table 6: Performance Evaluation of Polish Dataset using various Classifiers

Classifier	FPR	FNR	Prec.	Rec.	Accu.	F-score	Kappa
DL	0	0.30	100	99.69	99.70	99.84	95.81
MLP	32.25	3.57	99.85	96.42	96.30	98.10	13.22
LR	45.45	3.45	99.63	96.54	96.21	98.06	17.32
RBF Network	0	3.85	100	96.14	96.14	98.03	0

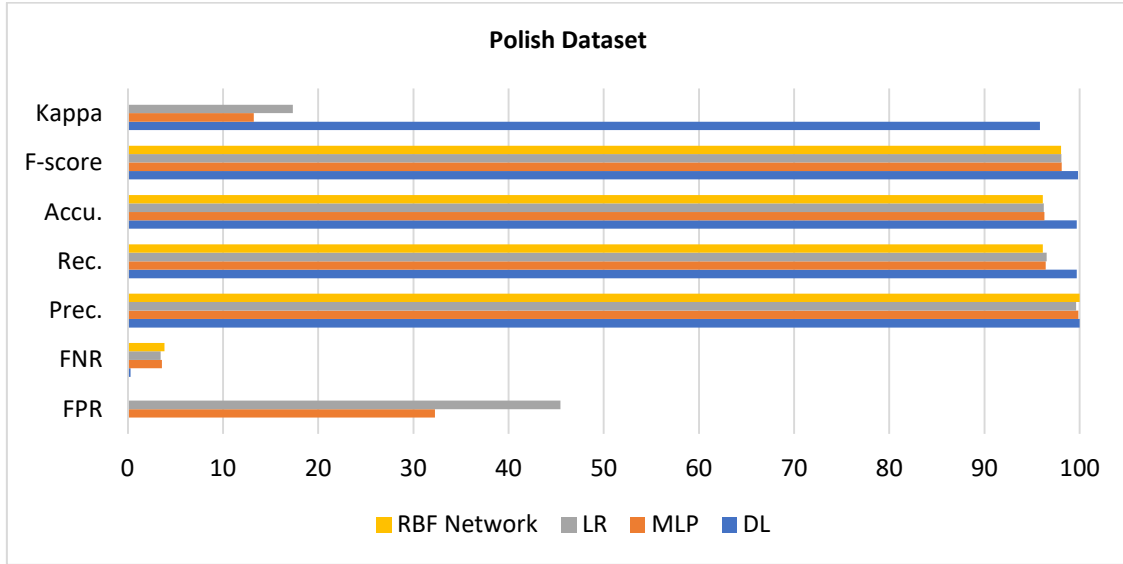


Fig. 6 Performance Evaluation of Polish Dataset using various Classifiers

Table 6 and Fig. 5 shows the classification results of different classifiers using Polish dataset. From the table, it is clear that the DL obtains the maximum classification results with an accuracy of 99.70 whereas other classifiers like MLP, LR and shows competitive results with an accuracy of 96.30, 96.21 and 96.14 respectively. At the same time, it is noted that the DL attains a minimum of zero FPR value and 0.30 FNR value. In contrast, the MLP and LR attained a

maximum FPR value of 32.25 and 45.45 respectively. Next, the kappa value indicates the perfect measure of agreement between expert’s opinion and classification algorithm. The highest kappa value of DL is 95.81 whereas the RBF network attains a lowest value of zero. In addition, the MLP and LR attain a minimum value of 13.22 and 17.32 respectively.

Table 7: Overall classification of Machine Layer Performances with different datasets

Classifier	Polish Dataset	Weislaw Dataset	German Dataset
DL	99.70	91.25	96.00
MLP	96.30	85.83	72.10
LR	96.21	83.75	75.20
RBF Network	96.14	73.75	74.00

From the overall results, it is concluded that the DL classifier is found to be efficient than other classifiers.

V. CONCLUSION

The Research study focused on the early prediction of Financial crisis Prediction (FCP) with maximum accuracy and performance levels using machine learning models in a successful manner. Initially the global bankruptcy datasets from German, Weislaw and Polish datasets are pre-processed and converted to testing sets after removal of outliers and normalisation process. The dataset is undergone

with comparative predictive analytics on various Machine Learning models like Deep Learning (DL), Multi-Layer Perceptron (MLP), Linear Regression (LR) and Radial Basis function (RBF) Network respectively. The Prediction is carried out in Rapid Miner and WEKA to provide solution for dataset with high accuracy. The results are obtained on the basis of performance measures like False Positive Rate (FPR), False Negative Rate (FNR), Precision, Recall, F-score and Kappa.

The final outcomes of Rapid Miner indicated that Deep Learning with 96%, 91.25 and 99.70% has outperformed other models where the results are very less compared to it. The other performance measures are also convincing for deep learning methods. Thus, the overall conclusion is derived as deep learning is the best model that is capable of detecting the outcome of the financial crisis prediction at the maximum level. This result will enhance further predictions that might be implemented in Artificial Intelligence with Internet of Things to automatically detect the crisis that might occur over a period of time. This research also suggested that other common supervised and unsupervised machine learning algorithms needs to be enhanced on the basis of performance to match the levels of deep learning techniques.

## REFERENCES

1. El, "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks," *J. Bank. Financ.*, 1994.
2. G. Wang, J. Ma, and S. Yang, "An improved boosting based on feature selection for corporate bankruptcy prediction," *Expert Syst. Appl.*, vol. 41, no. 5, pp. 2353–2361, 2014.
3. J. Sun and H. Li, "Financial distress prediction using support vector machines: Ensemble vs. individual," *Appl. Soft Comput. Vol.*, vol. 12, no. 8, pp. 2254–2265, 2012.
4. L. Zhou, "Knowledge-Based Systems Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods," vol. 41, pp. 16–25, 2013.
5. J. A. Ohlson, "Financial ratios and the probabilistic prediction of bankruptcy," *J. Account. Res.*, vol. 18, no. 1, pp. 109–131, 1980.
6. Serrano-Cinca and B. Gutiérrez-Nieto, "Partial least square discriminant analysis for bankruptcy prediction," *Decis. Support Syst.*, vol. 54, no. 3, pp. 1245–1255, 2013.
7. Fan and M. Palaniswami, "Selecting bankruptcy predictors using a support vector machine approach," ... , 2000. *IJCNN 2000, Proc. ....*, no. table 2, pp. 354–359, 2000.
8. West, "Neural network credit scoring models," *Comput. Oper. Res.*, vol. 27, no. 11–12, pp. 1131–1152, 2000.
9. F. Atiya and S. Member, "Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results," vol. 12, no. 4, pp. 929–935, 2001.
10. T. Lee, C. Chiu, and C. Lu, "Credit scoring using the hybrid neural discriminant technique....file danneggiato, guardare nella cartella origine," *Expert Syst. Appl.*, vol. 23, pp. 245–254, 2002.
11. 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," *Decis. Support Syst.*, vol. 37, no. 4, pp. 543–558, 2004.
12. J. H. Min and Y. C. Lee, "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters," *Expert Syst. Appl.*, vol. 28, no. 4, pp. 603–614, 2005.
13. K.-S. Shin, T. S. Lee, and H. Kim, "An application of support vector machines in bankruptcy prediction model," *Expert Syst. Appl.*, vol. 28, no. 1, pp. 127–135, 2005.
14. West, S. Dellana, and J. Qian, "Neural network ensemble strategies for financial decision applications," *Comput. Oper. Res.*, vol. 32, no. 10, pp. 2543–2559, 2005.
15. S.-H. Min, J. Lee, and I. Han, "Hybrid genetic algorithms and support vector machines for bankruptcy prediction," *Expert Syst. Appl.*, vol. 31, no. 3, pp. 652–660, 2006.
16. F. Tsai, "Financial decision support using neural networks and support vector machines," *Expert Syst.*, vol. 25, no. 4, pp. 380–393, 2008.
17. K.-S. Shin and Y.-J. Lee, "A genetic algorithm application in bankruptcy prediction modeling," *Expert Syst. Appl.*, vol. 23, no. 3, pp. 321–328, 2002.
18. M. J. Kim and I. Han, "The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms," *Expert Syst. Appl.*, vol. 25, no. 4, pp. 637–646, 2003.
19. Alfaro, N. García, M. Gámez, and D. Elizondo, "Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks," *Decis. Support Syst.*, vol. 45, no. 1, pp. 110–122, 2008.
20. Altman, G. Marco, and F. Varetto, "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)," *J. Bank. Financ.*, vol. 18, no. 3, pp. 505–529, 1994.
21. Tsai and J. Wu, "Using neural network ensembles for bankruptcy prediction and credit scoring," *Expert Syst. Appl.*, vol. 34, no. 4, pp. 2639–2649, 2008.
22. L. Zhou, "Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods," *Knowledge-Based Syst.*, vol. 41, pp. 16–25, 2013.
23. J. Abellán and C. J. Mantas, "Improving experimental studies about ensembles of classifiers for bankruptcy prediction and credit scoring," *Expert Syst. Appl.*, vol. 41, no. 8, pp. 3825–3830, 2014.
24. N. Chauhan, V. Ravi, and D. Karthik Chandra, "Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks," *Expert Syst. Appl.*, vol. 36, no. 4, pp. 7659–7665, 2009.
25. T. Xiong, S. Wang, A. Mayers, and E. Monga, "Personal bankruptcy prediction by mining credit card data," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 665–676, 2013.
26. L. Olson, D. Delen, and Y. Meng, "Comparative analysis of data mining methods for bankruptcy prediction," *Decis. Support Syst.*, vol. 52, no. 2, pp. 464–473, 2012.
27. C. F. Tsai and K. C. Cheng, "Simple instance selection for bankruptcy prediction," *Knowledge-Based Syst.*, vol. 27, pp. 333–342, 2012.
28. Fedorova, E. Gilenko, and S. Dovzhenko, "Expert Systems with Applications Bankruptcy prediction for Russian companies: Application of combined classifiers," vol. 40, pp. 7285–7293, 2013.
29. J. L. Iturriaga and I. P. Sanz, "Expert Systems with Applications Bankruptcy visualization and prediction using neural networks: A study of U. S. commercial banks," vol. 42, pp. 2857–2869, 2015.
30. M. Zi, S. K. Tomczak, and J. M. Tomczak, "Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction," vol. 58, pp. 93–101, 2016.
31. Sartori, A. Mazzucchelli, and A. Di Gregorio, "Bankruptcy forecasting using case-based reasoning: The CRePERIE approach," *Expert Syst. Appl.*, vol. 64, pp. 400–411, 2016.
32. M. Tseng and Y. C. Hu, "Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks," *Expert Syst. Appl.*, vol. 37, no. 3, pp. 1846–1853, 2010.
33. I.-P. Muhammad, "Business bankruptcy prediction models: A significant study of the Altman's Z-score model," *Asian J. Manag. Res.*, vol. 3, no. 1, pp. 212–219, 2012.
34. J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
35. J. H. Xue and D. M. Titterton, "Comment on 'on discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes,'" *Neural Process. Lett.*, vol. 28, no. 3, pp. 169–187, 2008.
36. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators. 2," *Neural Networks*, p. 359–366, 1989.
37. M. J. D. Powell, "Radial basis functions for multi-variable interpolation: A review," in *IMA Conference on Algorithms for the Approximation of Functions and Data*, RMCS Shrivvenham, UK, 1985.
38. Hans Hofmann, "German Credit Data," 1996. [Online]. Available: [https://archive.ics.uci.edu/ml/datasets/statlog+\(german+credit+data\)](https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data)).
39. S. Tomczak, "Polish companies bankruptcy data Data Set," UCI Machine Learning Repository. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Polish+companies+bankruptcy+data>.