The Study of the Predictive Power of Metaheuristic Algorithms to Provide a Model for Bankruptcy prediction

Ghazaleh Alibabaee
PhD student, Department of Accounting, Islamic Azad University, Damavand Branch, Damavand, Iran
ghazal.alibabaee@hotmail.com

Mohammad Hamed KhanMohammadi
Assistant Professor, Department of Accounting, Islamic Azad University, Damavand Branch, Damavand, Iran
(Corresponding Author)

Submit: 18/01/2021 Accept: 03/02/2021

ABSTRACT

Progress in technology as well as extensive environmental changes have accelerated the economy. Intelligent models are examples of these developments predicting the corporate bankruptcy in the future. Investors and creditors are very interested in predicting the corporate bankruptcy, looking for reliable strategies to identify distressed and bankrupt companies. In the present research, companies active in Tehran Stock Exchange in a 10-years period were investigated in terms of bankruptcy based on localized Kordestani-Tatli model based on Altman model, healthy and bankrupt companies were identified as, and those added to or excluded from the stock market during 10 years were eliminated. The research data were collected and refined by means of secondary data extracted from financial statements and through the databases of the Exchange Organization as well as the Central Bank. The models employed to assess the data and predict bankruptcy included six metaheuristic algorithms including gravitational, gray wolf, genetic, imperialist competition, whale and differential evolution algorithm in combination with neural network; and the quality of the predictive models were compared in terms of accuracy and other assessment criteria. Financial ratios in three groups of profitability, liquidity, and capital structure were considered as the input variables of the models with good predictive power in determining bankruptcy. Hence, from the six algorithms investigated in terms of accuracy, the proposed model of the current research includes the three genetic, imperialist competition, and gray wolf algorithms in combination with artificial neural network method having 86% correct prediction accuracy and being able to provide reliable results.

Keywords:

Bankruptcy Prediction, Artificial Intelligence (AI), Metaheuristic Algorithms, Financial Ratios.

1. Introduction

Fast technological progress and extensive environmental changes have accelerated the economy. Bankruptcy prediction models are one of the tools estimating the future of companies (Saqafi, 2002). Investors and creditors are very interested in predicting the corporate bankruptcy, since they incur high costs in case of bankruptcy. The models employed in the predictions each have their own strengths and weaknesses (Adnan & Homayun, 2002).

Bankruptcy may have a negative effect on both the financial institution and the global economy. In long term, businesses, investors, states, and academics have identified strategies to recognize a corporate potential bankruptcy risk to decline the economic loss of bankruptcy (Zywicki, 2008). Predicting the continuity of corporate activities and economic units is one of the significant tasks of many financial institutions. Such institutions and banks due to the decision to grant facilities, and individuals and investors due to making the right decision about their investments in companies require effective bankruptcy prediction models. In most studies performed on bankruptcy prediction, financial ratios have been considered, since the financial condition of the company can be realized by analyzing financial ratios.

Improvement of the economic environment and business environment is the key factor in preventing bankruptcy. However, in the meantime, bankruptcy prediction models act as alarm bells informing different individuals and groups to make the right decision before the event. In fact, both using statistical methods and intelligent methods based on past trends, these models start predicting the corporate bankruptcy in the future. Investigation of the course of studies carried out in the field of corporate bankruptcy prediction can reveal that in the global economic crisis, particularly in recent years due to the bankruptcy of many companies at the international level, the model had to be revised and modified models should be presented. Hence, in recent years, many researchers have developed hybrid and improved models. Generally, it can be expressed that given the increasing development of intelligent methods and algorithms as well as the improvement of computer hardware and software platforms, more optimal and accurate models are constantly presented and it is expected that this process will continue in the coming years (Qazanfari et al., 2018).

Now the question raised is that which of statistical models or intelligent hybrid systems may better predict the corporate bankruptcy. Given the research performed in this regard, intelligent hybrid systems are greatly capable of predicting compared to the statistical methods. Since the effectiveness of optimization algorithms differs in terms of convergence, computational speed, as well as the initial sensitivity, the results of six algorithms were compared with each other in the present study, and the corporate bankruptcy prediction was compared through using six metaheuristic algorithms including gravitational, gray wolf, genetic, imperialist competition, whale and differential evolution algorithm using financial ratios. We did this in order to simultaneously compare the prediction accuracy of the algorithms of companies and finally present the algorithm with higher prediction accuracy in distinguishing bankrupt companies from healthy ones.

- **Theoretical principles and review of research background**
  - Definitions, principles, and theoretical principles related to bankruptcy
  - Definitions of bankruptcy

In the term of commercial law, bankruptcy stands for the state of a trader who has stopped paying his debts and cannot fulfill his commercial obligations (the Commercial Code, Article 412). Studies on crises and bankruptcies indicate that corporates could step forward better and have more favorable conditions in their crisis and be safe from the possible crisis or bankruptcy risks if they had observed appropriate control systems on risk management, using internal auditing system, as well as adopting to rules and regulations for transparency of their financial statements. Corporate governance system is a process largely meeting the needs of company managers besides their stakeholders in controlling or preventing the crisis as well as helping them in crises (Saadat & Koushki, 2014).

- **Theoretical framework of financial bankruptcy**

Over the period 1930-1965, few studies have been carries out on predicting bankruptcy (Bellovary et al., 2007). Initial studies on the use of financial ratio analysis in corporate bankruptcy prediction are of the type of univariate analysis and their most significant one is Beaver model (1966). In 1966, in response to

Beaver,W.H, Tamari showed that the financial health of a corporate does not depend on only a single variable. The risk index introduced by Tamari is indeed a simple scoring system in which several financial ratios are employed as financial health measuring criteria. In 1968, Altman first used multiple diagnostic analytical method in order to predict business disability. In 1980, Ohlson introduced a Logit analysis model for bankruptcy prediction. The Logit model combines a number of features or specifications of the company, through which calculating a probability for that company. This number represents the probability of bankruptcy or vulnerability to bankruptcy.

The trend of studies indicates that although the statistical models could provide good predictions for corporate bankruptcy, limiting assumptions of some of these models, like linearity, normality, and independence of predictor variables affected their effectiveness (Raei & Fallahpour, 2008). Thus, to overcome these limitations and improve the performance of predictions, artificial intelligence (AI) techniques were gradually developed. Bankruptcy AI models primarily concentrate on the signs of business disability; they are generally multivariate, and the variables employed in them are extracted from the data in the company's financial statements (Aziz & Dar, 2006).

The neural networks were first used in the design of bankruptcy prediction models by Odem and Sharda in 1990. Moreover, in 1993, Jang succeeded to use the linguistic power of fuzzy systems and neural network training, introducing a system called the adaptive neuro-fuzzy inference system. These systems are known as ANFIS system (Vakilifard & Pilevari, 2013).

Generally, the reasons for bankruptcy may be categorized into two categories of inter-organizational and external-organizational reasons. External reasons are the specifications of the economic system, competition, changes in trade as well as improvements and transfers in public demand, trade fluctuations, financing, and accidents. Inter-organizational factors of bankruptcy of business units include factors that can be prevented by the business unit through acting appropriately. Most of these factors are resulted from wrong decision-makings and their responsibility is directly on the business unit. These factors include over-crediting, inefficient management, insufficient capital, betrayal, and fraud.

- **Definitions, Principles, and Theoretical Foundations of Metaheuristic Algorithms of the Research**

  - Gravitational Algorithm

In 2009, gravitational search algorithm was introduced by Rashedi et al. (Ebrahimzadeh et al., 2016). Gravitational search algorithm (gravitational optimization algorithm) has been presented with the inspiration of the concepts of mass and gravitational as well as simulation of the nature laws. In this method, each answer is regarded as a particle with a mass proportional to its function. I.e. better answers have more masses, in turn having a greater ability and possibility to absorb other masses. Based on Newton, the law of motion states that the acceleration of each particle is directly proportional to the total force applied on it and inversely proportional to its mass. Hence, in a group of objects, the smaller objects move towards the heavier ones. In this optimization algorithm, each object shows one of the answers to the problem. I.e., all answers move towards better answers (Hosseinipour et al., 2017).

This algorithm has outstanding specifications distinguishing it from other algorithms like low memory requirement, rapid convergence, as well as the use of several parameters. Gravitational search algorithm is employed to solve optimization problems in which the desired answers are objects in the problem space and the amount of objects is specified based on the objective function. The system space is first specified, including a multidimensional coordinate system in the problem space definition. After formation of the system, the rules governing it are specified and it is assumed that only the gravitational law and the laws of motion rule this system.

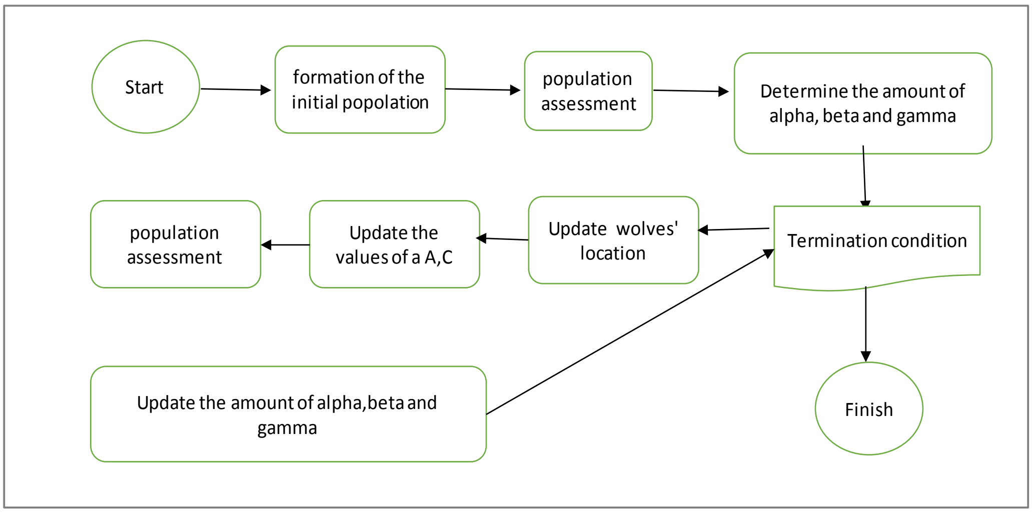
  - Gray Wolf Algorithm

In 2014, the Gray Wolf algorithm was introduced by Mirjalili and Lewis as a type of swarm intelligence optimization algorithm, simulating the social hierarchy and hunting behavior of the gray wolf in nature.

This algorithm has three main stages:

- Viewing the prey, tracking, and chasing it
- Approaching, surrounding the prey and misleading it until it stops moving
- Attack the prey

The type of wolves is determined at the onset of the algorithm. There are four types of wolves in the GWO algorithm: leader wolves or Alpha, Beta, Delta, and Gamma wolves. Alpha wolves are assumed as the algorithm's main driver. Beta and Delta wolves are often involved as Alpha wolves' assistants and the rest of the wolves are regarded as their followers. In the first stage of the algorithm, gray wolves surround the prey during the hunting. Over the hunting stage, the gray wolf attacks the surrounded prey. The hunting process is usually driven by Alpha but Beta and Delta wolves may sometimes participate in hunting (Misaqi et al., 2019).



The following Fig. shows the gray wolf algorithm:

- Genetic Algorithm

Inspired by the science of genetics and Darwin's theory of evolution, the Genetic algorithm (GA) was first presented by Professor Holland in 1975. This algorithm is a subset of evolutionary algorithms and a method of solving optimization problems (Mohammadi et al., 2019). GA is based on selection, survival, and evolution in natural environments. This method seeks to mathematically model the natural selection system, and the fundamental difference between it and other search methods is that it deals with a population of coded points instead of concentrating on finding the values of each point that may form the answer set. In this method, just the values of the objective function, which must first be clearly specified by a combination of all the targets, are used to direct the search, with no need to know how the variables alter. To optimize the parameters of artificial neural networks, the objective function is considered the minimization of MSE of all simulation and estimation data.

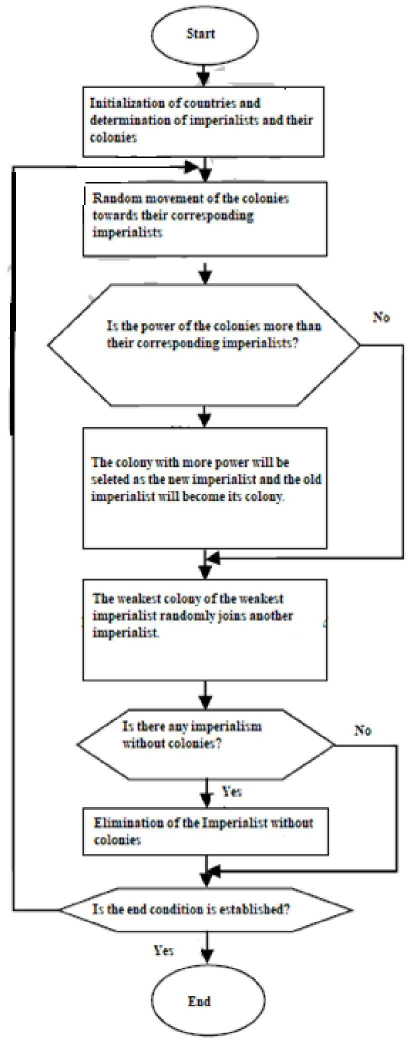
F=Minimize (MSE)

The key features of this algorithm are:

- Instead of a single point, it examines a large number of points in the space.
- Instead of directly using variables, they work with fields made up of variables and parameters.
- Instead of certain rules, in its search process, it uses possible rules.

- Imperialist Competition Algorithm

The Imperialist Competition Algorithm (ICA) refers to an emerging algorithm in the area of evolutionary computation derived from the human society's evolution trend. Like other evolutionary algorithms such as GA, this algorithm also starts with a number of random initial populations, each of which is called a country. Each country has its own specifications determining its location in the search space. Indeed, each country is defined as a point of this space. After scoring, the countries are sorted by their fitting value and a certain number of the best ones are published as the imperialist countries while other countries are considered as colonies. In this algorithm, the term power is used instead of the term fitting. Accordingly, the strongest countries are called imperialist and the weaker ones are introduced as colonies. Colonies are randomly divided between the colonizers. The more the power of the imperialist country is, the larger the number of colonized countries will be. Movement in the search space is in the way that each colony randomly moves under the effect of the power of its imperialist country and in parallel with it. Each iteration of the algorithm is called a decade over which there is an imperialism competition between the imperialists. In each decade, the weakest imperialist will lose its weakest colony. Other imperialists compete for ownership of this weak colony. The likelihood of assigning this colony to each of the imperialists will be proportional to the degree of their power. Finally, an imperialist become a colony with losing all its colonies and as a colony, it will be imperialized by another imperialist. The stages of the algorithm continue in the same way, until finally only one imperialist remains, taking the dominance of all countries. Moreover, different conditions, like the implementation of a predetermined number of iterations, can be considered as the ending conditions of the algorithm (Jamali et al., 2016).



- Whale Algorithm

In 2016, inspired by the social behavior of humpback whales, the Whale optimization algorithm was initially proposed by Mirjalili and Lewis. The Whale algorithm starts with a set of random solutions. In each iteration, the search factors update their position by means of three operators called prey surrounding, bubble-grid attack method (extraction phase), and prey search (exploration phase).

In prey surrounding, humpback whales identify prey and surround it. The Whale algorithm assumes that currently, prey is the best solution. After the best search factor is identified, other search actors update their location towards the best search factor. In the bubble-grid attack method (extraction phase), the humpback whale swims around the prey along a contractile circle and simultaneously in a spiral path. In order to model this simultaneous behavior, it has been assumed that the whale chooses between a contractile siege mechanism and a spiral model with a 50% probability to update the position of the whales over optimization. In the prey search (exploration phase), for updating the position of the search factor, random factor selection is used instead of employing the data of the best search factor (Mohammadi et al., 2019).

- Differential Evolution Algorithm

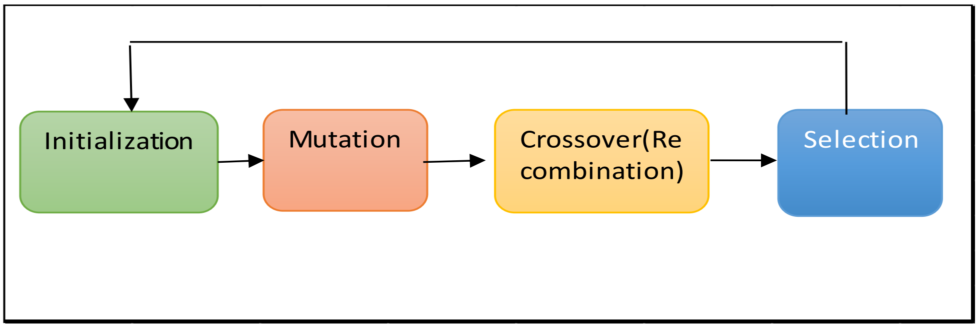
In recent years, population-based metaheuristic algorithms have been extensively employed to solve optimization problems. First introduced in 1995 by Storn and Price, the differential evolution algorithm can be referred to as one of these types of algorithms. The strategy of this algorithm in producing the initial population as well as continuing the evolution of the next generations, besides its approach to the competence of the objective function, is the same as the GA; nevertheless, there are some differences in how hybrids and mutation work. In this algorithm, unlike the GA, all members have the same opportunity to perform the combination operation; i.e., the chance of each chromosome being chosen for the combination operation does not depend on its merit value. Another specification of this algorithm is a memory storing information of appropriate answers in the current population.

Furthermore, the small number of control parameters responsible for directing the algorithm to obtain the optimal global response may be regarded as another advantage of this algorithm (Rashidi, 2017).

The following Fig. illustrates the operation process of differential evolution algorithm operators: (Mansouri et al., 2015).

This algorithm, like other population-based optimization algorithms, consists of two stages (initialization and evolution). In the first stage, the population is randomly generated if there is no data on the problem. In the evolutionary stage, individuals in the population repeatedly improve via mutations, hybridization, as well as the selection process, until the selection criterion for ending (termination) the algorithm is satisfied.

The differential evolution algorithm differs from other evolutionary algorithms in two aspects of the intersection order and mutation. In this way, a temporary response is first generated by means of the mutation operator and a new response is then generated by the intersection operator. Moreover, the step and length of the mutation are made by the distance and difference between the responses in the population (Farzan et al., 2018).



Research Background

To identify distressed and bankrupt financial companies in Tehran Stock Exchange **Kordestani and Tatli (2014)**, in a research under the title of Assessing the Bankruptcy Models Predictability (Comparing Primary and Modified Models) extracted a model from modifying coefficients of Altman, Springate, and Zmijewski models and their best explanatory variables y means of multivariate regression and logistic regression. Findings reveal that the initial models of Altman, Springate, and Zmijewski do not have the significant power to identify the distressed and bankrupt corporations in the Iranian environment, and the modified models are more powerful. Besides, the models adjusted based on information three years before the base year are more efficient compared to those adjusted based on the information five years before the base year. Moreover, the extracted models in accordance with the best explanatory variables in both audit analysis and logit analysis methods are capable of identifying the financially distressed and bankrupt companies at an overall accuracy level of 93%.

In the study by **Qazanfari, Rahimi Kia, and Asgari (2017)** under the title of corporate bankruptcy prediction based on hybrid intelligent systems with the implementation of a coherent and intelligent system based on neural network, learning support vector machines, and intensified learning, besides using imperialist competition optimization algorithm, cultural algorithms, and harmony search, it has been tried to eliminate the shortcomings of previous models as much as possible and at the international level. The findings indicate the superiority of the performance of hybridization of the backup vector machine with harmonic search optimization algorithms and imperialist competition in the absence of outliers.

Wang et al. (2014) in a study entitled, "An Improved Boosting Based on Feature Selection for Corporate Bankruptcy Prediction" have assessed the FS-Boosting method to choose the variables affecting the corporate bankruptcy prediction in two international databases. The results indicate that the accuracy of the model when using FS-Boosting compared to when not using it. The results of this study have been assessed for general classification accuracy, type I error, and type II error.

Jang et al. (2012) presented a new approach to properly adjust the factors influencing the performance of neural network models including input variables, number of hidden nodes, as well as the constant analysis value. Moreover, network search method and GA were continuously employed to properly adjust the collapse parameters' weight value and number of hidden nodes. This approach was applied to improve the performance of the neural network model for predicting the probability of corporate bankruptcy and its results were compared with the results of the existing bankruptcy prediction models like case-based reasoning model, decision tree, generalized linear model, multiple differentiation analysis model, and support vector machine and it was found that this model has significantly better performance compared to the abovementioned comparative models.

Liu (2008) examined three logistic regressions, decision tree, and neural networks models in terms of business failure prediction. The results revealed that in terms of wep1, the neural networks, as well as the logistic regression, all failed companies are correctly classified and the decision tree algorithm has the highest weep in classifying successful and unsuccessful corporates.

Kim and Kang (2012) in their study used the combination of GA with optimization models in corporate bankruptcy prediction. The variables used in this study included net profit to total assets ratio, financial costs to total assets ratio, net profit to sale ratio, current assets to total assets ratio, current assets to current debts ratio, and current debts to sales ratio. The results of their study revealed that combining GA with optimization models results in better bankruptcy prediction.

In their research, **Wu et al. (2007)** tried to develop a support vector machine (SVM) model using GA to optimize the factors of this method in the best way. They suggested that in terms of predictability and generalizability, the SVM model using their GA method is significantly superior to other methods.

Shin and Lee (2002) employed the genetic model to predict corporate financial distress, and their findings revealed that in addition to being suitable for financial distress prediction, the genetic model is easy for users to understand.

Research Hypotheses

The following hypothesis has been reviewed in this research:

Corporate bankruptcy predictability by means of a combination of neural network and GA, gravitational algorithm, gray wolf algorithm, imperialist competition algorithm, whale algorithm, and differential evolution algorithm are significantly different from each other.
The inputs of the abovementioned models have been the chosen financial ratios data set so that the accuracy of the models in corporate bankruptcy prediction can be tested.

- **Research Methodology**

This is a descriptive study in nature and content, analyzing the bankruptcy predictability models by means of the secondary data extracted from the financial statements of companies listed on the Tehran Stock Exchange. This research will be carried out in the framework of deductive-inductive reasoning. On the other hand, this study is of post-event (quasi-experimental) type, i.e. it is based on the analysis of past and historical data (financial statements of companies). Furthermore, this is a library and modeling study based on the analysis of panel data (data panel).

To choose the desired statistical population to test the data, the data of all companies listed on Tehran Stock Exchange from 2007 for 10 years were extracted in the first stage. In the second stage, banks, insurance, financial intermediation, besides companies added to or excluded from the Stock Exchange within 10 years were eliminated, leaving 408 companies in various industries that had remained stable for 10 years.
At the third stage, TKA-Z-Score was calculated for 408 companies (the index considered for bankruptcy is called the localized Kordestani-Tatli model based on the Altman model) and companies with TKA-Z-Score less than or equal to (-0.5) were presumed bankrupt and the companies that had TKA-Z-Score with these conditions for three consecutive years from the base year (2016) were chosen as bankrupt. Nevertheless, since there should be enough healthy and bankrupt companies in different industries to compare them, some industries were eliminated due to the following two reasons:

- Because of the absence of a bankrupt company in the industry, the industry was eliminated.
- Because of the lack of a healthy company in this industry, the industry was eliminated.

Out of the 34 existing industries, 18 were eliminated and 16 remained, leaving 168 companies out of the 408, 84 of which were healthy and 84 were bankrupt. Healthy companies were chosen as follows:

- The scope of information of this group was similar to that of the bankrupt companies.
- Their variables X1 to X4 were positive in the localized Kordestani-Tatli model based on Altman model.
- If a company was bankrupt in the base year but healthy in the previous year, it was not considered as a healthy company.

- **Research Variables and Model**
  - Hypothesis test method

$$\text{Bankruptcy}_{it} = \beta_{0} + \beta_{2}\, X3_{it} + \beta_{3}\, X4_{it} + \beta_{4}\, X5_{it} + \beta_{5}\, X7_{it} + \beta_{6}\, X13_{it} + \beta_{7} X1_{it} + \beta_{8}\, X6_{it} + \beta_{9}\, X8_{it} + \beta_{10}\, X16_{it} + \beta_{11}\, X\,17_{it} + \beta_{12}\, X20_{it} + \beta_{13}\, X\,19_{it} + \beta_{14}\, X30_{it} + \beta_{15}\, X31_{it} + \beta_{16}\, X41$$

In the above model, variables include:

Symbol	Type of variable	Variable
Bankruptcy	Dependent	Bankruptcy
X3	Independent	Current ratio
X4	Independent	Quick ratio
X5	Independent	Cost to interest ratio
X7	Independent	Debt ratio
X13	Independent	Cash ratio
X1	Independent	of long-term debt to equity ratio
X6	Independent	Debt to equity ratio
X8	Independent	Working capital to assets ratio
X16	Independent	Return to total assets ratio
X17	Independent	Gross profit margin
X20	Independent	Ratio of accumulated profit to total assets
X19	Independent	Profit margin before tax
X30	Independent	Financial Leverage
X31	Independent	Interest coverage ratio
X41	Independent	Equity to total assets ratio

- Research variables:
  - <u>Bankruptcy</u>

In the present research, there is dependent variable and the virtual variable is considered; if the company is bankrupt, the number one is considered, otherwise the number zero.
In this study, the financial bankruptcy is calculated in accordance with the localized Kordestani-Tatli model

based on the modified Altman models designed for the Iranian economic environment. The initial models of Altman, Springate, and Zmijewski are significant to identify distressed companies; they do not have financial bankruptcy in the Iranian environment and the modified models are more capable. Moreover, modified models based on information three years before the base year are more efficient compared to models modified based on the information five years before the base year. Like the modified models, the localized Kordestani-Tatli model is significantly able to identify healthy, bankrupt, and financially distressed companies in Iranian economic environment with the lowest error rate (assessing the bankruptcy model predictability with Kordestani, Tatli). The best variables explaining bankruptcy have been chosen based on the model (Auditing Knowledge Journal, 2014). The abovementioned model (audit analysis), like the modified Altman and Springate models up to three years before the base year, can significantly identify healthy, bankrupt, and financially distressed companies in the economic environment of Iran (Journal of Auditing Knowledge, 2014).

The localized Kordestani-Tatli model is as follows:

$$MDA=0.626 X_1+0.137 X_2+0.679 X_3+0.583 X_3$$

X_1=Profit (loss) accumulated to total assets ratio
X_2=Operating profit (loss) to total assets ratio
X_3=Net profit (loss) to total assets ratio
X_4=Total debt to total assets ratio

The methods of combining neural network and GA, gravitational algorithm, gray wolf algorithm, imperialist competition algorithm, Whale algorithm, and differential evolution algorithm have been used to predict bankruptcy, and the best method with higher accuracy has been announced as an appropriate method for prediction. Excel software and MATLAB software have been employed for the initial data processing and for simulation in predictive models, respectively.

- **Research Data Analysis**

The selected models were assessed and tested at this stage and their results were compared with each other. In the following, the criteria for evaluating the efficiency are first expressed and the used data are described. Subsequently, the models are tested; ultimately, their results are analyzed.

○ **Criteria for assessing prediction models**

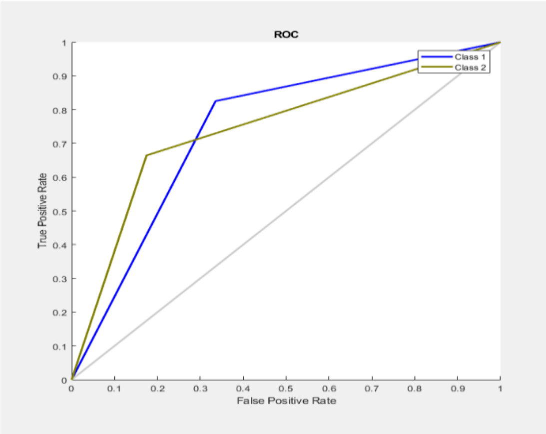
In the present study, 80% of the data were selected as training data and the rest as test data. The data selection method is random. The criteria for assessing the quality of error prediction models in the software are as follows, the answers obtained from these criteria are placed next to each other, and the assessment is performed for the models.

Description	Optimal answer
Misclassification Rate (MR)	The lowest value
Except cost of Misclassification (ECM)	The lowest value
Normalized Except cost of Misclassification	The lowest value
Sensitivity	The highest value
Specificity	The highest value
Accuracy	The highest value
Precision	The highest value
Recall	The highest value
F-measure	The highest value
Consistency	The highest value
Receiver operating characteristic	The highest value
Balance measure	The highest value

○ **Data set and research findings**

According to the prediction model in this research, the desired corporate's bankruptcy in the year has been defined as a dependent variable and 15 financial ratios have been used as independent variables for 168 selected companies as software input for prediction. At this stage, the output and data analysis are first presented in brief by method and data set, and subsequently, the simulation results and validation of the estimates have been discussed.

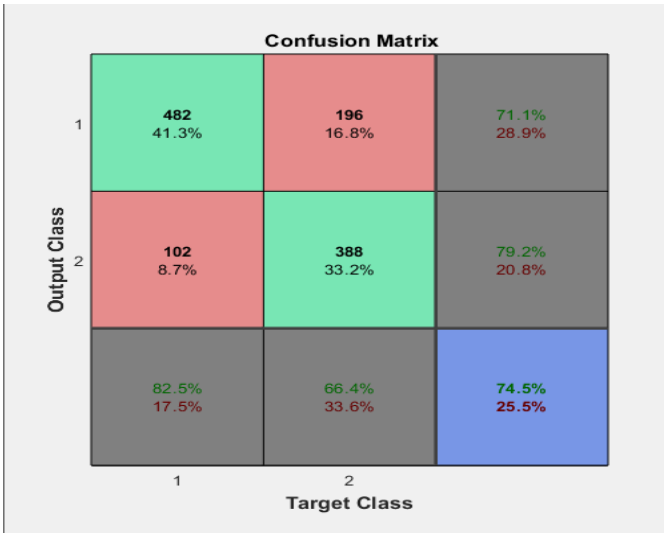
I. Gravitational Algorithm (GSA) Method Analysis



- ROC diagram obtained from the combination of neural network and gravitational algorithm

ROC Table					
Description		**TP-percentage**	**Description**	**FP-percentage**	**Description**
CLASS 1	Bankrupt	82.50	Correctly identified	33.60	Wrongly identified
CLASS 2	Healthy	66.40	Correctly identified	17.50	Wrongly identified

As it is obvious in the results, in combination of neural network and gravitational algorithm, it has succeeded to obtain the best result for the first class (bankrupt companies) with 82% accuracy, predicted the healthy companies with 66% accuracy, and from the total data, correctly identified 870 companies. Moreover, this method has generally succeeded to achieve an accuracy of 74.5%.

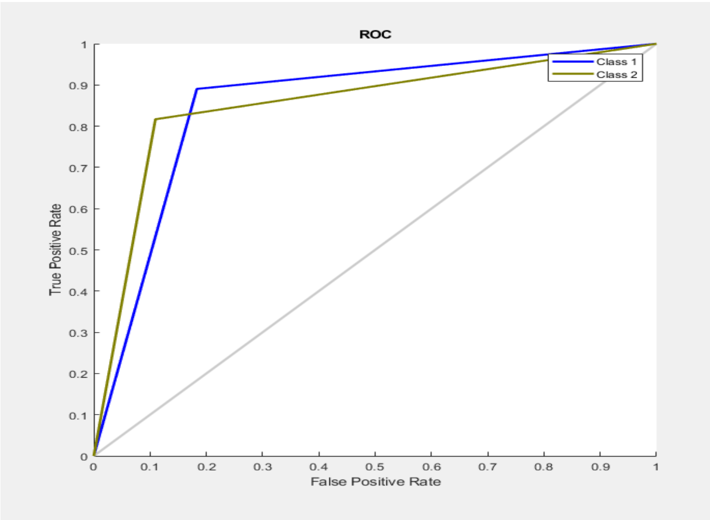


- Confusion diagram obtained from the combination of neural network and gravitational algorithm

Confusion Matrix			
N	Percentage	N	Percentage
٤٨٢	٤١,٣٠	١٩٦	١٦,٨٠
It is bankrupt and Correctly identified		It is healthy and wrongly identified	
١٠٢	٨,٧٠	٣٨٨	٣٣,٢٠
It is bankrupt and wrongly identified		It is healthy and Correctly identified	

74.5% of the total data was correctly identified.
25.5% of the total data was wrongly identified.

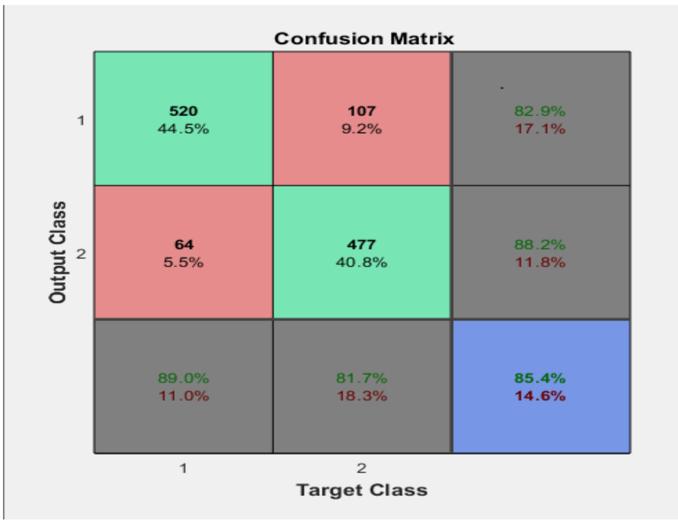
II. <u>Gray Wolf Algorithm Method Analysis</u>



- ROC diagram obtained from the combination of neural network and gray wolf algorithm

ROC Table					
Description		**TP-percentage**	**Description**	**FP-percentage**	**Description**
CLASS 1	Bankrupt	89.00	Correctly identified	18.30	Wrongly identified
CLASS 2	Healthy	81.70	Correctly identified	11.00	Wrongly identified

As it is obvious in the results, in combination of neural network and gray wolf algorithm, we have succeeded to obtain the best result for the first class (bankrupt companies) with 89% accuracy, predicted the healthy companies with 82% accuracy, and from the total data, correctly identified 997 companies. Moreover, this method has generally succeeded to achieve an accuracy of 85.4%.



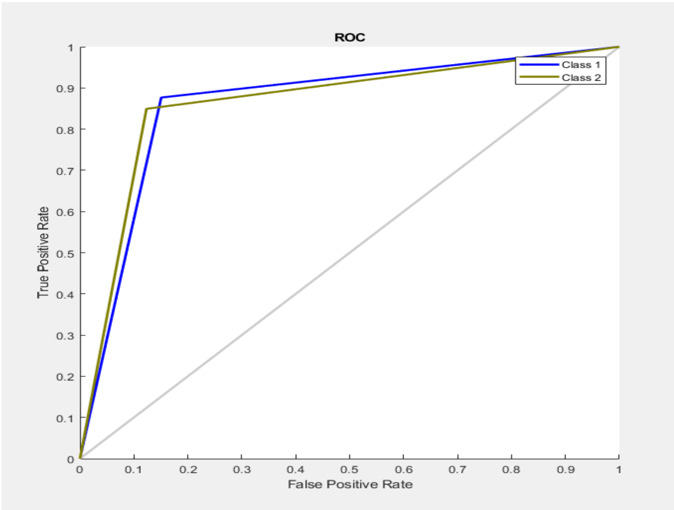
- Confusion diagram obtained from the combination of neural network and gray wolf algorithm

Confusion Matrix			
N	Percentage	N	Percentage
520	44.5	107	9.2
It is bankrupt and Correctly identified		It is healthy and wrongly identified	
64	5.5	477	40.8
It is bankrupt and wrongly identified		It is healthy and Correctly identified	

85.4% of the total data was correctly identified.

14.6% of the total data was wrongly identified

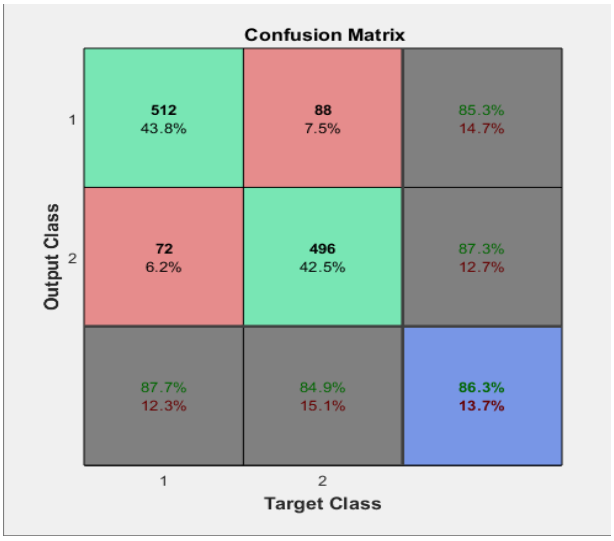
III. Genetic Algorithm (GA) Method Analysis



- ROC diagram obtained from the combination of neural network and genetic algorithm

ROC Table					
Description		TP-percentage	Description	FP-percentage	Description
CLASS 1	Bankrupt	87.70	Correctly identified	15.10	Wrongly identified
CLASS 2	Healthy	84.90	Correctly identified	12.30	Wrongly identified

As it is obvious in the results, in combination of neural network and genetic algorithm, we have succeeded to obtain the best result for the first class (bankrupt companies) with 87.70% accuracy, predicted the healthy companies with 84.90% accuracy, and from the total data, correctly identified 1008 companies, including 512 bankrupt companies and 496 healthy companies. Moreover, this method has generally succeeded to achieve an accuracy of 86.3%.

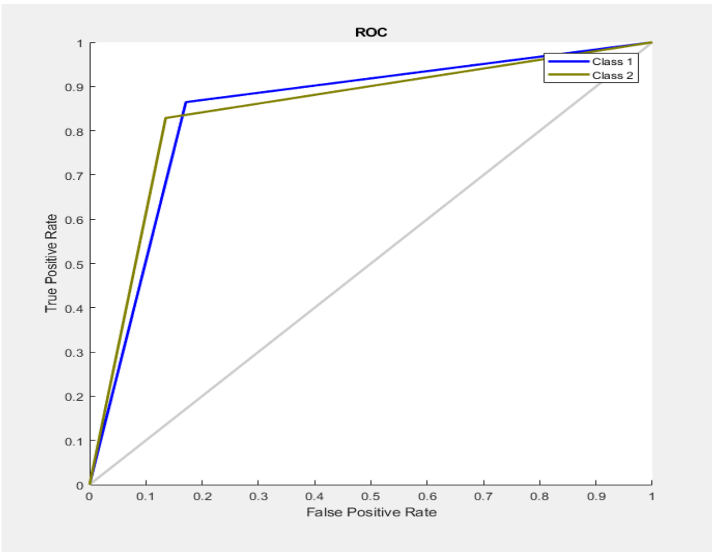


- Confusion diagram obtained from the combination of neural network and genetic algorithm

Confusion Matrix			
N	Percentage	N	Percentage
512	43.80	88	7.50
It is bankrupt and Correctly identified		It is healthy and wrongly identified	
72	6.20	496	42.50
It is bankrupt and wrongly identified		It is healthy and Correctly identified	

86.3% of the total data was correctly identified.
13.7% of the total data was wrongly identified

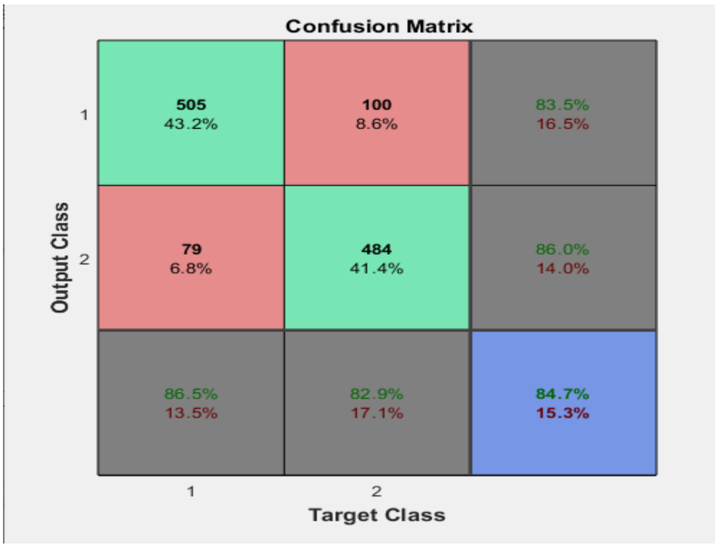
IV. <u>Imperialist Competition Algorithm Method Analysis</u>



- ROC diagram obtained from the combination of neural network and imperialist competition algorithm

ROC Table					
Description		TP-percentage	Description	FP-percentage	Description
CLASS 1	Bankrupt	86.50	Correctly identified	17.10	Wrongly identified
CLASS 2	Healthy	82.90	Correctly identified	13.50	Wrongly identified

As it is obvious in the results, in combination of neural network and imperialist completion algorithm, we have succeeded to obtain the best result for the first class (bankrupt companies) with 86.50% accuracy, predicted the healthy companies with 82.90% accuracy, and from the total data, correctly identified 989 companies. Moreover, this method has generally succeeded to achieve an accuracy of 84.7%.

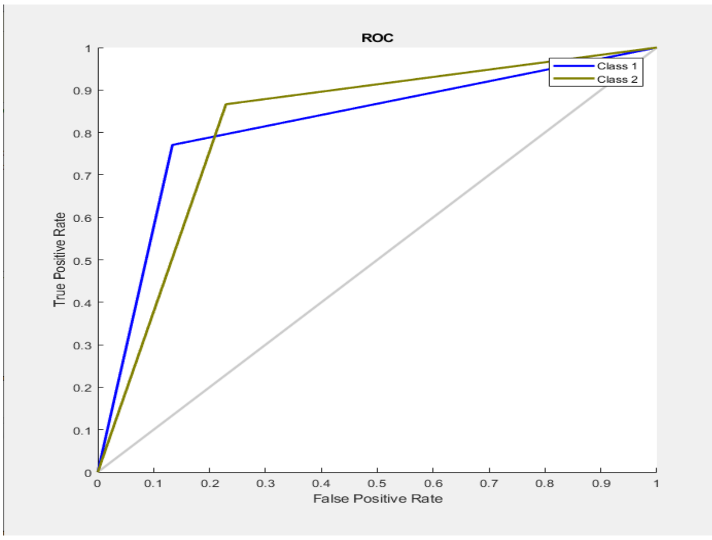


- Confusion diagram obtained from the combination of neural network and imperialist completion algorithm

Confusion Matrix			
N	Percentage	N	Percentage
505	43.2	100	8.6
It is bankrupt and Correctly identified		It is healthy and wrongly identified	
79	6.8	484	41.4
It is bankrupt and wrongly identified		It is healthy and Correctly identified	

84.7% of the total data was correctly identified.
15.3% of the total data was wrongly identified

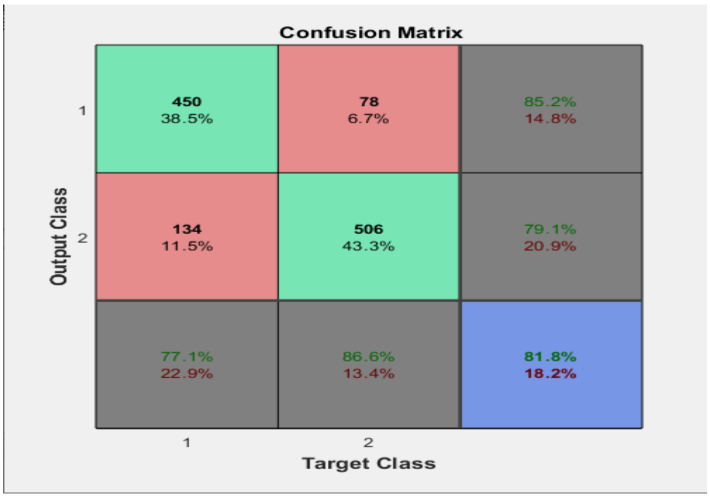
V. Whale Algorithm Method Analysis



- ROC diagram obtained from the combination of neural network and whale algorithm

ROC Table					
Description		TP-percentage	Description	FP-percentage	Description
CLASS 1	Bankrupt	86.50	Correctly identified	17.10	Wrongly identified
CLASS 2	Healthy	82.90	Correctly identified	13.50	Wrongly identified

As it is obvious in the results, in combination of neural network and whale algorithm, we have succeeded to obtain the best result for the first class (bankrupt companies) with 86.50% accuracy, predicted the healthy companies with 82.90% accuracy, and from the total data, correctly identified 956 companies. Moreover, this method has generally succeeded to achieve an accuracy of 81.8%.

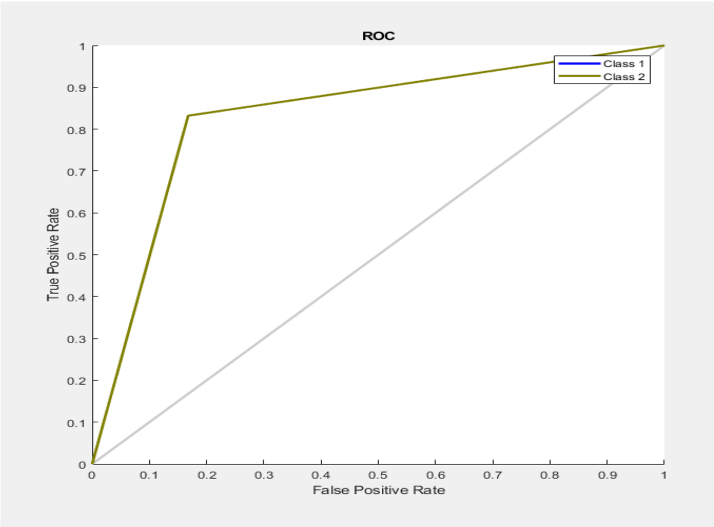


- Confusion diagram obtained from the combination of neural network and whale algorithm

Confusion Matrix			
N	Percentage	N	Percentage
450	38.50	78	6.70
It is bankrupt and Correctly identified		It is healthy and wrongly identified	
134	11.50	506	43.30
It is bankrupt and wrongly identified		It is healthy and Correctly identified	

81.8% of the total data was correctly identified.
18.2% of the total data was wrongly identified.

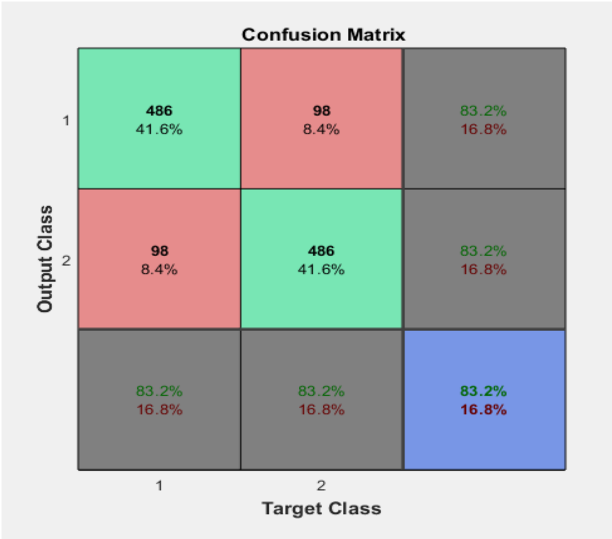
VI. Differential Evolution Algorithm Method Analysis



- ROC diagram obtained from the combination of neural network and differential evolution algorithm

ROC Table					
Description		TP-percentage	Description	FP-percentage	Description
CLASS 1	Bankrupt	83.20	Correctly identified	16.80	Wrongly identified
CLASS 2	Healthy	83.20	Correctly identified	16.80	Wrongly identified

As it is obvious in the results, in combination of neural network and differential evolution algorithm, we have succeeded to obtain equal result for both classes and predicted the results for both of the first class (bankrupt companies) and the second class (healthy companies) with 83.20% accuracy and from the total data, correctly identified 972 companies. Moreover, this method has generally succeeded to achieve an accuracy of 83.20%.



- Confusion diagram obtained from the combination of neural network and differential evolution algorithm

Confusion Matrix			
N	Percentage	N	Percentage
486	41.60	98	8.40
It is bankrupt and Correctly identified		It is healthy and wrongly identified	
98	8.40	486	41.60
It is bankrupt and wrongly identified		It is healthy and Correctly identified	

83.2% of the total data was correctly identified.
16.8% of the total data was wrongly identified

• Conclusion and Discussion

In the current research, the accuracy of corporate bankruptcy prediction by metaheuristic algorithms has been discussed and investigated. The results of the six algorithms were compared with each other by means of the quality assessment criteria of the error prediction models. Nevertheless, it is worth noting that the result of the neural network has been improved through the mentioned algorithms and the algorithm have been combined with the neural network in all methods; in fact, the results of this combination has been compared and the most accurate algorithm has been introduced.

The data set was 1168, in 168 companies of which the selected financial ratio has been continuously reviewed for 10 years to identify the bankrupt and healthy companies. Thus, in each of these algorithms, prediction has occurred with the percentage of accuracy stated in the table below to determine the accuracy of the model. Furthermore, other quality assessment criteria of the models are reviewed, summarized in the following table:

The six algorithms chosen in the present research were gravitational, gray wolf, genetic, imperialist competition and whale algorithm, as well as differential evolution algorithm. In order to test the desired hypothesis, the data have been assumed the same in all methods. The research hypothesis is "Corporate bankruptcy predictability by means of a combination of neural network and GA, gravitational algorithm, gray wolf algorithm, imperialist competition algorithm, whale algorithm, and differential evolution algorithm are significantly different from each other."

Description	Optimal answer	Gravitational	Gray wolf	Genetic	Imperialist competition	Whale	Differential Evolution
		GSA	GWO	GA	ICA	WOA	DE
Misclassification Rate (MR)	The lowest value	0.26	0.15	0.14	0.15	0.18	0.17
Except cost of Misclassification (ECM)	The lowest value	0.35	0.20	0.20	0.22	0.28	0.25
Normalized Except cost of Misclassification	The lowest value	0.18	0.10	0.10	0.11	0.14	0.13
Sensitivity	The highest value	0.79	0.87	0.88	0.86	0.79	0.83
Specificity	The highest value	0.71	0.83	0.85	0.83	0.85	0.83
Accuracy	The highest value	0.74	0.85	0.86	0.85	0.82	0.83
Precision	The highest value	0.83	0.88	0.89	0.86	0.77	0.83
Recall	The highest value	0.71	0.83	0.85	0.83	0.85	0.83
F-measure	The highest value	0.76	0.86	0.86	0.85	0.81	0.83
Consistency	The highest value	0.31	0.63	0.73	0.66	0.70	0.66
Receiver operating characteristic	The highest value	0.75	0.86	0.86	0.85	0.82	0.83
Balance measure	The highest value	0.75	0.85	0.86	0.85	0.82	0.83

The inputs of these models are a set of selected financial ratios in order to be able to test the accuracy of the models in corporate bankruptcy prediction.

Concerning the above hypothesis about the predictability of each method, since all six methods are AI and of metaheuristic algorithms type, they were successful in bankruptcy prediction. The highest accuracy was obtained by the combination of neural method and GA, followed by the combination of neural network and gray wolf algorithm and imperialist competition. A significant difference is observed between the results of AI methods and other statistical methods and AI methods have a higher predictability, they are useful approaches for optimization problems.

As illustrated in the table above, except for the gravitational algorithm, other methods were closer to each other in terms of accuracy. Respectively, GA had 86% accuracy, gray wolf algorithm and imperialist competition algorithm had 85%, differential evolution algorithm had 83%, whale algorithm had 82%, and gravitational algorithm had 74% accuracy in correctly identification of the total data. Furthermore, other quality assessment criteria of error prediction models were investigated in these methods. In the Misclassification Rate (MR), Except cost of Misclassification (ECM) and the Normalized Except cost of Misclassification, the lowest value will be better. This value has been 0.18 in the gravitational algorithm; but in other methods, they were closer to each other with lower values. The GA with the values 0.14, 0.20, and 0.10 has had the lowest value and been the best method, and then the imperialist competition algorithm and the gray wolf algorithm have obtained the best result.

Also in the F-measure, the study results are 0.86, 0.85, and 0.86, respectively, and the Receiver operating characteristic (ROC) shows 0.86, 0.85, and 0.86, respectively, confirming the three genetic, imperialist competition, and gray wolf algorithms. In the balance measure, the results of the study are 0.86, 0.85 and 0.85, also the Precision is 0.88, 0.89 and 0.86, respectively, which having the highest value among the six algorithms, the results show the success of three mentioned algorithms among other algorithms.

Concerning the input variable determining bankruptcy in these models, it is recommended to use the financial ratios as a fixed index in bankruptcy prediction models, since they have high bankruptcy predictability. In all areas of profitability, liquidity, and capital structure, financial ratios may provide useful information concerning the financial health of corporations and predict the future of companies for up to three years using the AI tools.

Approximation algorithms are able to find good (near-optimal) answers in a short time to solve difficult optimization problems. Meta-heuristic algorithms are one of the types of approximation optimization algorithms that have solutions for exit from local optimal points and may be usable for a variety of problems.

In AI iterative methods, it is tried to improve the results until obtaining the best results. A reason for improving the combination of neural methods and algorithms is that the neural network method’s optimal result is re-investigated and improved by the algorithms. The conducted studies have revealed that metaheuristic algorithms alone are not suitable methods for investigating bankruptcy prediction.

Hence, among the six methods studied, the current paper recommends three genetic, imperialist competitions, and gray wolf algorithms in combination with artificial neural network method to predict bankruptcy since they can provide reliable results.

References

* Jeong, C., Min, J. H., & Kim. M. S, 2012. A tuning method for the architecture of neural network models incorporating GAM and GA as applied to bankruptcy prediction. Expert Systems with Applications. 39(3): 3650-3658.
* Kim, Myoung-Jong & Dae-Ki Kang, 2012, Classifiers selection in ensembles using GAs for bankruptcy prediction. Expert Systems with Applications, pp. 1-7
* Kordestani & Tatli, Spring 2014, he Evaluate ability of Altman adjusted Model to Prediction Stages of Financial Distress Newton and Bankruptcy, Quarterly of Investment Knowledge, Issue 9, pp. 83-۱۰۰.
* Liou, F. M, 2008, Fraudulent financial reporting detection and business failure prediction models: A comparison. Managerial Auditing Journal. 23(7): 650- 662.
* Qazanfari, Rahimi Kia, and Asgari, 2017, Corporate bankruptcy prediction based on hybrid intelligent systems, Quarterly Journal of Financial Accounting and Auditing Research, Issue 37, pp. 159-194.
* Shin K, Lee Y. A GA Application in Bankruptcy prediction modeling, Expert systems with application 2002; 1-8.
* Wang,G., Ma, J., & Yang, S, 2014,“An Improved Boosting Based on Feature Selection for Corporate Bankruptcy Prediction”,Expert Systems with Applications, 41(5), PP.2353-2361
* Wu, C.H. Tzeng. G.H., Yeong-Jia G. and Fang W.C., 2007, A Real-Valued GA, To Optimize the Parameters of Support Vector Machine for Prediction Bankruptcy, Journal of Expert Systems with Application, Vol. 32, pp. 397-408