Available online at http://ijdea.srbiau.ac.ir
Int. J. Data Envelopment Analysis (ISSN 2345-458X)
Vol. 11, No. 1, Year 2023 Article ID IJDEA-00422, Pages 23-38
Research Article

International Journal of Data Envelopment Analysis Science and Research Branch (IAU)

Predicting the possibility of bankruptcy for the Companies admitted to the Tehran stock exchange by hybrid approach based on Data Envelopment Analysis and essential components analysis

M. Shahriari [*] [1], P. Hajizadeh[1], M. Hesam Asoodeh [2]

[1] Department of Industrial management, South Tehran Branch, Islamic Azad University, Tehran, Iran,

[2] Department of Technology Management, Islamic Azad University, Dubai, UAE.

Received 13 October 2022, Accepted 22 January 2023

Abstract

The task of forecasting corporate bankruptcy has garnered considerable attention from financial researchers. Given the consequences of a company's financial issues on its stakeholders, devising techniques and models to anticipate bankruptcy and financial instability has always been an appealing field of financial research. Company investors desire the security of their capital and expect a reasonable return commensurate with the risk they undertake. They anticipate being able to predict a company's financial crises ahead of time and safeguard their investment. To meet this need, researchers have conducted extensive research and used various methods and models to evaluate the financial performance of companies and of course predict their bankruptcy, but the point that should be noted is that none of these methods are not enough on their own, and they should be used together and of course using the professional judgments of elite experts. Among the methods that have attracted the attention of researchers in the last few years to facilitate the process of financial decisions, is the method of data envelopment analysis, which of course has had acceptable results. In this research, 52 listed manufacturing companies in the Tehran Stock Exchange were selected considering to available information into three groups: food and pharmaceutical companies, metal, automobile and machinery companies, and chemical and petrochemical companies, out of which, 21 companies (10 bankrupt and 11 healthy) companies were classified in the first group, 18 companies (10 bankrupt and 8 healthy) in the second group, and 13 companies (7 bankrupt and 6 healthy) in the third group.

This study aims to evaluate the efficacy of the data envelopment analysis model in predicting the financial insolvency of companies across different categories. To achieve this, we have employed the principal component technique, a statistical approach, to reduce the input dimension of the data envelopment analysis model. Our focus is on identifying the model's potential in anticipating financial distress for companies hovering near the brink of bankruptcy.

Keywords: Data envelopment analysis, Principal component analysis, Stock market, Bankruptcy

---

[*] Corresponding author: Email: shahriari.mr@gmail.com

1- Introduction

Altman, et al studied the problem of predicting bankruptcy of companies has significantly gained the interests of experts in the field of financial research. In the past, financial ratios and multiple diagnostic models have been used to predict bankruptcy, the most representative of which is the Altman model (Z-Score). In this research, Altman model could predict bankruptcy at one and two years before it is happening with the accuracies of %95 and %83, respectively [1].

Artificial neural networks can be mentioned among other methods used to predict bankruptcy in past researches. These models were more accurate compared to other prediction methods by Mehrani et al. [2].

Table 1 briefly describes the generalities of the most important past researches in the field of predicting corporate bankruptcy. Among the researches that has been done in recent years in the field of data envelopment analysis, the followings can be mentioned.

M. Shahriari uses the data envelopment analysis (DEA) to assess the progress/regression of decision-making units (DMUs) having a two-stage structure. The progress/regression of these DMU can be assessed in the first stage, the second stage, and the whole system. In the first stage, the progress/regression of bank branches in collecting resources and in the second stage their progress/regression in allocating the resources as well as gaining profit are calculated, and the combination of both types of data is ultimately analyzed. The progress/regression is calculated using two separate indices: the The Malmquist indexand the Meta-Malmquist index. This study applied the proposed models to 20 branches of a commercial bank with two-stage structure [3].

Hua and Bin investigated the relationships between DMUs and sub-DMUs assuming the existence of undesirable outputs. In addition, the NDEA model has been presented to calculate the total and partial efficiency of DMUs [4].

Badi' Zadeh et al. developed a DEA network model to calculate efficiency optimistically and pessimistically. Their proposed model was able to combine undesirable outputs and rank the supply chains in terms of efficiency scores. They demonstrated the effectiveness of their proposed model using a case study [5].

Akbarian et al. presented an integrated framework of balanced scorecard and DEA models to measure performance over time and along with strategies based on lag of key performance indicators (KPI) for a BSC model. Causal relationships over time between the perspectives of the BSC model were initially drawn as a dynamic BSC. Then, after identifying the structure of NDEA, a new objective function was developed to measure the efficiency of nine refineries affiliated to Iran's National Petroleum Refining and Distribution Company over time and along with strategies [6]. Rahmani et al. used a new technique based on hybrid method to evaluate the efficiency score of two-stage production processes where the data in the required model are imprecise. Their proposed method considers the expected distance of the values as fuzzy variable and uses the convex combination of their two endpoints to measure the efficiency of each step and the overall efficiency score for various α values [7].

Fernandez Castaro Since 1994, the data envelopment analysis model was added to the bankruptcy prediction models and it was found that this model as a supplement can be a good model for bankruptcy prediction along with other methods such as regression, and in 2004 this model was independently used for prediction and provided very good prediction accuracy

[8]. Afterwards, DEA models, which were previously only used to measure the efficiency of the same sets, due to the high accuracy that it provided in predicting bankruptcy in early researches and due to the advantages, it had, were considered more than before. The research conducted in the field of bankruptcy prediction using DEA in foreign countries are presented in Table 2.

In the field of bankruptcy prediction, far fewer researches have been done in Iran compared to outside Iran, where only one research has been done regarding forecasting using DEA. Below is the list of the most important researches done in Iran. In other researches, generally the models designed abroad have been used without changing the coefficients, which does not seem correct due to the difference in the environment (the difference in Iran's economic environment).

In past foreign and domestic researches, all companies were placed in one group and analyzed. Considering the nature of DEA that all members of the set (DMUs) must have the same nature, this issue is one of the weaknesses of previous researches. For this reason, in this research, the manufacturing companies of the stock exchange are selected as the statistical population, and they themselves are placed in three categories: food and pharmaceutical industries, metal and automobile industries, and machinery and chemical and petrochemical industries, that it increases the accuracy of the analysis and its higher generalizability, which can solve the main problem of previous researches.

Moreover, the use of principal component analysis technique helps to reduce the dimension in the inputs of DEA model to prevent the false increase of efficiency of DMUs.

Advantages such as simultaneously calculating the efficiency of companies in comparison with each other and their ranking and predicting their bankruptcy, and higher accuracy of this method compared to other methods based on the research conducted in recent years, requires more studies to check the power of this model in predicting bankruptcy, and this research has been done in this direction.

2- Statement of the problem

So far, much researches have been conducted on the prediction of bankruptcy of companies, and of course, different results have been obtained, which are discussed more in the background section of the research. However, the use of new methods and of course old models in new formats has always been the focus of financial researchers.

The simpler and more understandable models, the more attention they receive. One of the models that has recently received more attention and use for predicting the financial future of companies is the data envelopment analysis model, which provides a forecast for the future of the company (healthy or bankrupt) by measuring the efficiency of companies.

The main problem in past works is that all companies were analyzed in one group. Considering the nature of DEA that all the members of the set (DMUs) must have the same nature, this issue is considered one of the weaknesses of the previous researches, which in this research has been tried to solve this problem by categorizing the companies.

In such a situation, the predictability of each category can be evaluated separately which will provide a more complete result about the nature of this model in predicting the financial status of companies.

2-1- Integration of DEA and PCA

Since developed by Jenkins and colleagues, DEA has been used as an effective tool for evaluation and modeling. It is equal to the weighted ratio of the outputs to the weighted ratio of the inputs The weakness of this method is that the number of evaluated units is related to the number of input and output variables. Jenkins, et al delivered a study on the greater the number of variables, the less powerful the analysis will be to distinguish between efficient and inefficient units [9]. Therefore, it is necessary to reduce the number of variables for use in the DEA model. It is obvious that such a reduction should be in such a way that it has the least effect on the differentiation of efficient and inefficient units. Based on an empirical formula

$$(\text{Output} + \text{Input}) \times 3 < \text{Units to be evaluated.}$$

If the evaluation units (DMUs) are not more than three times the sum of inputs and outputs, we need to reduce the dimensions of the problem and purposefully reduce the outputs (otherwise, many units will be on the efficiency frontier and their efficiency will be the same, which will cause incorrect results, and as a result, lower quality of the model)

For this purpose, Jenkins and colleagues have used the partial covariance matrix to eliminate variables that are highly correlated with each other. In 2003, Alder and colleagues used the PCA method instead of the main outputs or inputs that enter the DEA model and replaced the main variables with input-oriented and output-oriented components Alder et al [10] set a similar approach has been used to evaluate privatized airline networks to measure airport quality [11] and to select variables and DEA models by Cinca et al [12]. Bruce et al. [13] have used PCA and DEA methods to evaluate performance in the Internet banking industry with an approach like that of Sinsa and colleagues. As a result of these researches, a combined model of DEA and PCA is used to reduce the dimensions of the data set in each of the categories and to rank the decision-making units and finally predict bankruptcy in each of the categories. For this purpose, the information related to inputs and outputs is first extracted and then the ratio of single outputs to single inputs is calculated and then the PCA method is applied on the ratio of single outputs to single inputs. With proper selection of the first few main components, the necessary reduction in the number of variables will be observed. In the following, the selected main components are used and analyzed as inputs to the DEA model. In many applications where the number of evaluated units is small compared to the number of input and output variables, so that other methods are not responsive in their analysis, this method works well.

2-2- Proposed methodology

As stated in the introduction, in past foreign and domestic researches, to predict bankruptcy through the DEA model, all companies were placed in one group and analyzed. Considering the nature of DEA that all members of the set (DMUs) must have the same nature, this issue is one of the weaknesses of previous researches. For this reason, in this research, the manufacturing companies of the stock exchange are selected as the statistical population, and they in return are classified into three categories: food and pharmaceutical companies, metal and automobile and machinery companies, and finally Chemical and petrochemical companies, which it increases the accuracy of the analysis and its generalizability. The basis of this classification is the production system and similar risks for each category. For example, food and pharmaceutical companies that are placed in the same category, both have a continuous production system and are similar in terms of raw material storage (perishable materials) and distribution and transportation of products, and as a result, their financial risks and non-financial are similar. The second category is similar in terms of the main raw material (metals) and the production, storage, and distribution system; The third category includes companies that are not included in the first and second groups and includes chemical and petrochemical factories that are not dependent on metal industries and most of them have a continuous production system and have been experimentally proven. In terms of financial risks, they are a function of economic conditions and oil prices. Of course, different categories have been made in different books and websites of the companies in the stock market; This general classification of production companies is only one of the types of classifications. For each category, first the inputs and outputs from the available sources, i.e., the financial statements of the companies (profit and loss statement and balance sheet) are extracted. Inputs include operating costs, current liabilities and financial costs, and outputs include net profit, operating profit, and current assets. Inputs and outputs have been selected according to past research and available sources.

Mohammad Gholiha,A., Designed a Data envelopment analysis with two-step structure which is fine for the companies can break-down their work flow as a two stage NDEA [14].

Then the ratio of each of the outputs to the inputs are obtained separately for each company (which finally results in 9 financial ratios), which according to the incremental nature of these ratios (given that the outputs are placed in the numerator, the larger the fraction, the more favorable it is), these ratios are considered a type of output, therefore, the DEA model used in this research does not have an input, and for this reason, we set the input of all decision-making units (DMU) as one. In fact, if any other fixed number is placed, there will be no change in the result).

Shahriari, M. extended a new VIKOR method as a compromise ranking approach to solve multiple criteria decision-making (MCDM) problems through intuitionistic fuzzy analysis. Using compromise method in MCDM problems contributes to the selection of an alternative as close as possible to the positive ideal solution and far away from the negative ideal solution, concurrently. Using Atanassov intuitionistic fuzzy sets (A-IFSs) may simultaneously express the degree of membership and non-membership to decision makers (DMs) to describe uncertain situations in decision-making problems [15].

I.M. Premachandra, et al utilizes an additive super-efficiency data envelopment analysis (DEA) model to introduce a novel evaluation index that utilizes two frontiers to forecast triumph or downfall of corporations. The proposed technique is tested on a random set of 1001 firms, comprising 50 large American insolvent companies picked from Altman's bankruptcy database, and 901 prosperous matching firms from various sectors, representing the largest bankrupt firms from 1991 to 2004. Our research findings reveal that the DEA model has limited capability to predict corporate failures, in comparison to predicting the prosperity of firms, and the evaluation index overcomes this weakness by providing decision-makers with several alternatives to attain varying precision levels of bankrupt, non-bankrupt, and overall predictions.[16].

M. Musavi et al exposed that the anticipation of corporate failure is a crucial aspect of auditing firms' risk and uncertainty management. The creation of

dependable models for predicting bankruptcy is essential in various decision-making processes. Despite numerous bankruptcy prediction models, assessing their relative performance often results in unidimensional and conflicting outcomes. In this study, we suggest a multi-criteria assessment approach using an orientation-free super-efficiency data envelopment analysis model to overcome this methodological challenge. Moreover, we conduct a thorough comparative analysis of various UK bankruptcy modeling frameworks, including our own models. We also investigate two significant research questions: firstly, whether some modeling frameworks outperform others by design and, secondly, how the selection and nature of explanatory variables impact modeling framework performance [17].
Anja Cielen et al presented increasingly acknowledged that optimization techniques can be useful in resolving several machine learning issues. This study contrasts the classification efficiency of a linear programming model, a data envelopment (DEA) model, and a rule induction (C5.0) model. Our results indicate that, in terms of accuracy and utilization, the DEA model outperforms the other models [18].
Due to the nature of inputs and outputs, the type of DEA model used in this research should be output-oriented BCC (due to the fact that the change in the inputs does not cause a change in the outputs in the same proportion and the efficiency is a variable scale, so the BCC model is a more suitable model, and considering that in this research, the DEA inputs are fixed and as a result, the goal is to maximize output to increase efficiency, output-oriented model should be used). In the following, the multiplicative and envelopment models of BCC are given.

$$
\begin{aligned}
&\min \sum_i v_i x_{ip} + v'_p = \phi_p \\
&s.t \\
&\sum_r u_r y_{rp} = 1 \\
&\sum_i v_i x_{ij} - \sum_r u_r y_{rj} + v'_p \geq 0 \quad \forall j \\
&u_r, v_i \geq \varepsilon, v'_p \ \ free
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
&\max \phi_p + \varepsilon\left[\sum_i s_i^- + \sum_r s_r^+\right] \\
&s.t. \\
&\sum_j \lambda_j x_{ij} + s_i^- = x_{ip} \quad \forall i \\
&\sum_j \lambda_j y_{rj} - s_r^+ = \phi_p y_{rp} \quad \forall r \\
&\sum_j \lambda_j = 1 \\
&\lambda_j, s_i^-, s_r^+ \geq 0, \phi_p \ \ free
\end{aligned}
\tag{2}
$$

The number obtained in the DEA model (efficiency) is a number between zero and one. According to the cutoff point obtained in the sample, the necessary prediction for the bankruptcy or health of the company is made. (In order to be able to use the calculated efficiency score for companies, it is necessary to introduce a point between zero and one that is the best point for separating healthy and bankrupt from each other (cut-off point model) in such a way that classification error should be minimized) If the efficiency number of the company is higher than the cutoff point, the company is healthy and if it is less, the company is predicted to be bankrupt.
The general assumption is that healthy and efficient companies have a higher efficiency score and are close to one. In fact, they will have a higher score compared to inefficient companies whose score should be close to zero.
As stated above, the goal is to measure the ability of the model to predict bankruptcy in 1 year and 2 years before the event. For this purpose, at first, the year of

bankruptcy (based on Iranian Law of Commerce No. 141) of each of the helpless companies is extracted as the "year of occurrence"; And in order for healthy companies to have a specific year of occurrence, after determining the year of occurrence of a helpless company, among healthy companies of the same group, companies that have a similar size (amount of assets) will also have the same year of occurrence. Then the required financial information for one year and two years before the year of occurrence is prepared and placed in the model. In order to obtain the cut-off point, the financial information on the year of occurrence of all three groups is included in the model, so that the efficiency score for each of the three groups is obtained separately, and the graph of the distribution of the obtained efficiency score for both healthy and bankrupt categories is drawn, where the point of intersection of these two graphs (where the least classification error occurs) will be the cutoff point, and the basis of prediction is placed, as stated, if the efficiency number of the company in each category is higher than the category's cutoff point, the company is healthy, and if it was less, the company is expected to be bankrupt. Of course, in one of the past researches, the number 0.5 (without special calculation and basis) was chosen as the cutoff point (meaning that if the obtained efficiency of each company is more than 0.5, the company would be healthy, and if it is less than 0.5, the company would be predicted to be bankrupt. But in this research, the same method as stated (for higher accuracy) is used.

This comparison is done separately for one year and two years before the year of occurrence so that the power of the model can be tested in predicting the event that happened in the past.

And finally, it will be explored that the mentioned model has correctly predicted the percentage of bankrupt companies and the percentage of healthy companies in one year and two years before the year of occurrence, in order to determine the accuracy of this model. Afterwards, using the ratio test at the 95% level, the research hypotheses are evaluated whether the companies are classified correctly by the model according to the expected percentage or not.

$$\frac{(p-p_0)}{\left(\frac{p_0 q_0}{n}\right)^{0.5}} \tag{3}$$

where p is the percentage of companies classified by the correct model, P0 is the determined accuracy percentage (expected percentage) which based on previous research is selected to be 0.5, and q0 is equal to 1-P0.

Next, at the confidence level of 95%, it is checked whether there is a significant difference between the predicted ratios of the three groups or not; If there is a difference, it can be concluded that this model can be useful for predicting the bankruptcy of some groups and not necessarily all companies.

$$\frac{p_1-p_2}{\left(\left(\frac{p_1 q_1}{n_1}\right)+\left(\frac{p_2 q_2}{n_2}\right)\right)^{0.5}} \tag{4}$$

where P1 is related to one of the three groups and P2 is related to another group, which in total three two-by-two comparisons for each year will be done. Based on the above explanation, the general procedure of the research is shown in Figure 1.

2-3- Research variables

In this research, independent variables are financial ratios that are used to classify companies into two groups: bankrupt and healthy. In fact, it can be said that the efficiency score obtained by using DEA model is also an independent variable, and finally, the belonging of a company to the group of healthy or bankrupt companies, which is a function of the efficiency score, is considered a dependent variable.

3- Implementing the proposed methodology and analyzing the findings

As stated before, the number of 52 manufacturing companies present in Tehran Stock Exchange in three groups of food and pharmaceutical industry companies, metal and automobile and machinery industry companies, and chemical and petrochemical industry companies, according to the information, 21 (10 bankrupt and 11 healthy) companies were selected in the first group, 18 (10 bankrupt and 8 healthy) companies were in the second group, and 13 (7 bankrupt and 6 healthy) companies were in the third group. The names and information of these companies and their related information have been extracted from the Codal website (www.Codal.ir).

First, to calculate the cutoff point of the companies in the first group, we put the financial information of the year of occurrence of the companies in this group into the model. The shape of the efficiency score distribution of healthy and bankrupt companies of the first group is shown in Figure 2.

2-3- Research variables

In this research, independent variables are financial ratios that are used to classify companies into two groups: bankrupt and healthy. In fact, it can be said that the efficiency score obtained by using DEA model is also an independent variable, and finally, the belonging of a company to the group of healthy or bankrupt companies, which is a function of the efficiency score, is considered a dependent variable.

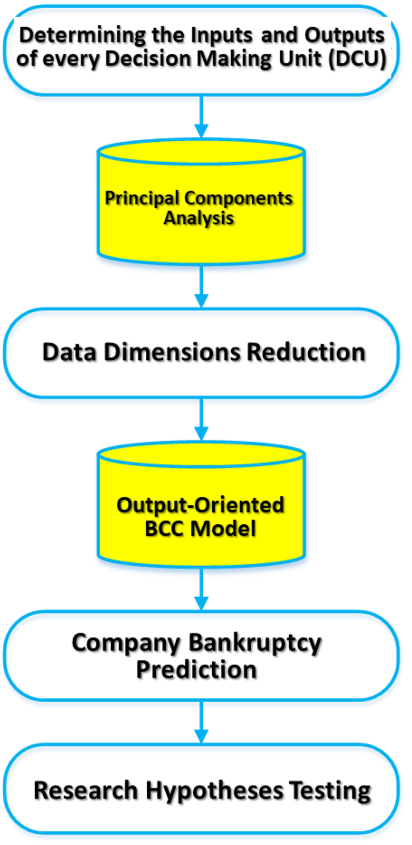

Figure (1) conceptual model of the research

3- Implementing the proposed methodology and analyzing the findings

As stated before, the number of 52 manufacturing companies present in Tehran Stock Exchange in three groups of food and pharmaceutical industry companies, metal and automobile and machinery industry companies, and chemical and petrochemical industry companies, according to the information, 21 (10 bankrupt and 11 healthy) companies were selected in the first group, 18 (10 bankrupt and 8 healthy) companies were in the second group, and 13 (7 bankrupt and 6 healthy) companies were in the third group. The names and information of these companies and their related information have been extracted from the Codal website (www.Codal.ir).

First, to calculate the cutoff point of the companies in the first group, we put the

financial information of the year of occurrence of the companies in this group into the model. The shape of the efficiency score distribution of healthy and bankrupt companies of the first group is shown in Figure 2.
According to Figure 1, the cutoff point of this group is equal to 0.59. In fact, as it can be seen from the graph, from the cutoff point onwards, as expected, the efficiency score of healthy companies becomes higher and closer to one, and the efficiency score of bankrupt companies becomes lower and closer to zero (in fact, the lowest classification error occurs)

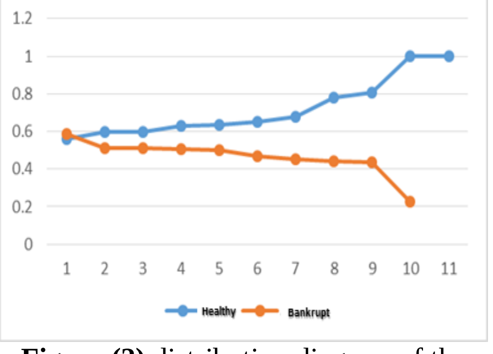

Figure (2) distribution diagram of the efficiency score of the first group companies

As stated in section 2-2, for each category, inputs and outputs were first extracted from available sources, i.e. companies' financial statements (profit and loss statement and balance sheet). Inputs include operating costs, current liabilities and financial costs, and outputs include net profit, operating profit and current assets. Then, the ratio of each of the outputs to the inputs is obtained separately for each company, which is a total of 9 ratios. As a result, it is very low due to the small number of decision-making units (companies) and causes errors in the DEA model. Therefore, they are reduced by PCA and then entered the output-oriented BCC DEA model, and finally the efficiency score is obtained, which is a number between 0 and 1.

Based on the calculations, one year before the occurrence, the model was able to identify 73% of the healthy companies and 90% of the bankrupt companies, and in the two years before the occurrence, it was able to correctly predict 73% of the healthy companies and 80% of the bankrupt companies.
Regarding the group of metal industries and automobile and machinery companies, Figure 3 shows the calculation of the cutoff point for this group.

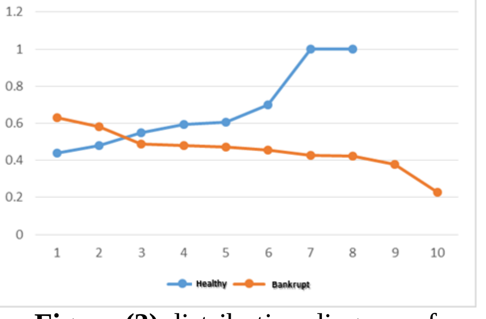

Figure (3) distribution diagram of efficiency score of second group companies

As can be seen, the value of the cutoff point for this group is equal to 0.52. One year before the occurrence, the model has been able to correctly predict 100% of healthy companies and 10% of bankrupt companies, and two years before the occurrence, 75% of healthy companies and 80% of bankrupt companies.
Regarding the group of chemical and petrochemical companies, chart 4 shows the calculation of the cutoff point for this group.

As can be seen, the value of the cutoff point for this group is equal to 0.65. One year before the event, the model was able to correctly predict 83% of healthy companies and 85% of bankrupt companies, and two years before the event, 34% of healthy companies and 71% of bankrupt companies.

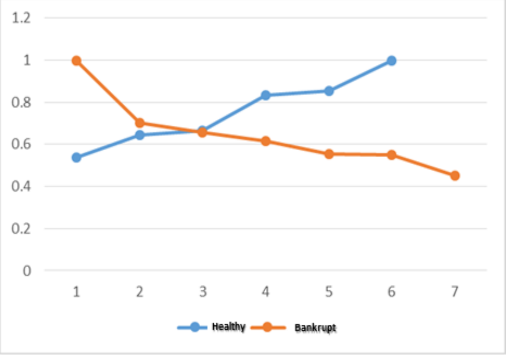

Figure (4) Distribution diagram of efficiency score for the third group of companies

As can be seen, the model has the necessary ability to predict the first group of bankrupt companies (on the verge of bankruptcy) in one year and two years before its occurrence; But this power of differentiation does not exist in the case of healthy companies.

The second hypothesis:

H_0: It is not possible to predict bankrupt companies in the metal industry and automobile and machinery companies in the stock exchange by using the combination of DEA and PCA models.

H1: It is possible to predict bankrupt companies in the metal industry and automobile and machinery companies in the stock exchange by using the combination of DEA and PCA models.

At this stage, we test the following hypotheses:

The first hypothesis:

H_0: It is not possible to predict bankrupt companies in the food and pharmaceutical industries in the stock exchange using the combination of DEA and PCA models.

H_1: The feasibility of predicting financially insolvent companies in the food and pharmaceutical sectors of the stock market was examined by employing a combination of DEA and PCA models. This hypothesis was evaluated using a ratio test with a 95% confidence level, and the findings are presented in Table 1.

It is evident that the model has the required capability to forecast the first group of companies nearing bankruptcy, one or two years before the actual event. However, this discrimination power is absent in the case of healthy companies.

The second hypothesis:

H_0: It is not possible to predict bankrupt companies in the metal industry and automobile and machinery companies in the stock exchange by using the combination of DEA and PCA models.

H1: It is possible to predict bankrupt companies in the metal industry and automobile and machinery companies in the stock exchange by using the combination of DEA and PCA models.

In the second hypothesis, the model has the ability to predict the critical situation of the second group companies in 2 years before, but it is not good in the previous year, and regarding the health of the companies in the previous year, the model was able to predict 100% correctly, but in two years before the year of occurrence, it could not provide a good result.

The third hypothesis:

H_0: It is not possible to predict bankrupt companies in the chemical and petrochemical industries in the stock exchange using the combination of DEA and PCA models.

H_1: It is possible to predict bankrupt companies in the chemical and petrochemical industries in the stock exchange by using the combination of DEA and PCA models.

Regarding the companies of the third group, the model had the ability to predict healthy and bankrupt companies well in one year before the occurrence, but it did not have the ability to distinguish healthy and bankrupt companies in two years before the occurrence.

Table (1) - the results obtained for hypothesis 1

Year	**Percentage of correct prediction by the model (percentage)**	**Expected percentage**	**Z Value**	**Result**
One year before the occurrence for healthy companies	73	50	1.526	H0 was supported
One year before the occurrence for bankrupt companies	90	50	2.530	H1 was supported
Two years before the occurrence for healthy companies	73	50	1.526	H0 was supported
Two years before the occurrence for bankrupt companies	80	50	1.897	H1 was supported

Table (2) - The results obtained for hypothesis 2

Year	**Percentage of correct prediction by the model (percentage)**	**Expected percentage**	**Z Value**	**Result**
One year before the occurrence for healthy companies	100	50	2.828	H1 is Confirmed
One year before the occurrence for bankrupt companies	10	50	2.530 -	H0 is Confirmed
Two years before the occurrence for healthy companies	75	50	1.414,	H0 is Confirmed
Two years before the occurrence for bankrupt companies	80	50	1.897	H1 is Confirmed

Table (3) - The results obtained for hypothesis 3

Year	Percentage of correct prediction by the model (percentage)	Expected percentage	Z Value	Result
One year before the occurrence for healthy companies	83	50	1.746	H1 is Confirmed
One year before the occurrence for bankrupt companies	85	50	1.852	H1 is Confirmed
Two years before the occurrence for healthy companies	34	50	0.847 -	H1 is Confirmed
Two years before the occurrence for bankrupt companies	71	50	1.111	H1 is Confirmed

Table (4) - the results obtained for hypothesis 4

Year	Comparison	Z Value	Result
One year before the occurrence	1 Vs 2	5.96	H1 is Confirmed
	1 Vs 3	0.303	H0 is supported
	2 Vs 3	4.546 -	H1 is Confirmed
Two years before the occurrence	1 Vs 2	0	H0 is Confirmed
	1 Vs 3	0.422	H0 is Confirmed
	2 Vs 3	0.422	H0 is Confirmed

The fourth hypothesis:

H_0: There is no significant difference between the results obtained from the three groups for the separation of healthy and bankrupt companies.

H_1: There is a significant difference between the results obtained from the three groups for the separation of healthy and bankrupt companies.

As we can see in the tables, there is a significant difference between the results of the three groups obtained in one year before the occurrence, while there is no significant difference between the results in the two years before the occurrence.

According to the results obtained in three groups, the following analyzes can be presented:

According to the structure of the food and pharmaceutical companies, if they are weak in their performance and unable to meet the needs of their customers, they lose them over time and continuously, and this issue is directly reflected in their finances index and therefore their financial performance can be an indicator to predict their future situation.

Regarding the companies of the second group, it should be said that due to the fluctuations in the prices of the main metals and the price of the currency, it is somewhat difficult to predict the financial crises of this group. Also, the cut-off point of this group was low, which is due to the low efficiency score of the companies in the sample in the year of occurrence, which led to the current results, and the result may be slightly different in other samples that are tested.

Fluctuations in oil prices and currency prices have a strong impact on companies whose main ingredient is oil or have competing imported goods. The extreme fluctuations in the price of oil in the ten-year period that is the time domain of this research has caused instability in the second and third group companies, and perhaps because of this, the model used in the research, whose main input is financial indicators, practically didn't have much power in estimating the situation and financial crises of these two groups.

4- Results

So far, various methods and models such as Altman, multiple diagnosis method, artificial neural networks, etc. have been used to predict bankruptcy and different results have been obtained; some methods have produced good results, and some have failed in making a successful prediction.

The use of the DEA model using the efficiency score is considered a new method for predicting bankruptcy, and in this short period of time, there have been successful and unsuccessful results in this matter, and therefore, there would be a need for more research to investigate the power of this model.

Despite the fact that the DEA model is used to calculate the efficiency of homogeneous and similar groups, in the works that had been done so far, all the investigated companies were placed in one group and evaluated; In this research, it was decided to make a small grouping and put the production companies that are active in similar fields in their own group and evaluate each group separately, it was also decided to use the more financial indicators than previous works. But since if the number of decision-making units (here, the production companies of each group) in DEA models are close to the total number of inputs and outputs of the model, they cause problems in the results of the model. Therefore, there was a need for a method to reduce the dimensions, that is why the PCA statistical method was used. As a result, the generalizability of the model was also practically increased, because it will be possible to use more groups of this model.

In a similar work in the past, where all companies were evaluated at once by the DEA model, it was shown that it is possible to predict bankruptcy up to 2 years before it occurs, but it was not clear which model is more capable of predicting which group of companies, and results were obtained in generic manner. While in the existing research we see that the results of the groups are different and the BCC model of DEA with the help of the PCA statistical technique has a better performance in predicting the bankruptcy of food and pharmaceutical companies and can predict up to 2 years before the occurrence of bankruptcy (based on Iranian Law of Commerce No.141). This model can predict the bankruptcy of chemical and petrochemical companies in 1 year before its occurrence, and its performance for predicting the bankruptcy of companies in the metal and automobile and machinery industries is less, and instead, it predicts the health of healthy companies in this group with high accuracy. The results obtained in this research show that the act of classifying companies based on their functions and the risks they are faced in specific groups and evaluating them separately using the model can be beneficial in obtaining acceptable results and provides researchers with more complete information.

What is important is that one can never rely on any of the different bankruptcy prediction methods (even though they have provided successful results in the past), including the method proposed in this research, because what is done is a prediction and not a prophecy, and for more success, one should always use several mathematical and statistical methods and models along with fundamental analyzes (interpreting and predicting the future based on political, economic, military, social issues, etc.) to increase the probability of a successful prediction.

References

[1] Altman, E, (1968); " Financial ratio, discriminal analysis and the prediction of corporate bankruptcy " the journal of finance, 23: 589-609

[2] Mehrani, Sasan and Mehrani, Kaveh and Monsefi, Yashar and Karmi, Gholamreza (2004); "An applied survey of Zemiski and Shirata's bankruptcy prediction models in companies listed on Tehran Stock Exchange" Accounting and Auditing Reviews, 12 (41)

[3] M. Shahriari, Malmquist Productivity Index for Two-stage Structures and its Applications in Bank Branches, International Journal of Industrial Mathematics (IJIM), (2011), Volume 4, Issue 4, Pages 11

[4] Hua, Z., & Bian, Y. (2008). Performance measurement for DEA network with undesirable factors. International Journal of Management and Decision Making, 9(2), 141-153.

[5] T. Badiezadeh, RF Saen, and T. Samavati, "Assessing sustainability of supply chains by double frontier network DEA: A big data approach," Comput. Oper. Res., vol. 98, pp. 284–290, 2018.

[6] Akbarian, Mojtaba, Esmaeil Najafi, Reza Tavakkoli-Moghaddam, and Farhad Hosseinzadeh-Lotfi. "A network-based data envelope analysis model in a dynamic balanced score card." Mathematical Problems in Engineering 2015 (2015).

[7] Rahmani, Amir, Mohsen Rostamy-malkhalifeh, and Farhad Hosseinzadeh Lotfi. "Evaluating performance of Two-Step Networks Using Fuzzy Data Envelopment Analysis." Revista Gestão & Tecnologia 20 (2020): 96-105

[8] Fernandez Castaro, A and Smith, P (1994); " Towards a general nonparametric model of corporate performance " OMEGA journal, 237-149.

[9] Jenkins, Larry, & Anderson, Murray (2003) "A multivariate statistical approach to reducing the number of variables in data envelopment analysis", European Journal of Operation Research, 147: 51-61

[10] Adler, N., & Golany, B. (2003) "Including principal component weights to improve discrimination in data envelopment analysis", Journal of the Operations Research Society of Japan, 46 (1): 66-73

[11] Adler, N., & Golany, B. (2001) "Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe", European Journal of Operational Research 132: 260-273

[12] Cinca, C. Serrano, & Molinero, CM (2004) "Selecting DEA specifications and ranking units via PCA", Journal of the Operational Research Society, 55(5): 521–528

[13] Bruce Ho, C., Wu, DD (2008) "Online banking performance evaluation using data envelopment analysis and principal component analysis", Computers & Operations Research, Article in Press

[14] Mohammad Gholiha, A. Hosseinzadeh Lotfi, F., Shahriari, M., Vaez-Ghsemi, M., (2021)., The Relationship between Financial

Investment and Malmquist Productivity Index Based on Data Envelopment Analysis in Network Structure (Case Study: Tehran Metro Stations), Journal of Investment Knowledge 10 (40), 103-118,
[15] Shahriari, M., (2017). Soft computing based on a modified MCDM approach under intuitionistic fuzzy sets, Iranian Journal of Fuzzy Systems 14 (1), 23-41

[16] lI.M. Premachandra a, Yao Chen b, DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment, Volume 39, Issue 6, ecember 2011, Pages 620-626, Omega.

[17] M. Mousavi a, Performance evaluation of bankruptcy prediction models: An orientation-free super-efficiency DEA-based framework, International Review of Financial Analysis, Volume 42, December 2015, Pages 64-75

[18] Anja Cielen, Ludo Peeters, Koen Vanhoof, Bankruptcy prediction using a data envelopment analysis, European Journal of Operational Research
Volume 154, Issue 2, 16 April 2004, Pages 526-532