1. Introduction

Listed companies’ financial distress prediction is important to both listed companies and investors. However, due to the uncertainty of business environment and strong competition, even companies with perfect operation mechanism have the possibility of business failure and financial bankruptcy. So whether listed companies’ financial distress can be forecasted effectively and timely is related to companies’ development, numerous investors’ interest, and the order of capital market [1–3].

Most topical studies have adopted a multiple-variable approach to the prediction of financial distress by combining accounting and nonaccounting data in a variety of statistical formulas [4–7]. While the predictive value of accounting information was based on samples of industrials or on nonindustrials, the misclassification rates were low. Hence the explanatory variables had a significant predictive power. Ratios based on accounting earnings reported cash flow and book debt figured prominently in various statistical formulas, especially those applied to the industrial sector such as univariate analysis, multiple discriminant analysis, logit, and probit model [8–11]. Although these methods use history samples to create diagnostic model, they cannot inductively learn from new data dynamically. This greatly affects the forecasting accuracy. More recently, many studies have demonstrated that artificial intelligence such as decision trees, neural networks, and support vector machine can be alternative methods for financial distress prediction [12–14]. Moreover, there are various studies on the comparison between statistical and machine learning methods in terms of their ability to predict financial data [15–18].

On the other hand, the recent researches have exploited multiple attribute decision making (MADM) methods in financial analysis to improve the final outputs [19–22].

This paper puts emphasis on optimizing the financial distress forecasting in the case of listed companies in Tehran Stock Exchange (TSE) of Iran by the hybrid approach which outperforms existing discrete models significantly. Because of the importance of financial ratios to describe a company’s situation, factor analysis is applied to summarize the effect of financial ratios; then the combinations of all of them are exploited. Subsequently, the extracted predictors are utilized to forecast financial distress in a hybrid approach through traditional statistical modeling distress and machine learning methods for classifying business. In this analysis, another important issue is homogenizing business via clustering method to improve prediction models. Also, MADM method
is used to distinguish the best model via different classification performances measures. Consequently, a comparison of the final results shows that the prediction of the financial distress is significantly consolidated.

The paper is organized as follows: Section 2 presents a short review of the literature in the field of financial distress forecasting. Section 3 briefly describes the applied methods. Then Section 4 explains the proposed approach and also the empirical evidence from Iran is presented. The paper ends with concluding remarks in Section 5.

2. Review of the Literature

The early prediction of distress is essential for companies and investors or lending institutions that wish to protect their financial investments. As a consequence, modeling, prediction, and classification of companies to determine whether they are potential candidates for financial distress have become key topics of debate and detailed research.

Corporate bankruptcy was first modeled, classified, and predicted by Beaver [23]. He defined financial distress as bankruptcy, insolvency, and liquidation for the benefit of a creditor, firms defaulting on loan obligations or firms miss preferred by dividend payments. In this study, “cash flow to total debt” had the highest discriminatory power of the ratios examined. Altman’s model is perhaps the best known of the early studies [4]. He developed a $Z$-score bankruptcy prediction model and determined a cut point of $Z$-score (2.675) to classify healthy and distressed firms. The results showed that the $Z$-score model had a sound prediction performance one year and two years before financial distress but did not indicate good prediction utility three to five years before financial distress. A number of authors such as Taffler [24, 25], Pantalone and Platt [26], Betts and Belhoul [27], and Piese and Wood [28] followed his work and applied the $Z$-score model into different markets, different time periods, and different industries. Also, Deakin [29] and Blum [30] used multiple-variable statistical techniques subsequent to Altman [31].

Furthermore, most recent studies have adopted a multiple-variable approach to the prediction of financial distress by combining accounting and nonaccounting data in a variety of statistical formulas. In the reviewed literature, 64% of all authors used statistical techniques whose overall predictive accuracy was 84%. 25% of the authors used machine learning models whose overall accuracy was 88%, and 11% of the authors used theoretical models whose accuracy was calculated as 85% [32]. Table 1 briefly presents some recent researches in financial distress forecasting.

<table>
<thead>
<tr>
<th>Method(s)</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applying support vector machines to bank bankruptcy analysis using practical steps [33]</td>
<td>Erdogan, 2013</td>
</tr>
<tr>
<td>Presenting particle swarm optimization techniques to obtain appropriate parameter settings for subtractive clustering and integrates the adaptive-network-based fuzzy inference system (ANFIS) [16]</td>
<td>Chen, 2013</td>
</tr>
<tr>
<td>Applying a simple hazard model to develop an early warning system of bank distress in the gulf cooperation council countries [34]</td>
<td>Maghyereh and Awartani, 2014</td>
</tr>
<tr>
<td>Presenting a statistics-based wrapper for SVM-based financial distress identification by using statistical indices of ranking-order information from predictive performances on various parameters [35]</td>
<td>Li et al., 2014</td>
</tr>
<tr>
<td>Examining the effect of the filter and wrapper based feature selection methods and applying different classification techniques [36]</td>
<td>Liang et al., 2015</td>
</tr>
</tbody>
</table>

3. Methodology

In this section, the methods applied in our paper are briefly described. Factor analysis, $K$-means method, discriminant analysis, logit model, decision trees, neural network, and TOPSIS are presented, respectively.

3.1. Factor Analysis. Factor analysis is a dimension reduction method of multivariate statistics, which explores the latent variables from manifest variables. Two methods for factor analysis are generally in use, principal component analysis and the maximum likelihood method. The main procedure of principal component analysis can be described in the following steps when applying factor analysis [40].

Step 1. Find the correlation matrix ($R$) or variance-covariance matrix for the objects to be assessed.

Step 2. Find the eigenvalues ($\lambda_k$, $k = 1, 2, \ldots, m$) and eigenvectors ($\beta_k = [\beta_{1k}, \beta_{2k}, \ldots, \beta_{mk}]$) for assessing the factor loading ($a_{ik} = \sqrt{\lambda_k} \beta_{ik}$) and the number of factors ($m$).

Step 3. Consider the eigenvalue ordering ($\lambda_1 > \cdots > \lambda_k > \cdots > \lambda_m$, $\lambda_m > 1$) to decide the number of common factors and pick the number of common factors to be extracted by a predetermined criterion.

Step 4. According to Kaiser [41], use Varimax criterion to find the rotated factor loading matrix, which provides additional insights for the rotation of factor-axis.
Step 5. Name the factor referring to the combination of manifest variables.

3.2. K-Means Method. This method clusters \( n \) objects into \( k \) (\( k < n \)) deterministic partitions by minimizing the total squared error function given by MacQueen (1967) [40, 42]:

\[
f = \min_{\mu_1, \ldots, \mu_k} \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2,
\]

where \( k \) is the number of clusters in the data, \( \mu_i \) is the center of cluster \( h \), and \( x \) is the data point in the cluster \( h \). Different solutions can be attained depending on the initial guess of cluster centers; therefore, the procedure should be repeated multiple times, and the final solution is selected as the one that gives the maximum separation between clusters.

3.3. Discriminant Analysis. Let \( X \) be a \( p \)-dimensional normal random vector belonging to class \( k \) if \( X \sim N_k(\mu_k, \Sigma) \), where \( \mu_k \) is a positive definite symmetric matrix. If \( \mu_1, \mu_2, \) and \( \Sigma \) are known, the optimal classification rule is Fisher’s linear discriminant rule,

\[
\delta_p (X) = I \left\{ \left( X - \mu_k \right)^T \Sigma^{-1} \mu - 1 \right\} > 0, \tag{2}
\]

where \( \mu_k = (\mu_1 + \mu_2)/2, \mu_d = (\mu_1 - \mu_2)/2, \) and \( I \) denotes the indicator function, with value 1 corresponding to classifying \( X \) to class 1 and 0 to class 2. Fisher’s rule is equivalent to the Bayes rule with equal prior probabilities for two classes. The misclassification rate of the optimal rule is

\[
R = 1 - \phi \left( \Delta^{1/2}_p \right), \tag{3}
\]

where \( \phi \) is the standard normal distribution function.

In practice, Fisher’s rule is typically not directly applicable, because the parameters are usually unknown and need to be estimated from the samples. Let \( \{X_{1,j}, j = 1, \ldots, n_1\} \) and \( \{X_{2,j}, j = 1, \ldots, n_2\} \) be independent and identically distributed random samples from \( N_k(\mu_1, \Sigma) \) and \( N_p(\mu_2, \Sigma) \), respectively. The maximum likelihood estimators of \( \mu_1 \) and \( \mu_2 \) and \( \Sigma \) are

\[
\overline{X}_k = \frac{1}{n_k} \sum_{j=1}^{n_k} X_{1,j}, \quad k = 1, 2, \tag{4}
\]

\[
S_n = \frac{1}{n(n-1)} \sum_{j=1}^{n_1} \sum_{j=1}^{n_2} (X_{k,j} - \overline{X}_k)^T (X_{k,j} - \overline{X}_k), \tag{5}
\]

where \( n = n_1 + n_2 \), and setting

\[
\hat{\mu}_a = \frac{\overline{X}_1 + \overline{X}_2}{2}, \tag{6}
\]

\[
\hat{\mu}_d = \frac{\overline{X}_1 - \overline{X}_2}{2},
\]

and \( \Sigma^{-1} = S^{-1}_n \) (or generalized inverse \( S_n \) when \( S^{-1}_n \) does not exist), Fisher’s rule becomes the classic LDA [40, 43]:

\[
\delta_{LDA} (X) = I \left\{ (X - \hat{\mu}_a)^T S^{-1}_n \hat{\mu}_a > 0 \right\}. \tag{7}
\]

3.4. Logit Model. Binary responses, for example, success and failure, are the most common form of categorical data and the most popular model for them is logit model. For a binary response, \( Y \), and a vector of explanatory variables, \( x \), let \( \pi(x) \) denote the success probability when \( X \) takes value \( x \). This probability is the parameter for the binomial distribution [44]. The logistic regression model has a linear form for the logit of this probability,

\[
\logit \pi(x) = \log \left( \frac{\pi(x)}{1 - \pi(x)} \right) = \alpha + \beta x. \tag{8}
\]

3.5. Decision Trees. A decision tree (DT) is a machine learning technique used in classification, clustering, and prediction tasks. A well-known tree-growing algorithm for generating DT is Quinlan’s ID3 [45]. It starts from the root node. The root node is one of the best attributes. The property values are then generated that correspond to each branch. Each branch generates a new node. For the best attributes according to the selection criteria, ID3 uses an entropy-based definition of the information gain to select the test attribute within the node. The entropy characterizes the purity of a sample set. Suppose \( S \) is a set of \( s \) data samples. We assume that the class label attribute has \( m \) different values, the definition of \( m \) different classes is \( C_i \) \( (i = 1, \ldots, m) \), and set \( S_j \) is the number of samples in class \( C_j \). Equation (8) is the given sample classification based on the expectations of the information:

\[
I (s_1, s_2, \ldots, s_m) = - \sum_{i=1}^{m} p_i \log_2 p_i, \tag{9}
\]

where \( p_i \) is the probability of any sample belonging to \( C_i \), which is estimated using \( S_i / S \).

The set attribute \( A \) has \( v \) different values \( \{a_1, a_2, \ldots, a_v\} \). A property can be divided into subsets \( S = \{s_1, s_2, \ldots, s_v\} \), where \( S_j \) contains a number of \( S \) values in this sample and they have a value of \( a_i \) in \( A \). If we select test attribute \( A \), these subsets correspond to set \( S_i \), which contains nodes derived from growing the branches. \( S_j \) assumes that \( S_{ij} \) is a subset of the samples of class \( C_j \). Thus, \( A \) can be divided into subsets of entropy or expected information, which is given by

\[
E (A) = \sum_{j=1}^{v} s_{ij} + s_{2j} + \cdots + s_{mj} I (s_{1j}, s_{2j}, \ldots, s_{mj}) / S, \tag{10}
\]

where the item \( (s_{1j} + s_{2j} + \cdots + s_{mj}) / S \) subset is on the right of the first \( j \) and is equal to the number of subsets of the sample divided by the total number of \( S \) in the sample. Equation (10) is a given subset for \( S_j \):

\[
I (s_{1j}, s_{2j}, \ldots, s_{mj}) = - \sum_{i=1}^{m} p_i \log_2 (p_i), \tag{11}
\]

where \( p_i = S_{ij} / S_j \) is a sample of \( S_j \) based on the probability of belonging to class \( C_j \). Equation (11) is a branch that will be used for encoding information:

\[
\text{Gain} (A) = I (s_1, s_2, \ldots, s_m) - E (A). \tag{12}
\]
In other words, Gain(𝐴) is attributable to a value of that property because of the expectations of the entropy of compression. Thus, a smaller entropy value leads to a lower correlation, whereas a higher corresponding information gain produces a subset of the division with higher purity. Therefore, the test attribute decision tree selects the properties with the highest information gain. This creates a node and marks the property, where each value of the property creates a branch and divides the sample accordingly.

The decision tree contains leaves, which indicate the value of the classification variable, and decision nodes, which specify the test to be carried out. For each outcome of a test, a leaf or a decision node is assigned until all the branches end in the leaves of the tree [45–47].

3.6. Neural Network. Neural network is a technique that imitates the functionality of the human brain using a set of interconnected vertices. It is based on an artificial representation of the human brain, through a directed acyclic graph with nodes (neurons) organized into layers. In typical feed-forward architecture, there are a layer of input nodes, a layer of output nodes, and a series of intermediate layers. The input signals are multiplied by their corresponding weights to give the value of output nodes, and a series of intermediate layers. The input signals are multiplied by their corresponding weights to give the value of output nodes, and as a series of intermediate layers. The input

\[ Y_j = w_0 + \sum_{i=1}^{n} w_j x_i, \quad (12) \]

where \( Y_j \) is weighted sum of input signals at node \( j \); \( w_0 \) is threshold (bias) value; \( w_j \) is the weight associated with the connection between node \( j \) and the input node \( i \); \( x_i \) is a value of input node \( i \); \( n \) is number of input nodes. A sigmoid activation function (13) is applied to the weighted sum:

\[ f(Y) = \frac{1}{1 + e^{-Y}}. \quad (13) \]

The value calculated from (13) is the output signal from node \( j \), which can be considered as the input signal to the next layer [48].

3.7. TOPSIS. TOPSIS (technique for order preference by similarity to an ideal solution) method is a popular approach to MADM (multiple attribute decision making) that has been widely used in the literature. It presented by Hwang and Yoon consists of the following steps [49].

**Step 1.** The decision matrix is normalized through the application of

\[ r_{ij} = \frac{W_{ij}}{\sqrt{\sum_{j=1}^{J} W_{ij}^2}}, \quad j = 1, 2, \ldots, J; \quad i = 1, 2, \ldots, n. \quad (14) \]

**Step 2.** A weighted normalized decision matrix is obtained by multiplying the normalized matrix with the weights of the criteria,

\[ V_{ij} = W^*_i r_{ij}, \quad j = 1, 2, \ldots, J; \quad i = 1, 2, \ldots, n. \quad (15) \]

**Step 3.** Positive indicator score \( (A^*) \) (maximum value) and negative indicator score \( (A^-) \) (minimum value) are determined by

\[ A^* = \{V^*_1, V^*_2, \ldots, V^*_n\}, \quad (16) \]

\[ A^- = \{V^-_1, V^-_2, \ldots, V^-_n\}. \]

**Step 4.** The distance of each alternative from \( A^* \) and \( A^- \) is calculated using

\[ d^*_i = \sqrt{\sum_{j=1}^{J} (V^*_j - V^*_i)^2}, \quad (17) \]

\[ d^-_i = \sqrt{\sum_{j=1}^{J} (V^-_j - V^-_i)^2}, \quad j = 1, 2, \ldots, J. \]

**Step 5.** The closeness coefficient for each alternative \( (CC_i) \) is calculated by applying

\[ CC_i = \frac{d^-_i}{d^*_i + d^-_i}, \quad i = 1, 2, \ldots, n. \quad (18) \]

**Step 6.** At the end of the analysis, the ranking of alternatives is made possible by comparing the \( CC_i \) values.

4. The Proposed Approach and Empirical Evidence

Corporate bankruptcy forecasting plays a central role in academic finance research, business practice, and government regulation to financial decision support. Consequently, accurate default probability prediction is extremely important.

The main purpose of this study is not only improving the prediction performance model through hybrid analysis approach but also employing a multiple attribute decision making (MADM) method to make optimum decision for choosing the best alternative classification. Figure 1 briefly presents the flowchart of the optimization approach.

As shown in Figure 1, the first step consists of a list of important and available financial ratios including liquidity.
measurement ratios, profitability indicator ratios, debt ratios, operating performance ratios, cash flow indicator ratios, and investment valuation ratios. In the proposed approach, all ratios would be considered because of the efficiency of forecasting approach. In fact, the factor analysis is used for dimension reduction when the numbers of predictors are high. If this method is not applied, it may cause overfitting during modeling. Also, by applying the factor analysis, our predictors are more influential and significant than before because they present more information on the listed companies.

Another suggestion in this research to present more sophisticated models is homogenizing of the company’s performance. It is fairly clear that there are a variety of businesses which can bring about the inefficiency of forecasting models. As a solution, this step tries to cluster businesses based on an influenced ratio and then exploit forecasting methods. A comparison among performance measurements presents remarkable improvement among recent financial distress modeling.

Subsequently, the extracted predictors from the factor analysis are utilized to forecast financial distress through traditional statistical modeling distress and machine learning methods in each cluster separately.

As the last but not the least step, based on the different classification performances measures, we try to choose the best model from the data set. It is because up till now there is no best classification method which can cover the best score in all evaluation measures. Different multiple attribute decision making (MADM) methods often produce different outcomes for selecting or ranking a set of decision alternatives involving multiple attributes. The TOPSIS is one of the famous MADM methods used to distinguish the best model. Also, other MADM methods can be applied. Consequently, the final results show that the prediction of the financial distress is significantly consolidated.

In the following, empirical evidence related to the proposed approach is also presented. Population under the study is the accepted manufacturing companies in Tehran Stock Exchange (TSE) for one year ended on March 21, 2011. The reason for this choice is the availability of financial information of these companies. There are 461 companies listed in TSE with 37 industry groups, of which 412 are manufacturing companies and 49 are the nonmanufacturing ones. The number of the manufacturing companies is more than other listed companies subject to granting more loans due to their extensive activities. A sample of 180 companies is chosen for this research.

In Tehran Stock Exchange, the measure for companies exiting capital market is the commercial law of 141 acts. According to those acts, companies are known as bankrupt whose retained losses are more than 50% of their capital. 58 companies are bankrupt under this law. The rest of nonbankrupt companies were randomly selected from the remaining list.

In this research, some ratios that have a high unity of views for 180 manufacturing companies quoted in Tehran Stock Exchange for one year (year ended on March 21, 2011) were used. The required data to calculate the ratios have been gathered from companies’ balance sheets and income statements. The financial ratios used in the prediction are listed in “The Definition of Variables Used.”

Popular discourse on financial distress prediction deals with the selection of important variables because of enormous financial ratios [27, 30]. Hence, in this study, to reduce the variables dimension and summarize their effects on factors by the analysis of correlation matrix, we applied the factor analysis. Consequently, Table 2 shows four common factors by the principal component method and Varimax rotation. They account for a cumulative proportion of 93 percent of the total sample variance.

It is fairly clear that the first factor represents the debt and cash flow conditions (a strong combination of DR, DER, ROS, ROE, and WCTA). The second factor almost symbolizes the liquidity conditions (a strong combination of TA, CR, and QR), and the third one approximately denotes the operating performance of the listed companies (a strong combination of SFA, STA). The last factor is the investment conditions (a combination of SI, TLDS, CLOE, and TLOE). For example, we have

$$F_1 = -0.06TA + 0.08CR + 0.06QR - DR - 0.75DER$$
$$+ 0.05SI - 0.99TLDS + 0.02SFA + 0.14STA$$
$$+ ROS + 0.98ROE - 0.03CLOE + 0.03TLOE$$
$$+ WCTA.$$  \[19\]

Then, thanks to the existing difference in the size and type of the industries, the listed companies have not been homogeneous, causing the inefficiency of prediction models. As a solution, we tried to cluster listed companies by using the K-means technique based on a ratio. The current ratio (CR) is a popular financial ratio used to test a company’s liquidity by deriving the proportion of current assets available to cover current liabilities. Therefore, it is crucial to improve distress prediction; the listed companies were clustered based on their capital positions. Table 3 shows the result. Hence, the listed

<table>
<thead>
<tr>
<th>Table 2: Estimated factor loading.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>TA</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>QR</td>
</tr>
<tr>
<td>DR</td>
</tr>
<tr>
<td>DER</td>
</tr>
<tr>
<td>SI</td>
</tr>
<tr>
<td>TLDS</td>
</tr>
<tr>
<td>SFA</td>
</tr>
<tr>
<td>STA</td>
</tr>
<tr>
<td>ROS</td>
</tr>
<tr>
<td>ROE</td>
</tr>
<tr>
<td>CLOE</td>
</tr>
<tr>
<td>TLOE</td>
</tr>
<tr>
<td>WCTA</td>
</tr>
</tbody>
</table>
Table 3: Cluster means of 3-mean method.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>0.974</td>
<td>3.425</td>
<td>9.620</td>
</tr>
</tbody>
</table>

Table 4: The linear discriminant functions for financial distress.

<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distress</td>
<td>No distress</td>
<td>Distress</td>
</tr>
<tr>
<td>Constant</td>
<td>−4.06</td>
<td>−1.61</td>
</tr>
<tr>
<td>F₁</td>
<td>5.04</td>
<td>4.79</td>
</tr>
<tr>
<td>F₂</td>
<td>0.69</td>
<td>−0.13</td>
</tr>
<tr>
<td>F₃</td>
<td>6.81</td>
<td>−0.53</td>
</tr>
<tr>
<td>F₄</td>
<td>3.21</td>
<td>2.11</td>
</tr>
</tbody>
</table>

Companied can be divided into three levels of current ratio (CR): small, average, and high.

In the next step, the usual statistical distress prediction methods are applied. First of all, the logit models are

\[
\logit \left[ \pi_{\text{clus}}(x) \right] = -2.73 - 34.66F₁ - 2.47F₂ + 0.29F₃ + 0.53F₄, \tag{20}
\]

where \( F₁ \) is the debt and cash flow conditions, \( F₂ \) is the liquidity conditions, \( F₃ \) is the operating performance, and \( F₄ \) is the investment conditions. All listed companies divided into 3rd cluster are not distressed. Also, the average error count of clustered distress is about 7.5 percent less than that without cluster analysis for logistic regression.

The next statistical method for bankruptcy prediction is discriminant analysis. Table 4 presents the linear discriminant functions for financial distress in each cluster. The average error count is significantly reduced in comparison with the general function.

In addition, to implement the machine learning methods, decision trees and neural network are applied where analysis for each cluster is not necessary separately. This is because both methods can be used for both classification and clustering purposes. Actually, in this part, the important issue is to utilize all ratios and extract factors as predictors separately.

To classify the distressed companies through decision tree, 100 companies were randomly used to build the tree and the rest for test (see Figure 2).

It means that if, for a company, \( F₁ \) (the debt and cash flow conditions) is greater than 0.41, then the company is not distressed, and if \( F₁ \) is less than 0.23 and \( F₂ \) (the liquidity conditions) is less than 0.16, then the company is distressed. To evaluate the built trees, we applied about 30 percent of data. The results showed that the tree of directly financial ratios had a greater misclassification error than that of factors (20%). In other words, comparison between two trees showed that to make tree by factors more information was used; hence a more reliable result was achieved.

At last, neural network is implemented by three hidden layers and hyperbolic function. The final result shows the overall accuracy of 94 percent, which is 12 percent less than all ratios directly used.

On the other hand, there are four models to forecast financial distress (in fact, there are more classification methods that can be applied), and we have to present the best one from our data set. As mentioned above, the TOPSIS method is used to choose the best model based on the different classification performances measures. Table 5 illustrates some classification performances measures for the applied methods, where TP is true positive, FP is false positive, TN is true negative, and FN is false negative and

\[
\text{Accuracy} = \frac{TP}{TP + FN} + \frac{FP}{TN} + FN + FN,
\]

\[
\text{Error rate} = \frac{FP}{TP} + \frac{FN}{TP} + FP + TN + FN,
\]

\[
\text{Precision} = \frac{TP}{TP + FN}, \tag{21}
\]

\[
\text{Sensitivity} = \frac{TP}{TN + FN},
\]

\[
\text{Specificity} = \frac{TN}{TN + FP},
\]

and AUC is the area under a receiver operating characteristic (ROC) curve, where a ROC space is defined by FP and TP as \( x \)- and \( y \)-axes, respectively, which depicts relative trade-offs between true positive and false positive.

As mentioned, there are various methods to choose an ideal alternative from MADM problems and as stated before, TOPSIS is one of them. In the TOPSIS approach, the best alternative is the nearest one to the ideal solution and the farthest one from the negative ideal solution. Also, it is assumed that all the criteria have identical weights and importance. Table 6 presents a brief calculation of this method, where \( V_j \) is normalized criteria (accuracy, error rate, precision, sensitivity, specificity, and AUC, resp.), \( d⁺ \) is the distance from the ideal alternative, \( d⁻ \) is the distance from the negative ideal alternative, and CC is the relative closeness to the ideal solution.

Based on the last column of Table 6, the decision trees, neural network, logit analysis, and discriminant analysis are the better models, respectively, while if decision maker wants to choose the best model by the information in Table 5, then the judgment may be probably different.

5. Conclusion

The enterprise bankruptcy forecasting has always been an important issue in the business and financial decision support. In this research, applying a hybrid approach is suggested to improve the prediction performance and give more supportive results.
First of all, factor analysis was used to determine and summarize some combinations of financial ratios correlated together. After that, K-means algorithm was used to cluster companies, homogenize them, and get much accurate results.

Later, to predict the financial distress, the multiple logistic regression analysis, the multiple discriminant analysis, the decision tree, and the neural network, that are all the famous methods in this field, were applied. Finally, the best model classifier was chosen with the help of multiple attribute decision making (MADM), the TOPSIS method.

The proposed approach, which has also been applied in Iranian companies in Tehran Stock Exchange which used to employ previous performance models, can be consolidated with the help of the hybrid analysis. The comparison among the used methods clearly showed that the decision tree and then the neural network had a remarkable performance in comparison to others.

The hybrid approach advanced provides insight into the complex interaction of the common bankruptcy prediction methods and suggests avenues for applying MADM methods in this area in the future research.

### Table 5: Classification performances measures.

<table>
<thead>
<tr>
<th></th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Accuracy</th>
<th>Error rate</th>
<th>Precision</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logit analysis</td>
<td>0.64</td>
<td>0.12</td>
<td>0.16</td>
<td>0.08</td>
<td>0.80</td>
<td>0.20</td>
<td>0.84</td>
<td>0.89</td>
<td>0.57</td>
<td>0.69</td>
</tr>
<tr>
<td>Discriminant analysis</td>
<td>0.60</td>
<td>0.13</td>
<td>0.17</td>
<td>0.10</td>
<td>0.77</td>
<td>0.23</td>
<td>0.82</td>
<td>0.86</td>
<td>0.57</td>
<td>0.65</td>
</tr>
<tr>
<td>Decision tree</td>
<td>0.74</td>
<td>0.01</td>
<td>0.20</td>
<td>0.05</td>
<td><strong>0.94</strong></td>
<td><strong>0.06</strong></td>
<td>0.94</td>
<td><strong>0.95</strong></td>
<td><strong>0.84</strong></td>
<td></td>
</tr>
<tr>
<td>Neural network</td>
<td>0.73</td>
<td>0.02</td>
<td>0.21</td>
<td>0.04</td>
<td><strong>0.94</strong></td>
<td><strong>0.06</strong></td>
<td>0.78</td>
<td><strong>0.95</strong></td>
<td>0.91</td>
<td>0.83</td>
</tr>
</tbody>
</table>

### Table 6: Calculation of the TOPSIS method.

<table>
<thead>
<tr>
<th></th>
<th>$V_1$</th>
<th>$V_2$</th>
<th>$V_3$</th>
<th>$V_4$</th>
<th>$V_5$</th>
<th>$V_6$</th>
<th>$d^*$</th>
<th>$d^-$</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logit</td>
<td>0.46</td>
<td>0.63</td>
<td>0.52</td>
<td>0.49</td>
<td>0.37</td>
<td>0.46</td>
<td>1.86</td>
<td>1.00</td>
<td><strong>0.35</strong></td>
</tr>
<tr>
<td>Disc</td>
<td>0.45</td>
<td>0.72</td>
<td>0.51</td>
<td>0.47</td>
<td>0.37</td>
<td>0.43</td>
<td>2.07</td>
<td>0.90</td>
<td><strong>0.30</strong></td>
</tr>
<tr>
<td>Deci</td>
<td>0.54</td>
<td>0.19</td>
<td>0.49</td>
<td>0.52</td>
<td>0.62</td>
<td>0.56</td>
<td>1.08</td>
<td>1.85</td>
<td><strong>0.63</strong></td>
</tr>
<tr>
<td>Neu</td>
<td>0.54</td>
<td>0.19</td>
<td>0.48</td>
<td>0.52</td>
<td>0.59</td>
<td>0.55</td>
<td>1.11</td>
<td>1.81</td>
<td><strong>0.62</strong></td>
</tr>
</tbody>
</table>
The Definition of Variables Used

TA: The total asset
CR: The ratio of the current assets to the current liabilities
QR: The ratio of the amount of cash and equivalents, short, and accounts receivable, term investments to the current liabilities
DR: The ratio of the total liabilities to the total assets
DER: The ratio of the total liabilities to the total owner’s equity
SI: The ratio of the sales to the number of the inventories
TLDS: The ratio of the total liabilities to the daily sales
SFA: The ratio of the sales to the fixed assets
STA: The ratio of the sales to the total assets
ROS: The ratio of the net income to the sales
ROE: The ratio of the net income to the average owner’s equity
CLOE: The ratio of the current liabilities to the owner’s equity
TLOE: The ratio of the total liabilities to the owner’s equity
WCTA: The ratio of the working capital to the total assets.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References


