Unveiling the diverse efficacy of artificial neural networks and logistic regression: A comparative analysis in predicting financial distress

Amine Sabek
Investment bets and sustainable development stakes in border areas, University of Tamanghasset, B.P 10034 Tamanghasset Airport road, Algeria, sabek.amine@univ-tam.dz

Abstract
The prediction of financial distress has emerged as a significant concern over a prolonged period spanning more than half a century. This subject has garnered considerable attention owing to the precise outcomes derived from its predictive models. The main objective of this study is to predict financial distress using two types of Artificial Neural Networks (ANN) compared to the Logistic Regression (LR), and this will be done by relying on the data of 12 Algerian companies for the period 2015-2019. The reason for choosing these two types of networks in particular, is attributed to the fact that Elman Neural Network (ENN) is commonly used network, in contrast to the Feed-forward Distributed Time Delay Neural Network (FFDTDNN). Regarding the choice of these companies as a study sample, can be attributed to the similarity in the temporal range covered by their financial statements, coupled with their approximate parity in terms of asset size. This study concluded that the ENN model outperformed the LR model in predicting financial distress with a classification accuracy of 100%. On the other hand, the LR model outperformed the FFDTDNN with a classification accuracy of 83.33%. Therefore, it can be asserted that ANNs cannot be regarded as superior to Logistic Regression (LR) in all statuses. Instead, it is accurate to affirm that specific types of ANNs exhibit greater efficacy than LR in predicting financial distress, while other types demonstrate relatively diminished effectiveness.

Keywords: artificial neural network, financial distress, logistic regression, prediction.

JEL classification: C45, C53, G17, G33.
DOI: 10.2478/crebss-2023-0002

Received: May 04, 2023
Accepted: July 04, 2023

Introduction

There are multiple definitions of financial distress, which include various financial situations. Numerous studies have been conducted on this subject during the last 50 years. Carmichael described financial distress as a company's inability to satisfy its obligations. He includes insufficient liquidity, insufficient capital, nonpayment of debt (Paule-Vianez et al., 2019). Financial distress is a situation in which a company is unable to meet or has difficulty paying off financial obligations to its creditors, which may result in the enterprise's bankruptcy, so accurate financial distress prediction models have a significant influence on various corporate stakeholders in the decision-making process (Xie et al., 2011). When a company experiences a temporary lack of liquidity and the difficulties that follow in meeting its financial obligations on time and in full, it is said to be in "financial distress." Financial distress can take one of two forms: either the company defaults on a debt payment or makes an effort to restructure its debt in order to avoid the default situation (Shisia et al., 2014). It could also be described as the final phase of organizational decline before bankruptcy. As a result, financial hardship is distinct from bankruptcy because it describes a time when a borrower is unable to fulfill a debt to creditors, but bankruptcy is an official statement of a firm's financial situation in which it may stop operating or restructure. Bankruptcy may result when financial trouble is not addressed, but it is not a given (Puro, 2019). Last but not least, financial distress is defined as having negative earnings per share for listed companies, occurring when a company's interest cover is lower than 0.7, there is a decline in fixed assets or a decrease in share capital, as well as when a company's net worth falls below half of its share capital (Zhiyong, 2014). Although many people use the term bankruptcy to refer to a failed business, a firm is not legally bankrupt until it has been declared bankrupt by applicable law (Mayliza et al., 2020).

Financial distress prediction has now become an absolute necessity for all companies in different fields of activity, as it helps to avoid many potential financial risks that may lead to the end of the company's activity, and unlike the developed world countries, Algeria is still suffering from the financial distress effects in terrible silence without moving a finger to deal with this phenomenon seriously enough. Until now, the financial managers of these companies are still avoiding the idea of relying on the modern financial analysis methods and dispensing with traditional methods that have lost their effectiveness compared to the results achieved by relying on various statistical and artificial intelligence techniques (International Monetary Fund, 2014).

The growing interest in financial distress prediction among companies operating in developed countries stems from the positive outcomes yielded by various modern forecasting models. These models have enabled companies to avert additional costs and losses that could have severely disrupted their financial systems. It is important to note that the consequences of such imbalances extend beyond distressed companies, affecting other entities within the state's economy due to shared interests. Consequently, the implications gradually impact the overall state economy, leading to potential bankruptcies and an influx of debts as the state intervenes to assist distressed or bankrupt companies within its jurisdiction. Given the increasing importance of forecasting financial distress, it has become a prominent topic within the fields of finance and accounting. Consequently, this study holds significance as it aims to urge companies, particularly those operating in non-developed countries, to take financial distress prediction seriously rather than relying on chance when managing their financial aspects.

To keep pace with the evolving landscape, the techniques employed in predicting financial distress continue to advance. Initially, traditional financial analysis methods constituted the primary means of diagnosing financial distress. However, contemporary methods now rely on a combination of statistical techniques and artificial intelligence, which have proven to possess robust classification and prediction abilities compared to the statistical techniques (Odom, Sharda, 1990; Wilson, Sharda, 1994; Jo et al., 1997; Min, Lee, 2005; Salehi, Pour, 2016; Zieba et al., 2016; Belas et al., 2017; Karas, Reznakova, 2017).

One could contend that Logit is the preeminent method widely employed for forecasting corporate distress. Although Logit demonstrates a higher rate of success compared to MDA, it has yet to reach the same levels of achievement as NN (Callejón et al., 2013). These techniques have garnered significant attention from financial researchers, who recognize their efficacy in various scientific domains and their potential applicability in the financial realm.

The primary objective of this paper is to build new models utilizing different combinations of carefully selected financial ratios. These models aim to facilitate a comparative analysis between artificial intelligence and statistical methods, specifically examining artificial neural networks against logistic regression. As we mentioned earlier, all previous studies that compared ANNs and statistical techniques concluded that the first technique is better than the second in classification, and to the best of our knowledge, we did not find a comparative study that proves that LR can be better than ANNs in predicting financial distress. By doing so, this study seeks to determine whether all types of ANNs outperform logistic regression in classification tasks in general and, specifically, in predicting financial distress?

Literature review

A neural network is an intricate framework engineered to mimic the cognitive processes of the human brain in executing a particular activity or function that holds significance. Typically, this network is constructed using electronic components or emulated through software on a digital computer (Hardinata, Warsito, 2018). Artificial neural networks (ANNs) can be described as highly simplified models of the intricate interactions among brain cells. They are recognized as significantly simplified models of the human nervous system, possessing notable capabilities such as learning, generalization, and abstraction. Nevertheless, advancements in technology have recently rendered ANN models a feasible alternative for addressing various decision problems, presenting promising prospects for enhancing models related to financial activities. One such area where ANN models hold potential is in the realm of forecasting financial distress within corporations (Sudarsanam, 2016). The ANN is a nonparametric modeling tool renowned for its adaptability. It possesses the ability to accurately map intricate functions. Typically, an ANN comprises multiple layers housing numerous computing elements referred to as nodes. Each node acquires input signals from other nodes. After performing localized signal processing through a transfer function, the node transmits the transformed signal to other nodes or yields the final outcome (Zhang et al., 1999). The functioning of an ANN is contingent upon the combination of weights and the input-output for the units. These functions can be categorized into three distinct types: sigmoid, threshold, and linear. The selection of a neuron's transfer function is based on the desire to enhance or streamline the network's overall performance (Ibiwoye et al., 2012). The ANN comprises interconnected nonlinear nodes that engage in parallel communication. The connection weights are modifiable, thereby enabling the ANN to acquire knowledge directly from exemplars, eliminating the need for an analytical approach to problem-solving. The predominant modes of learning within ANN encompass unsupervised learning and supervised learning (Chen, Du, 2009). While ANN ensembles possess advantageous qualities such as efficiency, robustness, and adaptability, rendering it a valuable asset in classification, decision support, financial analysis, and credit scoring tasks, certain researchers have demonstrated that ensembles consisting of multiple neural network classifiers do not consistently outperform a singular, optimal neural network classifier (Tang et al., 2019).

Statistical methods, originally developed for analyzing limited data sets, are inadequate for handling large volumes of data and intricate associations. The key crucial factors of machine learning algorithms compared to conventional computer algorithms lie in their capacity to enhance performance and exhibit dynamic transformations through the incorporation of fresh data. Machine learning algorithms are categorized into various divisions based on their specific purposes, such as classification, clustering, pattern recognition, and correlation analysis. These algorithms can be employed to predict financial distress or failure (Ağdeniz, Yıldız, 2016). Several scholarly studies have successfully utilized NN methodologies to tackle various challenges related to classification. In addition to bankruptcy, the implementation of NN techniques has been applied to address a range of issues, including capital structure, ineffective management, adverse economic impacts, and volatility (Callejón et al., 2013). ANNs possess suitable capabilities that render them highly effective in predicting bankruptcy. Their remarkable attributes, including their adeptness in handling non-linear data and their ability to self-learn, contribute to their widespread utilization in financial contexts. A comprehensive survey addressing bankruptcy prediction is presented, focusing primarily on ANN models. Furthermore, recent advancements have explored various intelligent techniques such as rough sets, case-based reasoning, fuzzy set theory, support vector machines, decision trees, and, which have received considerable attention and extensive examination in this domain (Ribeiro et al., 2010). The Elman neural network belongs to the category of feedback neural networks and builds upon the BP ANN by integrating an extra hidden layer called the "undertake" layer. This specific layer operates as a delay operator, empowering the network with memory capabilities. As a result, the network system acquires the ability to effectively adjust to time-dependent dynamic characteristics while ensuring resilient overall stability (Jia et al., 2014). Time Delay Neural Networks (TDNNs) present a significant improvement compared to traditional multi-layer perceptrons through the integration of time-delayed connections. These networks are specifically designed for tackling sequence recognition tasks that necessitate only a limited recollection of past events.

The training of TDNNs commonly involves the extensive incorporation of delayed connections over time, effectively converting the training procedure into that of a feedforward network (Cancelliere, Gemello, 1996). Logistic Regression (LR) is a member of the generalised linear models family, GLMs offer a consistent framework for modelling response from any exponential family distribution, such as Gaussian, Binomial, or Poisson. In GLM framework, the model is quantified from a binary target variable, Y which represents the status of a loan over the outcome window where bad (defaulted) loan is labelled as 0 and good loan is labelled as 1 and related to the linear combination of predictor $\beta 1 * X1 + \ldots + \beta m * Xm$ (Bayraci, Susuz, 2019). LR is widely employed within the financial sector for the development of credit risk assessment models, making it a prominent statistical technique. The LR model offers several notable benefits, including its reliable performance, ease of understanding, and simple applicability. Additionally, LR outperforms linear regression due to its ability to overcome specific challenges encountered in the latter approach. Notably, linear regression may produce negative or probability values exceeding 1, which contradicts the fundamental nature of probability (Uzair et al., 2019). In 1980, Ohlson (1980) conducted a research study employing "Logit" or Multiple LR to construct a model for forecasting bankruptcy. Ohlson asserted that his study offered a notable advantage by enabling the identification of whether a company would declare bankruptcy prior to or subsequent to the disclosure of financial statements. Ohlson further noted that preceding studies did not explicitly tackle the aspect of timing in their analyses. While contemporary intelligence techniques have gained significant popularity in recent years, it is noteworthy to acknowledge the enduring relevance of pioneer statistical methods, such as discriminant analysis, in the realm of corporate bankruptcy prediction modeling. Linear classification algorithms, including linear discriminant analysis and LR, are widely favored in this domain. All of these approaches strive to identify the optimal linear amalgamation of explanatory input variables (Ribeiro et al., 2010).

Since the 1930s, scholars have undertaken numerous trials aimed at assessing a range of plausible methodologies in order to address the demand for accurate predictions. The outcomes of these experiments have substantially enhanced our comprehension of the field of forecasting (Osho, Idowu, 2018). The financial distress prediction is of significant interest to the diverse set of stakeholders associated with the company, encompassing regulators, creditors, investors, and lenders. Particularly, stakeholders who hold company shares within their derivatives portfolio require timely access to this information to assess the likelihood of financial distress (Kapil, Agarwal, 2019). The primary objective of a Financial Distress Prediction Model is to forecast the likelihood of future financial distress for a company. The conventional statistical models were discriminant analysis and the logit model. However, these traditional linear techniques, though straightforward, lack practicality, rendering them inadequate for developing a robust model capable of generating real-time predictions (El-Bannany et al., 2020).

There exist two categories of failure prediction models, namely statistical-based models and algorithm-driven models employing Machine Learning (ML) techniques. Initially, statistical methods were employed by early researchers in the field of bankruptcy prediction. Despite the continued utilization of statistical methods, certain scholars have embraced the application of Machine Learning techniques, including neural networks, to forecast corporate failures (Bonello et al., 2018). The onset of the second phase, which commenced in the late 1980s, marked a pivotal moment wherein numerous scholars embarked on investigations aimed at ascertaining the efficacy of non-parametric methodologies in prognosticating bankruptcy risk. Notably, this period witnessed the rise of non-linear techniques such as, Support Vector Machines, Artificial Neural Networks, Nave Bayesian Classifier, and k-Nearest Neighbor which consistently outperformed the prevailing methods of the time (Mousavi et al., 2012). Both statistical methods and artificial intelligence methods have distinct advantages and disadvantages. Multiple discriminant analysis (MDA), a commonly used statistical technique, provides a clear advantage in terms of interpretability. However, its application is limited by strict statistical assumptions, and it functions as a fixed determination model. In contrast, artificial intelligence methods like the back-propagation neural network do not rely on assumptions about probability distributions, making it valuable for modeling non-linear systems. As a result, many researchers choose to utilize three hidden layer backpropagation neural networks for predicting financial crises (Cheng et al., 2018). The pioneering research conducted by (Odom, Sharda, 1990) marked the inception of bankruptcy prediction employing NNs. Their study involved the development of a neural network model aimed at forecasting bankruptcy. Furthermore, the research presented a comparative analysis of two approaches, namely discriminant analysis and the neural network method, to assess their respective predictive capabilities. The outcomes demonstrated the neural network's proficiency in generating precise forecasts. A dataset comprising 165 pairs of firms was utilized by (Koh, Tan, 1999) to construct a NN model incorporating six financial ratios. Remarkably, the NN model accurately predicted all of the examined cases. These findings imply that neural networks possess considerable potential for both research and practical implementation, suggesting a promising avenue for future exploration. Zhou et al. (2010) sought to investigate the comparative performance of Support Vector Machine (SVM) and Logistic regression in accurately predicting instances of individual loan defaults. The researchers found that SVM surpasses the Logistic regression model in effectively fitting intricate feature patterns without the need for incorporating high-power features, thereby demonstrating its superior ability to accurately predict loan defaults. In a research undertaken by Obare and Muraya (2018) an investigation was carried out to assess the effectiveness of support vector machine and logistic regression models in forecasting individual loan defaults in Kenya. The results demonstrated that the support vector machine model exhibited superior performance in predictive accuracy compared to the logistic regression model. Barboza et al. (2017) present an extensive examination of statistical methods, namely Linear Discriminant Analysis, and machine learning approaches, specifically Support Vector Machine. The study also encompasses a comparative analysis of these techniques. Gregova et al. (2020) conducted a comparative examination of models generated using three different approaches (LR, random forest, and neural network models) with the aim of determining the model that displayed the highest predictive accuracy in detecting financial distress. The results revealed that all models demonstrated significant discrimination accuracy and performed similarly. Nevertheless, neural network models exhibited superior outcomes across all performance measures.

Lahmiri and Bekiros (2019) introduced the Generalized Regression Topology as a rapid approach for identifying the optimal neural architecture. On the other hand, within the realm of Machine Learning, logistic regression techniques were employed for predictive purposes. These techniques relied upon an Artificial Neural Network comprising a single layer, evolved through the incorporation of financial ratios, in order to ascertain the most optimal configuration for the neural network (Mukeri et al., 2020).

Methodology

This paper aims to compare artificial intelligence with statistics. To achieve this objective, Neural Networks and Logistic Regression were compared in order to identify the optimal model for predicting financial distress in Algerian companies. We chose two types of neural networks, Elman NN, the commonly used network (Jia et al., 2014), and the Feed-forward Distributed Time Delay NN, and to the best of our knowledge, this network is uncommonly used.

Artificial neural network was chosen as the test subject in this study because it is commonly used for financial distress prediction among the other intelligent models, and it is considered the leading technique. The same applies to logistic regression for statistical models. Since ANNs were able to outperform LR in most cases according to previous studies (We referred to this earlier with citations), we aspire to verify, are all types of artificial neural networks really better than logistic regression in classification in general, and in predicting financial distress in particular?

Sample

The analysis could be extended to a larger sample of companies from more countries. We relied on a small sample because of the difficulty of obtaining the financial statements of Algerian companies. However, we compensated for this by extending the study period to five years (2015-2019), with the aim of increasing the number of financial cases. Therefore, we used a data of (12) Algerian economic companies, where the total of financial cases reached 60 cases. The financial data is divided into a training sample which concerns data of the period between (2015-2017). 36 financial cases were allocated to this sample, divided into 6 cases of financial distress and 30 cases of non-financial distress. The test sample that concerns data of the period between (2018-2019) for the purposes of evaluating the efficiency level of the models. 24 financial cases were allocated to this sample, divided into 8 cases of financial distress and 16 cases of non-financial distress.

Among the disadvantages of using a small sample is the results cannot be generalized, and relying on a large sample will reduce the accuracy of the best model. However, the same will apply to the less accurate models. Therefore the matter will be relatively from the researcher's point of view. The larger the sample, the lower the accuracy of the best model, and with it the accuracy of other models. We used a set of data belonging to different economic sectors. Among these companies, (4) are listed on the Algerian Stock Exchange, whereas, the other (7) companies, their data were obtained by resorting to a competent official Algerian authority after undertaking not to mention the names of these companies in order to preserve their confidentiality and privacy. Finally, Sycma's company data was obtained after submitting an official request to its financial department.

Distress dataset description

Table 2 shows the actual financial status of the companies under study for the period between (2015-2019), where the description of the financial status differed between distress (D) and non-distress (N-D), which allows the models to understand, comprehend, and determine the status of distress or not based on the financial ratios that were previously selected with high accuracy. The aim of the training process is to make the models predict the company's actual status with the least possible degree of error, which we will confirm after testing the predictive ability of the three models using the test sample.

It should also be noted that the actual financial status was assessed according to the self evaluation based on the indications of the financial ratios, according to the Table 1.

Most of the previous studies relied on the financial ratios in building the prediction models using the variables from the capital, assets, management, earnings, liquidity, sensibility. In this paper, we relied on 21 financial ratios, which proved their great ability to predict financial distress, according to previous studies.

Table 1 Actual financial status's Indicators

Indicator	Distress	Non-Distress
Net income/Assets	Negative	Positive
Current Assets/Current Liabilities	Less than 1	Greater than 1
Debts/Assets	Greater than 1	Less than 1
Sales/Assets	Low	High
Cash/Assets	Less than 0	Greater than 0
Current Assets/Assets	Low	High
Net income/Sales	Less than 0	Greater than 0
Net income/Equity	Less than 0	Greater than 0
Equity/Assets	Less than 0	Greater than 0

Table 2 Description of the actual financial status

Company	2015	2016	2017	2018	2019
Company A	N-D	N-D	N-D	N-D	N-D
Company B	N-D	N-D	N-D	N-D	N-D
Company C	N-D	N-D	N-D	N-D	N-D
Company D	N-D	N-D	N-D	N-D	N-D
Company E	N-D	N-D	N-D	N-D	N-D
Company F	N-D	N-D	N-D	N-D	N-D
Company G	D	D	D	D	D
Aurassi Company	N-D	N-D	N-D	D	D
Biopharm Company	N-D	N-D	N-D	N-D	N-D
Ruiba Company	N-D	N-D	D	D	D
Sycma Company	D	N-D	D	D	D
Saidal Company	N-D	N-D	N-D	N-D	N-D

Table 3 Finanacial variables

Financial variables	
Current Assets/Current Liabilities	Net income/Fixed assets
Profits before taxes/Current liabilities	Net income/Sales
Current Liabilities/Assets	Current Liabilities/Current Assets
Net income/Assets	Net working capital/Assets
Sales/Assets	Debts/Assets
Fixed assets/Assets	Sales/Net working capital
Current Assets/Assets	Net income/Net working capital
Equity/Assets	Sales/Invested Capital
Non-current liabilities/Current assets	Net working capital/Sales
Net income/Equity	Cash/Assets
Sales/Equity	/

These financial ratios were derived from several traditional statistical models, and we aimed to rely on the largest possible number of financial ratios, in order to raise the level of the independent variable's impact on the dependent variable. We did not test these financial ratios statistically, because they are basically derived from combinations of statistical models dedicated to predicting financial distress, and these ratios were previously chosen using a statistical method among many other ratios to be independent variables, because they have the highest impact on the dependent variable (prediction of financial failure).

Results

Training the logistic regression model

Using Matlab, the financial data of the period between (2015-2017) were included in order to train the LR model to understand the study's objective, which is to predict financial distress two years before it occurs. The data related to the period between (2018-2019) were allocated to test the ability of LR model in predicting financial distress. The following results were reached, as shown in Table 4.

Table 4 Training results of the logistic regression model

Accuracy	Prediction speed	Training time
77.8	230 obs/sec	29.76 sec

The table 4 shows the LR training process provided by Matlab. What really matters to us is the accuracy, as we note that the LR model achieved a classification accuracy rate in the training phase of 77.8%, which can be said to be fairly acceptable, and it is probably to be greatly improved in the testing phase. Because the test sample size (40%) is smaller than the training sample (60%).

Testing the logistic regression model

To test the model's ability to predict financial distress, we will rely on the test sample. It should be noted that these data were completely excluded from the training phase and the model had never been dealt with before, which ensures the testing results credibility. In the following, the results obtained in the testing phase can be clarified, as shown in Table 5.

Table 5 Testing results of the logistic regression model

	2018		2019	
	Y	Y Output	Y	Y Output
Company A	1	1	1	1
Company B	1	1	1	1
Company C	1	1	1	1
Company D	1	1	1	1
Company E	1	1	1	0
Company F	1	1	1	1
Company G	0	0	0	0
Aurassi Company	0	1	0	0
Biopharm Company	1	1	1	1
Ruiba Company	0	1	0	0
Sycma Company	0	0	0	1
Saidal Company	1	1	1	1

Table 5 compares the actual financial status with the expected values. We note that the LR model was able to predict correctly in most cases, but it failed to determine the actual value in some cases, and this will cause a decrease in its accuracy. In order to further clarify the results shown in the table 5, we will rely on Table 6 to evaluate the logistic regression model classification accuracy.

Through Table 6, the vision becomes clearer to us, where we can note that the LR model achieved a classification accuracy of 83.33%, which is an acceptable. We also note that the accuracy has improved compared to the training phase, and this is considered a suitable. However, the model showed great weakness in identifying financial distress cases, as it failed to identify (3) cases, and it was able to identify (5). On the other hand, it was able to identify cases of non-distress very well, by determining (15) cases of non-distress, and it failed to determine one case.

Table 6 Classification accuracy of the logistic regression model

Observed		Predicted		
		Y		Percentage Correct
		0	1	
Actual Y	0	5	3	62.5
	1	1	15	93.75
Accuracy				**83.33**

Designing the artificial neural networks models

Using Matlab, the neural network models were built by defining the basics of the models structure in accordance with the understanding of the study's objective and reaching accurate results compared to the actual financial status. Knowing that the process of determining the parameters of the training process was carried out according to scientific foundations, and on the basis of repeated attempts in order to reach the optimal results, as shown in Table 7.

Table 7 Variants of the design phase

Elements of the design stage	Description
Input neurons	21 neurons
Hidden layer neurons	12 neurons
Output neurons	1 neuron
Learning rate	0.01
Push rate	0.7
Lowest permissible value of error	0.00001
Training function	TrainGDX
Transfer function	LOGSIG
Adoption learning function	LearnGDM
Performance function	Mean square error

It is noteworthy that the design variables remain consistent across both networks, and despite extensive experimentation with numerous other variables, the optimal outcomes were achieved by employing the variables specified in Table 7. After completing the design phase of neural network models, the financial data of the period between (2015-2017) was included in order to train the models to understand the study objective, which is to predict financial distress two years before it occurs.

Training the artificial neural networks models

In the following, the results achieved after training the models can be clarified, as shown in Table 8. We note from the Table 8 that the FFDTDNN achieved moderate results in the training phase. The accuracy rate was 83.33%, which is the same percentage achieved by the LR model in the testing phase. We notice that the network continued to face serious problems in determining the state of non-distress despite the completion of the training process. Will this problem remain in the testing phase? This is what we look forward to knowing in the next stage. The training process of the Elman NN model was completed without an error in classification.

Therefore, it can be said that the ENN was able to build the suitable weights that correspond to the desired objective. Thus, the ability of this model to determine the financial distress cases can now be tested.

Table 8 Training results of the neural networks

	Feed-forward distributed time delay NN					
	2015		2016		2017	
	0	1	0	1	0	1
0	0	2	0	1	0	3
1	0	10	0	11	0	9
Accuracy	83.34		91.67		75	
Overall	83.33					
	Elman NN					
	2015		2016		2017	
	0	1	0	1	0	1
0	2	0	1	0	3	0
1	0	10	0	11	0	9
Accuracy	100		100		100	
Overall	100					

Testing the artificial neural networks models

As a final phase, the ANN models can be tested based on the financial data of the period between (2018-2019). In the following, the results obtained in the testing phase can be clarified, as shown in Table 9.

Table 9 Testing results of the neural networks

	Feed-forward distributed time delay NN				Elman NN			
	2018		2019		2018		2019	
	Y	Y Output	Y	Y Output	Y	Y Output	Y	Y Output
Company A	1	0.9997979	1	0.9997853	1	0.9998131	1	0.9998518
Company B	1	0.9999887	1	0.9999950	1	0.9999609	1	0.9999964
Company C	1	0.9999940	1	0.9999893	1	0.9999853	1	0.9999763
Company D	1	0.9998988	1	0.9999393	1	0.9998425	1	0.9999035
Company E	1	0.9999976	1	0.9999984	1	0.9999972	1	0.9999945
Company F	1	1.00E+00	1	0.9998962	1	0.9999432	1	0.9994441
Company G	0	0.500000	0	0.5000003	0	2.13E-07	0	3.576E-07
Aurassi	0	0.5000022	0	0.5001582	0	3.144E-07	0	0.0044820
Biopharm	1	0.9999956	1	0.9999946	1	0.999992	1	0.9999907
Ruiba	0	0.5011415	0	0.5000024	0	0.0413313	0	0.0376234
Sycma	0	0.5008013	0	0.5000007	0	0.0396334	0	6.69E-06
Saidal	1	0.9995117	1	0.9982154	1	0.9996925	1	0.9992919

It can be seen from Table 9 that the FFDTDNN model was able to achieve generally acceptable results, but some of these results do not significantly correspond to the actual financial status. The network also showed weakness towards identifying cases of financial distress, exactly as the training phase.

The results of the ENN in the testing phase confirm the results of its training phase, as it achieved a classification accuracy of 100%, and we conclude that the network was able to predict values that are almost identical to the actual financial status in all cases. Therefore, the network is suitable for predicting financial distress. In order to evaluate the ANN models classification accuracy more clearly, we will rely on Table 10.

Table 10 Classification accuracy of the neural networks

Feed-forward distributed time delay NN				
Observed		Predicted		
		Y		Percentage Correct
		0	1	
Actual Y	0	0	8	0
	1	0	16	100
Overall Percentage				**66.68**
Elman neural network				
Observed		Predicted		
		Y		Percentage Correct
		0	1	
Actual Y	0	8	0	100
	1	0	16	100
Overall Percentage				**100**

We conclude from Table 10 that the FFDTDNN model was unable to achieve a suitable overall classification accuracy rate compared to the ENN and logistic regression models, as its accuracy in predicting financial distress was only 66.68%, and it was able to correctly identify (0) cases of financial distress. On the other hand, it was able to accurately classify (16) cases of non-distress. We note that the ENN achieved very suitable results, and outperformed all other models in terms of classification accuracy, as its accuracy in predicting financial distress was 100%, and it was able to correctly identify (8) cases of financial distress, and to accurately classify (16) cases of non-distress.

Comparison

In order to identify the optimal model, we will rely mainly on the comparison element between the overall classification accuracy rate. In addition, we will also rely on the comparison between the prediction accuracy measures if the classification accuracy was equal, as shown in the Table 11.

It appears to us from Table 11, which aims to compare the classification accuracy of the three models, and their prediction accuracy measures (Mse, Rmse, Mae). We can easily conclude that the ENN model has a high ability to distinguish between distress and non-distress cases by noting its complete classification accuracy of 100%, compared to the other models. However, prediction accuracy measures point to the same and confirm the strength of the model in predicting financial distress. A classification accuracy of 100% is considered an excellent percentage, but it raises a kind of question: Could this network achieve the same percentage if the test sample was larger? I think it is impossible, because although the network achieved a classification accuracy of 100%, it also achieved error rates, and these rates are probably to be higher if we rely on a larger test sample, and this certainly means that the classification accuracy of the ENN will also decrease.

The coefficient of determination R^2 is employed in assessing the level of correlation between the independent variables and the dependent variable, as well as the correlation between the observed values and the predicted values. We note that R^2 is high for the ENN, and this is expected. We also note that the quality of

FFDTD NN is statistically better than the quality of LR, and as it seems, the value of R^2 is associated with error rates, and is not associated with classification accuracy. The lower the error rates, the higher the statistical quality of the model. Although the classification accuracy of LR is better than FFDTD NN, but its R^2 was lower.

Table 11 Classification accuracy comparison

Measure	LR	FFDTD NN	E NN
Accuracy	83.33	66.68	100
Missclassification	16.67	33.32	0
R^2	0.166	0.391	0.393
MSE	0.166666	0.083421	0.000196
RMSE	0.408248	0.288827	0.014017
MAE	0.166666	0.166879	0.005225
Precision	83.34	66.67	100
Recall	93.75	100	100
Specificity	62.5	0	100
F1-Score	0.8824	0.8	1
ROC-AUC	0.688	0.918	0.930

We note that both measures Precision and Recall confirm the same results, which are the superiority of ENN, the weakness of FFDTDNN, and average results for the LR model. For Specificity, the same again, and we note that both LR and FFDTDNN, show extreme difficulty in identifying non-distress statuses. As for F1-Score value, the higher the F1-Score, the better the predictive strength of the model. We note that its value for ENN is perfect, and a lower level for the other two models, knowing that its value for the LR was better than FFDTDNN. With regard to the value of ROC-AUC, its value is 0.930 m for ENN, which is an excellent value, because the higher the ROC-AUC, the better the predictive accuracy of the model. We also notice that the value of ROC-AUC for FFDTDNN is much better than its value for the LR. Thus, the adoption of this criterion leads to the validity of the classification, but with a high error rate and a decrease in the value of ROC-AUC, and this is what happened with the LR. The reason for this will be explained in the next paragraph.

Although the LR model could not achieve better classification accuracy than ENN, it was able to achieve better results than FFDTDNN model errors, even though its error measures are higher than the FFDTDNN. This is because of the criterion adopted in calculating the validity of the classification, as achieving a value greater than (0.5) for non-distress cases, it is considered a correct classification, and achieving a value less than (0.5) is considered an incorrect classification. The opposit for the distress cases, as achieving a value greater than (0.5), it is considered a incorrect classification, and achieving a value less than (0.5) is considered an correct classification. Thus, the adoption of this criterion leads to the validity of the classification, but with a high error rate and a decrease in the value of ROC-AUC, and this is what happened with the LR. For example, if LR model classified the no-distress status correctly, but the value was 0.51, the error rate would be 0.49, which is high. To be more clear, we present the ROC Curve of the three models.

In order to make the comparison process between the three models more accurate, although the final result has been decided in the results of Table 11, we decided to compare the accuracy of the three models in predicting financial distress a year before the distress, then two years before its occourence, which can be clarified in Table 12.

Figure 1 ROC Curve of the models

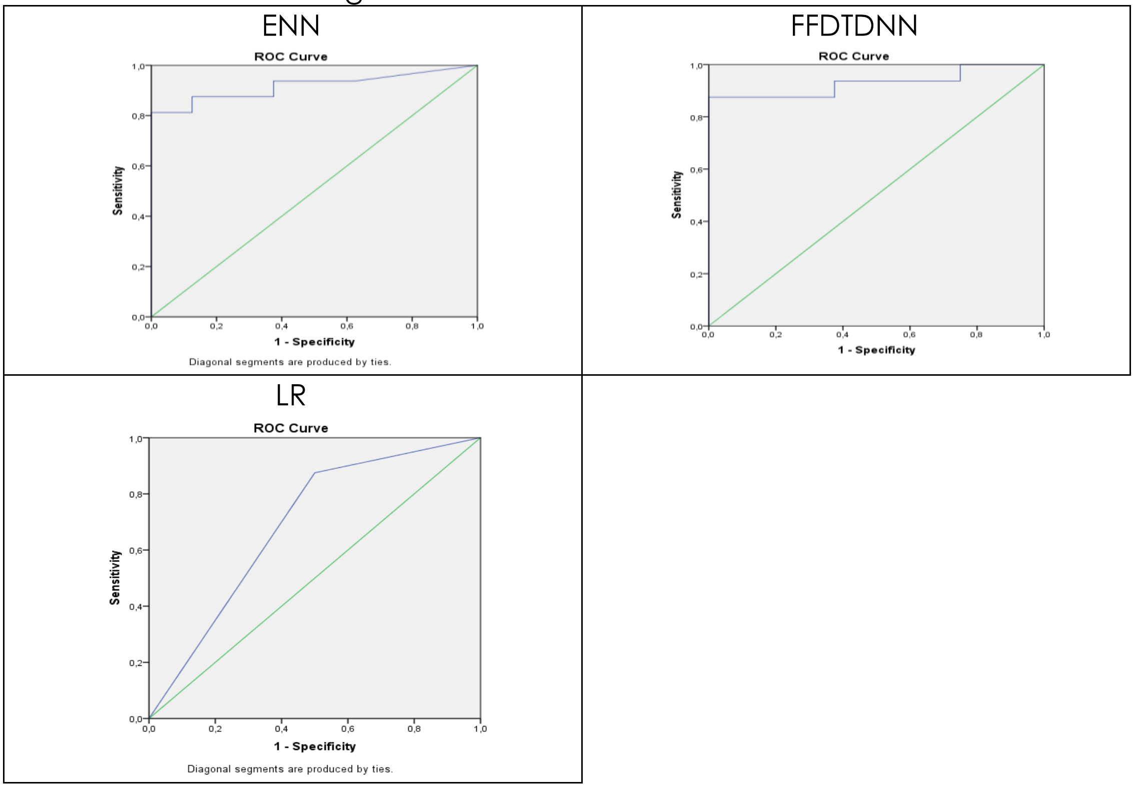

Table 12 Classification accuracy comparison of the N-1 & N-2

	N-1			
	Accuracy	MSE	RMSE	MAE
LR	83.34	0.167	0.408	0.167
FFDTD NN	66.67	0.0835	0.289	0.167
ENN	100	0.000273	0.0165	0.00681
	N-2			
	Accuracy	MSE	RMSE	MAE
LR	83.34	0.167	0.408	0.167
FFDTD NN	66.67	0.0835	0.289	0.167
ENN	100	0.00012	0.0109	0.00364

We note from Table 12 that the ENN model was able to outperform the FFDTDNN model and the LR in predicting financial distress a year before its occurrence, and the same for two years. We also note that the ENN model was able to achieve a complete classification accuracy rate of 100% for predicting distress a year, two years before its occurrence with a very low error values, but it is completely different for the other two models, and the results were moderate regarding the LR, and low regarding the FFDTDNN. Although the LR model was able to achieve better classification accuracy than FFDTDNN, its error measures are higher than FFDTDNN, whether it is for N-1 or N-2.

Conclusion

Finally, the theoretical aspects surrounding the prediction of financial distress were addressed. Subsequently, the LR model's ability to accurately predict financial distress in Algerian companies was compared to artificial neural intelligence models.

The results obtained from this comparison yielded valuable insights. By evaluating and comparing the models using important statistical and mathematical measures presented in Table 11, the superiority of the ENN model was evident. Hence, the ENN model is considered the optimal choice for predicting financial distress in Algerian companies. Notably, it achieved favorable results one and two years prior to the occurrence of distress. It is important to note that although the ENN model achieved a classification accuracy of 100% and minimal error values, it is unlikely to achieve the same percentage with a larger sample. This is due to the increased probability of errors as the number of financial cases increases. Therefore, in the case of a larger sample, lower classification accuracy and higher error values are expected for all models, not just a specific one.

LR model yielded less accurate results. However, it was able to outperform the FFDTDNN model, even though it is a statistical model, and its error values were higher than the FFDTDNN errors values. This is because of the criterion adopted in calculating the validity of the classification. As achieving a value greater than (0.5) for non-distress cases, it is considered a correct classification, and achieving a value less than (0.5) is considered an incorrect classification. The opposite for the distress cases.

Previous studies have proven many times that artificial intelligence models, especially artificial neural networks, are superior to statistical models. This is correct, but not in all cases, and we concluded that LR can achieve better results than some types of networks. These findings are relevant to individuals interested in forecasting, and predicting financial distress in particular. The key takeaway is that selecting the suitable technique is crucial in the forecasting process, regardless of whether it is a statistical or intelligent approach. The chosen technique should align with the objective of the prediction process.

Besides these valuable results, the study presents few limitations. For example. The analysis could be extended to a larger sample of companies from more countries. This is as a result of the difficulty of obtaining the financial statements of Algerian companies. Therefore, future studies in this field could consider firms from more countries and alternative methods such as the use of a hybrid intelligent multi-models in predicting financial distress.

References

1. Ağdeniz, Ş., Yıldız, B. (2016). A Comparative Study of Machine Learning Algorithms As an Audit Tool in Financial Failure Prediction. *TIDE AcademIA Research*, Vol. 1, No. 1, pp. 11-32.
2. Barboza, F., Kimura, H., Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, Vol. 83, pp. 405-417.
3. Bayraci, S., Susuz, O. (2019). A Deep Neural Network (DNN) based classification model in application to loan default prediction. *Theoretical and Applied Economics*, pp. 75-84.
4. Belas, J., Misankova, M., Schonfeld, J., Gavurova, B. (2017). Credit risk management: financial safety and sustainability aspects. *Journal of Security and Sustainability Issues*, Vol. 7, pp. 79-93.
5. Bonello, J., Bredart, X., Vella, V. (2018). Machine Learning Models For Predicting Financial Distress. *Journal of Research in Economics*, Vol. 2, pp. 174-185.
6. Callejón, A., Casado, A., Fernández, M., Peláez, J. (2013). A System of Insolvency Prediction for industrial companies using a financial alternative model with neural networks. *Int. J. Comput. Intell. Syst*, Vol. 6, No. 1, pp. 29-37.
7. Cancelliere, R., Gemello, R. (1996). Efficient training of Time Delay Neural Networks for sequential patterns. *Neurocomputing*, Vol. 10.

8. Chen, W., Du, Y. (2009). Using neural networks and data mining techniques for the financial distress prediction model. *Expert Systems with Applications*, Vol. 36, No. 2, pp. 4075-4086.
9. Cheng, C., Chan, C., Yang, J. (2018). A Seasonal Time-Series Model Based on Gene Expression Programming for Predicting Financial Distress. *Computational Intelligence and Neuroscience*, Vol. 2018, pp. 1-14.
10. El-Bannany, M., Sreedharan, M., Ahmed, M. (2020). A Robust Deep Learning Model for Financial Distress Prediction. *International Journal of Advanced Computer Science and Applications*, Vol. 11, No. 2, pp. 170-175.
11. Gregova, K., Valaskova, K., Adamko, P., Tumpach, M., Jaros, J. (2020). Predicting Financial Distress of Slovak Enterprises:Comparison of Selected Traditional and Learning Algorithms Methods. *Sustainability*, Vol. 12, No. 10, pp. 1-17.
12. Hardinata, L., Warsito, B. S. (2018). Bankruptcy prediction based on financial ratios using Jordan Recurrent Neural Networks: a case study in Polish companies. *Journal of Physics*, Vol. 1025, No. 1, pp. 1-6.
13. Ibiwoye, A., Ajibola, O., Sogunro, A. (2012). Artificial Neural Network Model for Predicting Insurance Insolvency. *International Journal of Management and Business Research*, Vol. 2, No. 1, pp. 59-68.
14. International Monetary Fund (2014). Algeria: Financial system stability assessment. Washington, D.C: IMF Country Report No. 14/161.
15. Jia, W., Zhao, D., Shen, T., Tang, Y., Zhao, Y. (2014). Study on Optimized Elman Neural Network Classification Algorithm Based on PLS and CA. *Computational Intelligence and Neuroscience*, Vol. 2014, pp. 1-13.
16. Jo, H., Han, I., Lee, H. (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems with Applications*, Vol. 13, No. 2, pp. 97-108.
17. Kapil, S., Agarwal, S. (2019). Assessing Bankruptcy of Indian Listed Firms Using Bankruptcy Models. Decision Tree and Neural Network. *International journal of business and economics*, Vol. 4, No. 1, pp. 112-136.
18. Karas, M., Reznakova, M. (2017). Predicting the bankruptcy of construction companies: A CART-based model. *Engineering Economics*, Vol. 28, pp. 145-154.
19. Koh, H. C., Tan, S. S. (1999). A neural network approach to the prediction of going concern status. *Accounting and Business Research*, Vol. 29, No. 3, pp. 211-216.
20. Lahmiri, S., Bekiros, S. (2019). Can machine learning approaches predict corporate bankruptcy? evidence from a qualitative experimental design. *Quantitative Finance*, Vol. 19, No. 9, pp. 1569-1577.
21. Mayliza, C., Manurung, A., Hutahayan, B. (2020). Analysis of The Effect of Financial Ratios to Probability Default of Indonesia's Coal Mining Company. *Journal of Applied Finance & Banking*, Vol. 10, No. 5, pp. 167-179.
22. Min, J., Lee, Y. (2005). Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters. *Expert Systems with Applications*, Vol. 28, pp. 603-614.
23. Mousavi, S., Amini, M., Raftar, M. (2012). Data mining techniques and predicting corporate financial distress. *Interdisciplinary Journal of Contemporary Research in Business*, Vol. 3, No. 12, pp. 61-68.
24. Mukeri, A., Shaikh, H., Gaikwad, D. (2020). Financial Data Analysis Using Expert Bayesian Framework For Bankruptcy Prediction. *Arxiv*, Vol. 3, pp. 1-24.
25. Obare, D., Muraya, M. (2018). Comparison of Accuracy of Support Vector Machine Model and Logistic Regression Model in Predicting Individual Loan Defaults. *American Journal of Applied Mathematics and Statistics*, Vol. 6, No. 6, pp. 266-271.
26. Odom, M. D., Sharda, R. (1990). A neural network model for bankruptcy prediction. IJCNN International Joint Conference on neural networks, pp. 163-168.
27. Ohlson, A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, Vol. 18, No. 1, pp. 109-131.

28. Osho, A., Idowu, A. (2018). Relevance of Accounting Theory in Forecasting Techniques and Default Prediction in an Organization in Nigeria. *European Journal of Business and Management*, Vol. 10, No. 29, pp. 116-129.
29. Paule-Vianez, J., Arias-Nicolas, J., Perez, J. (2019). Bayesian Networks To Predict Financial Distress In Spanish Banking. *Revista Electronica de Comunicaciones y Trabajos de ASEPUMA*, Vol. 20, pp. 131-154.
30. Puro, N. (2019). Financial Distress and Bankruptcy Prediction: A Comparison of Three Financial Distress Prediction Models in Acute Care Hospitals. *Journal of Health Care Finance*, Fall 2019, pp. 1-15.
31. Ribeiro, B., Silva, C., Vieira, A., Gaspar-Cunha, A. (2010). Financial distress model prediction using SVM+. The 2010 International Joint Conference on Neural Networks (IJCNN), pp. 1-7.
32. Salehi, M., Pour, M. D. (2016). Bankruptcy prediction of listed companies on the Tehran Stock exchange. *International Journal of Law and Management*, Vol. 58, pp. 545-561.
33. Shisia, A., Sang, W., Waitindi, S., Okibo, W. (2014). An In-Depth Analysis of the Altman's Failure Prediction Model on Corporate Financial Distress in Uchumi Supermarket in Kenya. *European Journal of Business and Management*, Vol. 6, No. 23, pp. 27-41.
34. Sudarsanam, S. (2016). A Fuzzy Neural Network Model for Bankruptcy Prediction. *Journal of Engineering Computers & Applied Sciences*, Vol. 5, No. 6, pp. 33-40.
35. Tang, Y., Ji, J., Zhu, Y., Gao, S., Tang, Z., Todo, Y. (2019). A Differential Evolution-Oriented Pruning Neural Network Model for Bankruptcy Prediction. *Complexity*, Vol. 2019, pp. 1-21.
36. Uzair, A., Ilyas, T., Asim, S., Kadhar, B. (2019). An Empirical Study on Loan Default Prediction Models. *Journal of Computational and Theoretical Nanoscience*, Vol. 16, No. 8, pp. 3483-3488.
37. Wilson, R., Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, Vol. 11, No, 5, pp. 545-557.
38. Xie, C., Luo, C., Yu, X. (2011). Financial distress prediction based on SVM and MDA methods: the case of Chinese listed companies. *Quality and Quantity: International Journal of Methodology*, Vol. 45, No. 3, pp. 671-686.
39. Zhang, G., Hu, M., Patuwo, B., Indro, D. (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research*, Vol. 116, No. 1, pp. 16-32.
40. Zhiyong, L. (2014). Predicting Financial Distress Using Corporate Efficiency and Corporate Governance Measures. Phd thesis.
41. Zhou, L., Lai, K., Yu, L. (2010). Least Squares Support Vector Machines Ensemble Models for Credit Scoring. *Expert Systems with Applications*, Vol. 37, pp. 127-133.
42. Zieba, M., Tomczak, S. K., Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, Vol. 58, pp. 93-101.

About the author

Amine Sabek is a young researcher, recently obtained a PhD in management sciences, majoring in financial management, University of Tamanghasset. The title of his doctoral thesis is related to the use of artificial intelligence techniques in predicting financial distress compared to statistical techniques. He was a member of the Economic Development Studies Laboratory, then became a member of the Investment Bets and Sustainable Development Stakes in Border Areas Laboratory. He obtained a master's degree in management sciences in 2018, majoring in financial management, University of Guelma. He has university experience in teaching mathematics, participated in many forums, and published many articles. He is interested in the topic of predicting financial distress using artificial intelligence and other topics. He has a spirit of teamwork and ambition to develop his CV and create new academic relationships with new scientific collaborations. Author can be contacted at *sabek.amine@univ-tam.dz*.