Journal of Retailing and Consumer Services 27 (2015) 11–23

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

journal homepage: www.elsevier.com/locate/jretconser

A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring

Fatemeh Nemati Koutanaei [a], Hedieh Sajedi [b,*], Mohammad Khanbabaei [c]

[a] Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Saveh, Iran
[b] Department of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran
[c] Department of Information Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form
30 June 2015
Accepted 4 July 2015

Keywords:
Credit scoring
Classification
Feature selection
Ensemble learning
Data mining

ABSTRACT

Data mining techniques have numerous applications in credit scoring of customers in the banking field. One of the most popular data mining techniques is the classification method. Previous researches have demonstrated that using the feature selection (FS) algorithms and ensemble classifiers can improve the banks' performance in credit scoring problems. In this domain, the main issue is the simultaneous and the hybrid utilization of several FS and ensemble learning classification algorithms with respect to their parameters setting, in order to achieve a higher performance in the proposed model. As a result, the present paper has developed a hybrid data mining model of feature selection and ensemble learning classification algorithms on the basis of three stages. The first stage, as expected, deals with the data gathering and pre-processing. In the second stage, four FS algorithms are employed, including principal component analysis (PCA), genetic algorithm (GA), information gain ratio, and relief attribute evaluation function. In here, parameters setting of FS methods is based on the classification accuracy resulted from the implementation of the support vector machine (SVM) classification algorithm. After choosing the appropriate model for each selected feature, they are applied to the base and ensemble classification algorithms. In this stage, the best FS algorithm with its parameters setting is indicated for the modeling stage of the proposed model. In the third stage, the classification algorithms are employed for the dataset prepared from each FS algorithm. The results exhibited that in the second stage, PCA algorithm is the best FS algorithm. In the third stage, the classification results showed that the artificial neural network (ANN) adaptive boosting (AdaBoost) method has higher classification accuracy. Ultimately, the paper verified and proposed the hybrid model as an operative and strong model for performing credit scoring.

1. Introduction

Recently, banks and financial institutions have extensively started to consider the credit risk of their customers. In order to differentiate customers for offering credit services to them and managing their risks, banks have needed to apply credit scoring systems in their procedures (Gray and Fan, 2008). Lately, non-parametric approaches and data mining practices have been used in the area of customer credit scoring. The statistical methods, non-parametric methods, and artificial intelligence (AI) approaches have been suggested in order to provision the credit scoring developments. In addition, ensemble credit scoring methods have been used in many studies. It should be mentioned that a noticeable number of researches have shown that ensemble learning classification approaches in credit scoring have a better performance in comparison with single classifiers. With respect to the review of these studies, there are nine main approaches in credit scoring researches as provided in the following:

1. Single-classifier credit scoring models.
2. Multiple-classifier credit scoring models.
3. Credit scoring models based on statistical methods.
4. Credit scoring models based on AI methods.
5. Linear and non-linear credit scoring models.
6. Parametric credit scoring models, including linear probability model, discriminant analysis model, probit and logit models, etc.
7. Non-parametric (data mining) credit scoring models, including decision tree, K nearest-neighbor (KNN) model, expert system, ANN, fuzzy logic, GA, etc.
8. Ensemble learning credit scoring models.
9. Hybrid credit scoring models.

* Corresponding author.
E-mail address: hhsajedi@ut.ac.ir (H. Sajedi).

http://dx.doi.org/10.1016/j.jretconser.2015.07.003
0969-6989/

Many researchers have employed the above-mentioned approaches in their investigations. Hu and Ansell (2007) utilized some algorithms, including Naïve Bayes, logistic regression (LR), recursive partitioning, ANN, and sequential minimal optimization (SMO) in their study. In a study by Min and Lee (2008), they applied the credit scoring model based on data envelopment analysis (DEA). In another study, link analysis ranking method with the SVM was used for credit scoring (Xu et al., 2009). Setiono et al. (2009) used GA to optimize the KNN classification algorithm in credit scoring. Moreover, Yeh and Lien (2009) compared the data mining techniques, including KNN, LR, discriminant analysis, Naïve Bayes, ANN, and decision trees. Zhou et al. (2009) used direct search for parameters selection in the SVM classification algorithm. In a study by Ping and Yongheng (2011), neighborhood rough set and the SVM-based classifier were used for credit scoring. In another study (Kao et al., 2012), Bayesian latent variable model with classification regression tree was employed. Vukovic et al. (2012) used the preference theory functions in the case-based reasoning (CBR) model for credit scoring model. Danenas and Garsva (2015) applied particle swarm optimization (PSO) for the optimal linear SVM classifier selection in the domain of credit risk.

As cited above, recently, the ensemble credit scoring models have been used in a number of researches. Tsai and Wu (2008) applied multilayer perceptron (MLP) neural network ensembles for the credit scoring problem. In an investigation by Nanni and Lumini (2009), an ensemble of classifiers, including bootstrap aggregating (Bagging), Random Subspace, Class Switching, and Random Forest, was involved in the credit scoring. In addition, the ensemble of classifiers, including ANN, decision tree, Naïve Bayes, KNN, and logistic discriminant analysis was applied by Twala (2010). Hsieh and Hung (2010) utilized bagging ensemble classifier, including ANN, SVM, and Bayesian network. In another study by Paleologo et al. (2010), the subagging ensemble classifier, including kernel SVM, KNN, decision trees, AdaBoost, and subagged classifiers, was used in the credit scoring. Wang and Ma (2012) proposed a hybrid ensemble learning approach using SVM as a base learner for enterprise credit risk assessment.

Several studies have deployed the FS approach in their credit scoring models. Wang and Huang (2009) applied evolutionary-based FS approaches in a case study of credit approval data. Tsai (2009) compared five famous FS methods used in bankruptcy prediction, which were *t*-test, correlation matrix, stepwise regression, PCA, and factor analysis, in order to examine their performance by using MLP neural networks. Chen and Li (2010) proposed a combined strategy of FS approaches, including Linear Discriminant Analysis (LDA), rough set theory, decision tree, F-score and SVM classification model in credit scoring. In a research by Wang et al. (2012), the rough set and scatter search meta heuristic in FS were used for credit scoring. Chen (2012) developed an integrated FS and a cumulative probability distribution approach based on rough sets in credit rating classification. Hajek and Michalak (2013) suggested an approach to combine the mixed and individual FS methods with well-known machine learning models, such as MLP, radial basis function (RBF), SVM, Naive Bayes, random forest, LDA, and nearest mean classifier in corporate credit rating prediction. Oreski and Oreski (2014) presented a new hybrid GA with ANN to identify an optimum feature subset in order to increase the classification accuracy and scalability in credit risk assessments. Liang et al. (2015) deployed three filters including LDA, *t*-test, and linear regression, and two wrappers including GA and PSO based FS methods, combined with six different prediction models, namely linear SVM, RBF SVM, KNN, Naive Bayes, classification and regression tree (CART), and MLP under some experiments in bankruptcy and credit scoring datasets.

The above-mentioned studies have been stated from three main viewpoints as follows: (1) General credit scoring studies (2) Ensemble credit scoring studies (3) FS based credit scoring studies. This article is differentiated from the rest of the papers due to the simultaneous consideration of these three viewpoints. It is worth mentioning that past studies have considered only one or two of these viewpoints. In addition, it should be stated that the main aim of this article is to propose a proper FS algorithm and an appropriate base and ensemble classifier via three types of evaluation approaches, i.e., the SVM classification accuracy (only for FS), classification accuracy, and the area under the receiver operating characteristic curve (AUC) for classifiers and parameters setting (for both) in the context of hybrid credit scoring model. Moreover, many studies have not examined the effect of several FS methods and classifier parameter setting on the credit scoring problem. As another distinguished aspect, on the basis of the aforementioned attentions, in the present paper, nine approaches are combined in order to build a new hybrid FS and ensemble learning credit scoring model. The proposed model is a combination of FS techniques and several base (single) classifiers and ensemble classifiers in the parametric (owing to the Naïve Bayes algorithm) and non-parametric approaches of credit scoring. The parameters setting of four FS algorithms and two types of classification algorithms (base and ensemble) are used. For each FS algorithm, the performance is examined in terms of the SVM classification accuracy measure. The SVM is an influential learning method for classification problems. As cited by Brown and Mues (2012), “it is based on construction of maximum-margin separating hyper plane in some transformed feature space”. It should be indicated that SVM is one of the most popular techniques used in the literature. Then, SVM is utilized to evaluate the performance FS algorithms. Moreover, the classification algorithms are compared according to the classification accuracy and AUC measures. For experimental results, the dataset of the ‘Export Development Bank of Iran’ is used. In the hybrid model, four FS algorithms are used as follows: (1) PCA; (2) GA; (3) Information gain ratio; and (4) Relief algorithm. Furthermore, two types of classification algorithms prevalent in the previous studies are as follows: (1) Base classification algorithms: Naïve Bayes, CART decision tree, SVM, and ANN; (2) Ensemble classification algorithms: bagging, AdaBoost, random forest, and staking. The results can confirm that the hybrid model of credit scoring has a robust functioning in comparison with the other classification algorithms presented in this paper.

As an abstract representation, the main contributions of this study reflected in the proposed model are as follows:

1. Providing a comprehensive study by comparing different FS methods and classifiers, with respect to the credit scoring problem.
2. Hybrid simultaneous use of three general, ensemble, and FS based credit scoring approaches.
3. Using FS algorithms and comparing their performance with the aid of the accuracy measure of the SVM classification algorithm and also the accuracy and AUC measures of the base and ensemble classifiers.
4. Employing the parameters setting procedures for FS and classification algorithms in order to improve the credit scoring performance with an iterative manner.
5. Simultaneous use and comparison of the base and ensemble learning classification algorithms in the proposed credit scoring model.

6. Using and comparing nine approaches of credit scoring models employed in the literature in an integrated framework.
7. Whereas credit scoring in the studies is mostly based on real customers, in the current study, the credit scoring model is built based on legal customers.

This paper has been structured as follows. In Section 2, the related methods used in this paper are briefly described. In Section 3, the experimental design is presented, including the dataset description and pre-processing, performance evaluation, and development of the hybrid credit scoring model. By using a case study, the experimental results and discussions are elaborated in Section 4. Finally, Section 5 is devoted to the conclusions as well as the future recommendations of the paper.

2. Background

2.1. Credit scoring

Thomas defined credit scoring as a process of recognizing bank customers in order to grant them credit based on a set of predefined criteria (Yu et al., 2009). In general, according to Ong et al. (2005), credit scoring models have several advantages as provided here: (1) Decreasing costs of credit analysis; (2) Assignment of credits with an effective and rapid decision making process; (3) Higher probability of credit repayment; and (4) Lower possible risks. In 1936, Fisher expressed the concept of statistical discrimination analysis, which was the basis of the credit scoring field. Afterwards, David Durand in 1941 applied several methods to differentiate between the good and bad loans. Then in 1960, credit cards appeared and banks started to apply credit scoring for their businesses. In 1980, banking experts encouraged using credit scoring in credit cards and it was a starting point to apply credit scoring methods in other products (Thomas, 2000). In the beginning, credit scoring was implemented by a judging attitude of an expert by reviewing the application form and expressing a 'yes' or 'no' argument as the final decision. They applied 3Cs, 4Cs, or 5Cs (i.e., the character, capital, collateral, capacity, and conditions of the customer) in their customer credit scoring evaluations (Thomas, 2000). The investigation carried out by Yu et al. (2009) in credit scoring depicted many statistical and optimization methods were employed in recent studies, such as LDA, logistic analysis, probit analysis, linear programming, integer programming, KNN, and classification tree. Newly, a number of studies in credit scoring have concentrated on the AI techniques, such as ANN, evolutionary computation (EC), GA, and SVM (Yu et al., 2009), which are more capable to distinguish between good and bad customers than the statistical and optimization methods. In addition, the ensemble and hybrid credit scoring models have been employed by researchers, and their findings have exposed that these approaches can present a higher performance than the single and statistical approaches of credit scoring, depicted in the literature.

2.2. Feature selection algorithms

Feature selection algorithms, generally as pre-processing methods of model creation, can be used to increase the classification performance. They have a number of benefits as follows (Salappa et al., 2007): (1) Decreasing the noise in dataset; (2) Reducing the computational cost in order to successfully acquire proper models; (3) Helping to better understand the final models in the classification algorithms; (4) Simple application; and (5) Assisting in updating the model. There are three important subjects in FS: (1) Evaluation measure; (2) Searching behavior; and (3) Stopping rule. There are five types of evaluation criteria: (1) Information; (2) Dependency; (3) Distance; (4) Consistency; and (5) Classification accuracy. Moreover, there are typically three types of search methods in the FS: (1) Complete; (2) Heuristic; and (3) Random search (Wang and Li, 2008). Furthermore, the stopping rules have been provided by Wang and Li (2008): (1) Determine the maximum iteration number; (2) Do not change the performance after adding or removing a feature; and (3) The ideal feature sub-division has been found.

In this paper, three types of FS algorithms are applied, which are GA, relief method, and information gain ratio. Furthermore, one type of FS algorithm (i.e., PCA) is employed. In the following, the algorithms are described.

1. Genetic algorithm feature selection: In this method, one chromosome is a set (group) of the features of bank customers. Gene is a customer's feature that its type of encoding is binary, and the values of (1) and (0) respectively mean that there is and is not a particular feature in the set of credit scoring features. Goldberg strategy is used to discover an ideal set of variables (features). Thereafter, the subset evaluator function with n-fold cross-validation is applied to evaluate the input variables. Additionally, the subset of features is assessed according to the classification accuracy measure of the SVM algorithm. Finally, the initial population, maximum number of generations, mutation, crossover probability, cross validation, and random seed number were 20, 20, 0.01, 0.9, 10, and 1, respectively.
2. Relief method feature selection: as quoted in Rapid Miner software (version 4.1 beta 2), (RapidMiner version 4.1 beta 2, 2001–2007), this method “evaluates the value of a feature by repetitively sampling a case and considering the value of the specified feature for the nearest case of the same and different class”.
3. Information gain ratio feature selection: it is based on the information entropy concept (Brown and Mues, 2012). By considering the WEKA machine learning tool (Waikato Environment for Knowledge Analysis), in this method, the value of an attribute is assessed by determining the information gain ratio (entropy difference) with respect to the class. Information gain ratio (Class, Attribute) is equal to H (Class) minus H (Class | Attribute) (Witten and Frank, 2005).
4. Principal component analysis feature selection: PCA is a transformation process to reduce the number of features by extraction of the new independent features (Sustersic et al., 2009). In this statistical dimensionality reduction technique, the correlated features can be combined as principal components. According to Sustersic et al. (2009), there are several principal components as eigenvector of the variance–covariance matrix of the original variables.

2.3. Base (single) classification algorithms

1. Decision tree: this classification algorithm is one of the most popular algorithms used by researchers in credit scoring studies. Decision tree is a model with a top-down tree structure that contain several nodes, branches, and leaves. Each node belongs to a variable or an attribute. On the other hand, branches divide the data into smaller datasets, and leaves encompass the class value that assigns every observation to one of the leaves (Brown and Mues, 2012). In this study, the decision tree classifier CART is applied, which contains two

branches for each decision node. In order to split the training data, the CART classifies the transactions into the subsets of transactions with similar values for the target variable. CART employs an exhaustive search of all attributes and all possible splitting values for each decision node in the tree growing process under a recursive manner in order to achieve the optimal splitting measure (Larose, 2005).

2. Artificial neural network: this widespread classification algorithm is based on non-parametric approaches and has been frequently used in credit scoring problems. ANN is appropriate for non-linear problems with a suitable flexibility. In the current study, the most extensively used category of ANN is utilized (i.e., MLP), which contains three layers. The input layer includes neurons for all input variables and the output layer has one neuron. The relationship between neurons in layers is based on a weight and bias and they are applied for training the network. Each neuron in the hidden and output layers is activated by an activation function. During training the network, the weights and biases are adjusted to decrease an objective function and increase the classification accuracy. The training procedure is an iterative process based on the gradient descent learning with learning rate criterion (Brown and Mues, 2012).
3. Support vector machine: SVM has been developed by Vapnik and is a supervised and non-parametric machine learning algorithm. Recently, SVM has been used in numerous credit scoring studies. As depicted by Ping and Yongheng (2011), the basic approach of SVM is based on minimization of the structural risk by constructing an optimum separating hyper plane: $w \,.\, x + b = 0$. In this study, the linear, polynomial, and RBF kernels are used to optimize the hyper plane.
4. Naïve Bayes: this algorithm is based on the parametric approaches and probabilistic learning method. As cited by Twala (2010), “this classifier applies the Bayes rule to calculate the probability of class label C_i given all attributes A_j and predict the class with the highest posterior probability”. The probability of a class value *Ci* given an instance *X* for *n* observations is displayed in Eq. (1).

$$p(C_i|X) = \prod_{j=1}^{n} p\left(A_j|C_i\right).\ \ p(C_i) \tag{1}$$

2.4. Ensemble classification algorithms

There are many applications of ensemble classification algorithms in credit scoring. Ensemble learning is based on the machine learning approach, where several learning algorithms can be employed in order to solve one problem (Wang et al., 2011). It is in contrast to the single classification algorithms. Recently, ensemble learning algorithms have been applied in credit scoring studies. In Section 1, some of the studies in ensemble credit scoring models have been demonstrated. In continuance, four main ensemble learning models are described, i.e., bagging, AdaBoost, stacking, and random forest.

1. Bagging: this algorithm was developed by Breiman (1996). Bagging is based on the majority voting concept, where different training data subsets are randomly used for training a dissimilar base learner of the similar manner (Wang et al., 2011).
2. AdaBoost: this algorithm was advanced by Freund and Schapire (1996), and is a famous member of the boosting approach. It creates base classifiers via sequential bootstrap samples gained by weighting the training transactions via numerous iterations. Weighting is adjusted by misclassification related to the base classifier (Marques et al., 2012).
3. Stacking: it is used to combine different learning algorithms in order to achieve higher prediction accuracy by building base learners (Wang et al., 2011). As cited by Wang et al. (2011), “stacking combines the prediction of the multiple base-level base learners by a meta-level base learner” (Wang et al., 2011).
4. Random forest: it is a collection of un-pruned decision trees and a tree generation procedure via random FS, which has been trained according to the bootstrap trials of the training data. It is based on the tree voting procedure of the most popular class (Brown and Mues, 2012).

3. Experimental design

3.1. Dataset description and data pre-processing

The dataset used for assessing the performance of the proposed model in this paper is dependent on the legal customers of the 'Export Development Bank of Iran' and within the period of two years. In continuance, the variables employed in the proposed model are described (see Appendix A). The type of variables has been determined as discrete (D) or continuous (C) as follows:

Return on sales (C), total quality score (out of 130) (C), claims cycle (C), cycle operation (C), cycle inventories (C), risk of target markets (D), experience with the bank (D), history of the company/background activity (D), records of senior managers (D), legal personality (D), seasonal factors (D), activity in the domestic market (D), territory of the foreign market (D), working capital/circulating capital (C), current assets flow (C), non-current assets flow (C), interest rate of fixed capital (C), interest rate of investments (C), equity dividend rate (C), current liabilities to equity ratio (C), non-current liabilities to equity ratio (C), percentage of covering the financing costs (C), current ratio (C), non-current assets to equity ratio (C), quick ratio (C), equity ratio (C), debt ratio (C), credit expert 1's verdict (D), credit expert 2's verdict (D), and fulfillment of obligations (D) (target variable).

In the dataset, the target variable is a two-class problem and can be defined as follows: good and bad customers are those that their repayments are made before and after two months, respectively. The dataset included the records of 1100 legal customers, and the number of features was 59.

In order to use the dataset for the proposed model, data preprocessing should be done. In here, there are some data pre-processing methods (without any sequence in the operations) applied for data preparation and cleaning as listed in the following: (1) Removing some of the records as well as the un-valued features because of having missing values; (2) Integrating data; (3) Data transformation (i.e., (a) transforming two features into one feature, (b) converting quantitative variables into financial ratios, and the alternate discrete variables into binary variables); (4) Eliminating some of the ineffective features; (5) Normalization; (6) Determining the correlation between two variables by Pearson’s test; (7) Data visualization; (8) New feature creation; and (9) Outlier detection (finding outlier by the Shewhart control chart and replacing them with a medium value for continuous features and a mode value for discrete features).

After data pre-processing, 30 features were selected for the final experiment. In addition, the number of records after data preparation decreased to 777 records (customers).

3.2. Performance evaluation

In order to evaluate the experiments, some measures including the classification accuracy and the AUC were applied. The description of these measures can be clarified with respect to the

Predicted classification \ Actual classification	Positive (Good customers)	Negative (Bad customers)
Positive (Good customers)	True Positive (TP)	False Positive (FP)
Negative (Bad customers)	False Negative (FN)	True Negative (TN)

Fig. 1. Confusion matrix (Wang et al., 2011; Liang et al., 2015).

confusion matrix as illustrated in Fig. 1. The abbreviations in the confusion matrix are as follows: *TP*: true positive; *TN*: true negative; *FP*: false positive; and *FN*: false negative. The definition of these abbreviations can be explained with respect to the confusion matrix, as shown in Fig. 1. The classification accuracy has been depicted in Eq. (2) (Wang et al., 2011; Liang et al., 2015)

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{2}$$

The second evaluation measure applied in the paper is the AUC. As explained by Brown and Mues (2012), the receiver operating characteristic curve (ROC) is a two-dimensional diagram, where in each dimension, one concept is illustrated. The diagram shows the interaction between the true positive rate (sensitivity) and the false positive rate (1-specificity). In order to compare several classifiers, the AUC is evaluated and if the AUC for a classifier is greater than the other ones, its classification performance would be better.

3.3. The proposed model

In this section, the hybrid data mining model for credit scoring is described. Fig. 2 shows the block diagram of the proposed model. It contains three main stages: (1) Data gathering and pre-processing; (2) Feature selection; and (3) Modeling (classification). In the first stage, after data gathering, some pre-processing methods are utilized (see Section 3.1). In the second stage, four FS algorithms are used in order to achieve better features in the subset in order to gain a higher performance of the classification algorithms in the next stage (i.e., in the modeling). The FS algorithms are as the following: (1) GA; (2) Relief method; (3) Information gain ratio; and (4) PCA. In this stage, parameters setting of all FS techniques is carried out and the acquired results (i.e., the selected features) from the FS methods are used for the SVM algorithm. The optimal parameters are determined by the SVM classification accuracy measure for each FS algorithm and this is implemented through the iteration process of parameters setting. In this regard, 70% and 30% of the data were used for training and testing, respectively. The accuracy index was used to compare the FS algorithms. The output of the second stage is selection of the best FS algorithm as well as the best selected features subset which are applied for classification algorithms in the modeling stage. In the third stage, four base classifiers were used; i.e., CART, ANN, SVM, and Naïve Bayes, and also ensemble classification algorithms, such as Naïve Bayes–AdaBoost, CART–AdaBoost, ANN–AdaBoost, SVM–AdaBoost, Naïve Bayes-bagging, CART-bagging, ANN-bagging, SVM-bagging, random forest, and stacking.

The training (90%) and testing (10%) datasets were used for the building base and ensemble credit scoring models via 10-fold cross-validation. In order to evaluate the classification algorithms, two measures are applied as follows: (1) classification accuracy; and (2) the AUC. Via these measures, the best classification algorithm is used for credit scoring of bank customer. It is noticeable that for weaker models, with respect to the iterative process, better parameters values are selected and utilized for the classification algorithms in the modeling stage.

As stated in the introduction section, several studies have employed hybrid and ensemble strategies for their credit scoring problems. Moreover, they have used several single and multiple classifiers to resolve credit scoring difficulties. FS algorithms were applied to achieve a higher performance out of the credit scoring classification model. However, the main shortage of these studies is the need to ensure a comprehensive approach for the simultaneous application of several FS algorithms, including single, multiple, and ensemble classifiers, in a hybrid data mining framework. In this regard, it is evident that the essential unique aspect of this framework is the use of a parameters setting procedure in the FS and modeling stages of the proposed model. This can increase the confidence level of the final model to be used in the credit scoring problems. In addition, because of using different datasets in the credit scoring problems, only some of the FS algorithms can be consistent with the dataset for obtaining a better classification performance. Furthermore, it would be better to use several FS algorithms and expand a parameters setting procedure. Besides, this problem is true for applying several classification algorithms in credit scoring studies. The proposed model in this study aims to decrease the above-mentioned problems by considering the strength points of previous studies.

4. Experimental results and discussions

In this section, the sample dataset of the 'Export Development Bank of Iran' is used for evaluation of the model. Rapid Miner data mining software (version 6.0.8) is used for the experimentation. This software is appropriate for data mining problems and contains a group of machine learning procedures. In Rapid Miner, excluding the parameters setting cited in the following tables, other parameters values of the FS algorithms are default according to the common mode in the literature.

In the proposed model, after data gathering and pre-processing in the first stage (see Section 3.1), the FS algorithms are applied and their parameters are adjusted with respect to the SVM classification accuracy in the second stage. In continuance, the results of parameters setting (adjusting) for four FS algorithms are described.

4.1. Genetic algorithm and parameters setting

In genetic algorithm FS, the 'population size' parameter was changed. Maximum number of generations, mutation and crossover values were 30, 0.01, and 0.5, respectively. The output of FS is used for SVM classification model. The results are demonstrated in Table 1.

With respect to Table 1, the selected features resulted from the values of the 'population size' parameter (5, 10, and 15) in the SVM classification were employed, and the outcomes showed that the 'population size' parameter value (5) has higher classification accuracy and AUC simultaneously than others. In Table 1, it can be seen that the SVM classification accuracy and AUC values of Model 1 (population size of 5) are 88.41% and 90.45%, respectively. Therefore, it would be better to use its selected features in the base and ensemble learning classification, in continuance of the proposed model. The selected features using GA are stated as follows: non-current liabilities to equity ratio, interest rate of investments, equity dividend rate, risk of target markets, seasonal factors, history of the company/ background activity, records of senior

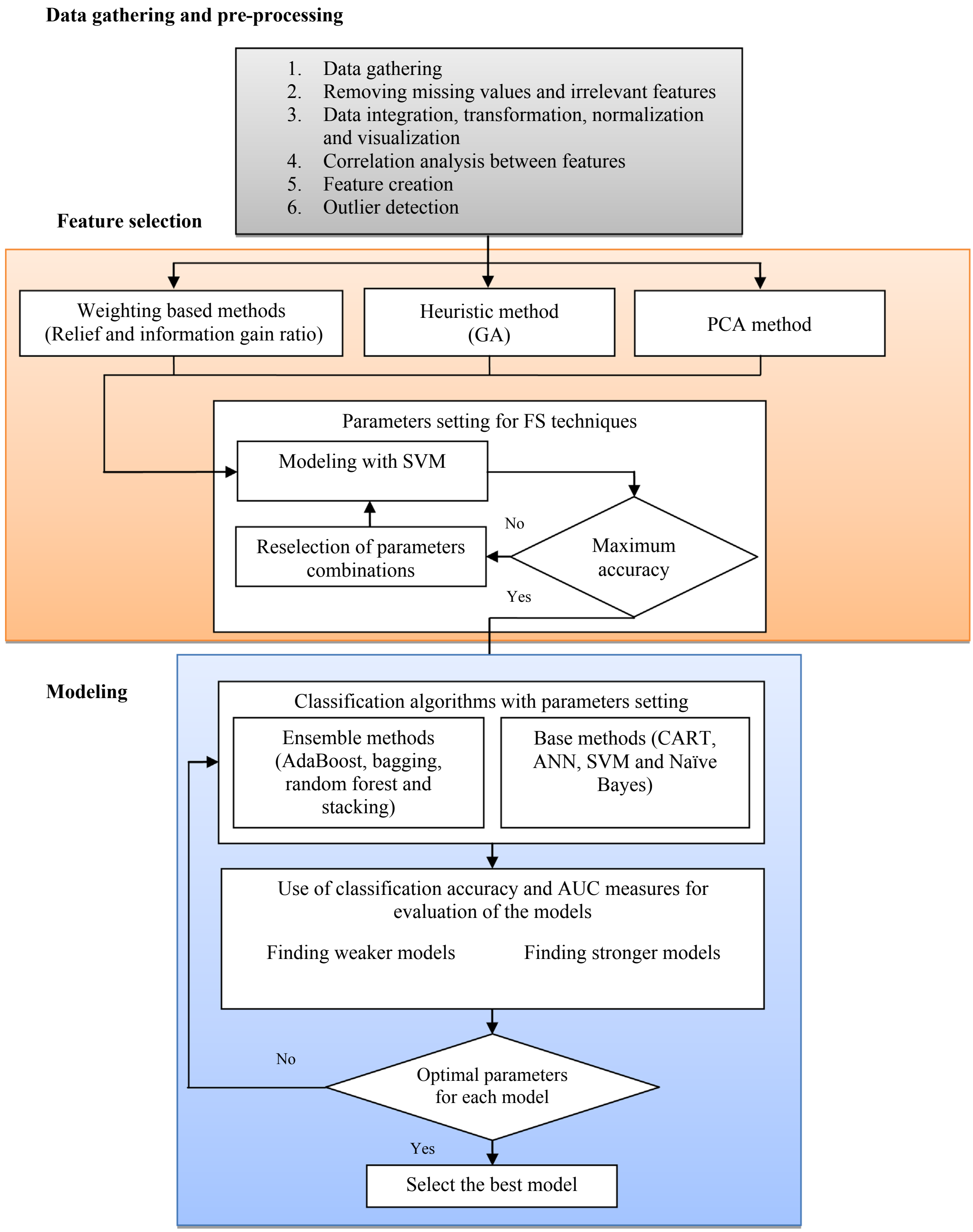

Fig. 2. The block diagram of the proposed model.

managers, credit expert 2's verdict, and total quality score (out of 130).

4.2. Relief method and parameters setting

In the relief method FS, the number of selected features and the nearest neighbors were changed. The selected features resulted in each change were used for SVM classification. The results for

Table 1
Parameters setting in FS based on GA.

Model	Population size	Accuracy (%)	AUC (%)
1	5	88.41	90.45
2	10	87.70	90.17
3	15	88.41	90.13

Table 2
Parameters setting in FS based on relief method.

Model	Number of selected features	Number of nearest neighbors	Accuracy (%)	AUC (%)
1	10	5	21.46	50.00
2	10	10	78.54	50.00
3	10	15	78.54	50.00
4	10	20	78.54	50.00
5	10	25	78.54	50.00
6	15	5	78.54	50.00
7	15	10	78.54	50.00
8	15	15	78.54	50.00
9	15	20	78.54	52.00
10	15	25	78.54	52.00
11	20	5	78.54	50.00
12	20	10	78.54	50.00
13	20	15	78.54	50.00
14	20	20	78.54	50.00
15	20	25	78.54	50.00

Table 3
Parameters setting in FS based on information gain ratio.

Model	Number of selected features	Accuracy (%)	AUC (%)
1	10	90.99	86.27
2	13	77.25	55.81
3	16	76.39	50.00
4	19	78.97	58.71
5	22	80.69	52.22
6	25	78.11	50.00

classification are demonstrated in Table 2.

With respect to Table 2, models 9 and 10 have higher performance in the SVM classification. The values of the accuracy and AUC measures are 78.54% and 52%, respectively. Therefore, the selected features of models 9 and 10 can be used for classification models. As can be noticed, Table 2 presents that the AUC value for each model, except models 9 and 10, is equal to each other; because by changing the 'number of selected features' and the 'number of nearest neighbors' parameters, the AUC value for each model does not change. After using the relief method, more important selected features, where their weights were more than 0.1 value, were obtained as follows respectively: credit expert 1's verdict, legal personality, total quality score (out of 130), credit expert 2's verdict, territory of the foreign market, activity in the domestic market, equity ratio, debt ratio, current liabilities to

Table 4
Parameters setting in FS based on PCA.

Model	Number of factor	Accuracy (%)	AUC (%)
1	10	87.12	83.73
2	20	87.12	84.26
3	30	87.12	84.14
4	40	87.12	84.17
5	50	87.12	84.10

equity ratio, non-current assets to equity ratio, interest rate of investments, working capital/ circulating capital, seasonal factors, risk of target markets, and experience with the bank.

4.3. Information gain ratio method and parameters setting

In the information gain ratio FS, only the number of selected features was changed. The selected features resulted in each change were used for SVM classification. The results are demonstrated in Table 3.

According to Table 3, model 1 has a higher performance in the SVM classification. The values of the accuracy and AUC measures are 90.99% and 86.27%, respectively. Then, the selected features of model 1 are used for classification models. Using information gain ratio FS, more important selected features where their weights is more than 0.1 value, are presented respectively, as follows: total quality score (out of 130), equity dividend rate, non-current liabilities to equity ratio, non-current assets to equity ratio, current assets flow, percentage of covering the financing costs, current liabilities to equity ratio, debt ratio, equity ratio, non-current assets flow, credit expert 1's verdict, quick ratio, current ratio, working capital/ circulating capital, and credit expert 2's verdict.

4.4. Principal component analysis method and parameters setting

In the PCA feature reduction, the number of factor was altered. The selected features resulted in each change were used for SVM classification. The results are demonstrated in Table 4.

As presented in Table 4, model 2 has a higher performance in the SVM classification. The values of accuracy and AUC measures are 87.12% and 84.26%, respectively. Then, the selected features of model 2 were applied for classification models. As can be discerned, Table 4 presents that the accuracy value for each model is equal to each other; because changing the parameter of 'number of factor' does not change the accuracy value for each model. The PCA feature selection used only 19 continuous variables in its

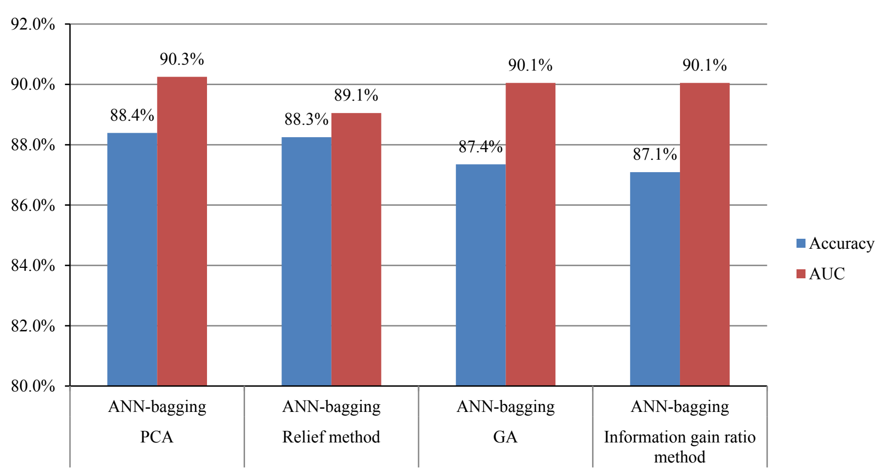

Fig. 3. The results of FS algorithms.

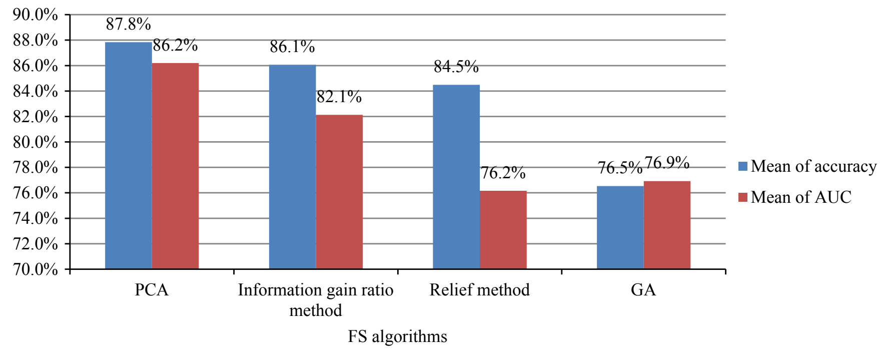

Fig. 4. The results of FS algorithms via the mean of accuracy and the mean of AUC measures in all of the classification algorithms.

computations and after implementation of the algorithm, 13 new variables were extracted. The new extracted variables have been deployed from the combination of previous 19 variables. In the new extracted variables, the mean of the relative importance of previous variables is computed and the more important selected features, where their weights were more than 0.15 value, were presented as follows respectively: non-current liabilities to equity ratio, claims cycle, non-current assets to equity ratio, current assets flow, non-current assets flow, equity dividend rate, return on sales, interest rate of fixed capital, total quality score (out of 130), working capital/ circulating capital, and current liabilities to equity ratio.

After execution of FS and adjusting the parameters, in order to find a suitable subset of features, all base and ensemble classification algorithms are implemented for each of the best model of FS algorithms resulted from Tables 1–4. Fig. 3 depicts the best classification results for each FS algorithm. For example, by using GA feature selection in all classification algorithms, ANN-bagging algorithm has the best performance with the accuracy and AUC measures of 87.4% and 90.1%, respectively.

As depicted in Fig. 3, for all FS algorithms, ANN-bagging classification model is the best in the classification performance. In addition, it is obvious that ANN-bagging resulted from the PCA feature selection algorithm has better results in contrast to the other FS algorithms. The values of accuracy and AUC measures are 88.4% and 90.3%, respectively. Then, the PCA is the best FS (creation) algorithm to be used in the model and classification algorithms.

As another results, Fig. 4 shows the mean of accuracy and the mean of AUC measures for four FS algorithms in all the classification algorithms. As it is clear, PCA has the best performance more than the rest of algorithms in both measures. Accuracy and AUC values are 87.8% and 86.2%, respectively. Then, by considering Figs. 3 and 4, we selected and applied this FS algorithm for our base and ensemble learning classification models.

After selecting the best FS algorithm along with its parameters setting, it was applied for implementation of classification algorithms. In continuance, the classification results with parameters setting for each classification algorithm are demonstrated. Parameters setting was employed for each classification algorithm and the best was selected for credit scoring. In Rapid Miner, excluding parameters setting, the other parameters values of the base and ensemble classifiers are default with respect to the common mode in the literature. In the CART decision tree, the 'confidence for pruning' and 'minimal size for split' were adjusted from zero to 0.5 and 4 to 40, respectively, with 4 and 3 steps. The number of folds for reduced error pruning was 3. Moreover, the measure for selecting features is 'Gini coefficient'. In the ANN, the 'learning rate' and 'epochs' (training cycles) were adjusted from 0.4 to 0.6 and 400 to 600, respectively with 2 steps. The number of hidden layers and the momentum rate were 1 and 0.2, respectively. Furthermore, it should be stated that the types of the input and output layers are linear and sigmoid, respectively. In SVM, the 'kernel types' are as follows: 1. Linear; 2. Polynomial; and 3. RBF. Additionally, the 'fixed coefficient' was adjusted from −100 to 100 with 6 steps. The remaining parameters are default. In the AdaBoost, 'iterations' were adjusted with 5, 8, and 10 values. In bagging, the 'sample ratio' was adjusted from 0.7 to 1 with 3 steps. In the random forest, the 'number of trees' was adjusted from 8 to 10 with 2 steps. Also, the 'minimal size for split' was adjusted from 4 to 15 with 2 steps. The 'confidence for pruning' was adjusted from 0.2 to 0.5 with 3 steps. The minimal gain and maximal depth were 0.1 and 20, respectively. Eventually, the criterion for selecting features was 'Gini coefficient'. In stacking, the base learners were as follows: (1) CART; (2) SVM; and (3) ANN. Moreover, the stacking model learner was Naïve Bayes. At last, cross validation in stacking was 10 folds. The best model of each classification algorithm after parameters setting is shown in Tables 5–12.

As the final result, Fig. 5 compares several base and ensemble classification algorithms by two measures of accuracy and AUC. As

Table 5
The best model of Naïve Bayes.

Accuracy (%)	AUC (%)
82.05	79.81

Table 6
The best model of CART.

Confidence for pruning	Minimal size for split	Accuracy (%)	AUC (%)
0.125	16	85.90	89.31

Table 7
The best model of ANN.

Learning rate	Epochs	Accuracy (%)	AUC (%)
0.6	400	87.18	84.54

Table 8
The best model of SVM.

Kernel type	Coefficient	Accuracy (%)	AUC (%)
RBF	0	85.90	81.85

Table 9
The best models of AdaBoost.

Row	Model	Iterations	Accuracy (%)	AUC (%)
1	ANN–AdaBoost	8	91.03	91.20
2	CART–AdaBoost	8	83.33	84.72
3	Naïve Bayes–AdaBoost	8	82.05	83.84
4	SVM–AdaBoost	8	85.90	75.28

Table 10
The best models of bagging.

Row	Model	Sample ratio	Accuracy (%)	AUC (%)
1	ANN-bagging	1	88.46	90.28
2	CART-bagging	1	84.62	89.31
3	Naïve Bayes-bagging	0.7	83.33	83.70
4	SVM-bagging	0.8	85.90	81.39

illustrated in Fig. 5, the ANN–AdaBoost classification algorithm is the best classification algorithm for credit scoring, since its accuracy and AUC values are 91.0% and 91.2%, respectively. The ANN-bagging could achieve the next ranks. The worst algorithms are Naïve Bayes and SVM–AdaBoost because of the lowest values of accuracy and AUC measures, respectively.

Figs. 6 and 7 compare and sort the results of classification accuracy and AUC measures for the classification algorithms in a visual representation, respectively.

In order to compare the classification algorithms with the simultaneous application of two measures, the mean value of accuracy and AUC measures were computed. The result is displayed in Fig. 8.

With respect to the experimental results, the following conclusions can be expressed:

1. In FS algorithms, the PCA algorithm has better performance than the other algorithms, because of the fact that the selected (constructed) features by PCA used for classification algorithms could create better results in the classification accuracy and AUC measures, as confirmed by Figs. 3 and 4.
2. As shown in Fig. 8, ensemble learning algorithms have better performance than the base learners. Moreover, it is proved in Figs. 6 and 7 for the accuracy and AUC measures, respectively. This finding supports the results of studies mentioned in the literature. But it is important that which base algorithm is better to corporate with the ensemble algorithm. For example, SVM could not have a good corporation with the ensemble learning strategy. This has been acknowledged in Figs. 6–8.
3. According to Figs. 7 and 8, ANN has better results than the other base learners, i.e., Naïve Bayes and SVM, excluding CART. But with respect to Fig. 6, ANN is the best for the classification accuracy measure than the other base learners. Moreover, the ANN–AdaBoost has the best performance than the base and ensemble learning algorithms, which has been verified in Figs. 6–8. In other words, the best base classifier to be employed in ensemble learning algorithms is ANN.

Table 11
The best model of random forest.

Confidence for pruning	Minimal size for split	Number of trees	Accuracy (%)	AUC (%)
0.2	4	10	87.18	89.12

Table 12
The best model of stacking.

Accuracy (%)	AUC (%)
85.90	82.41

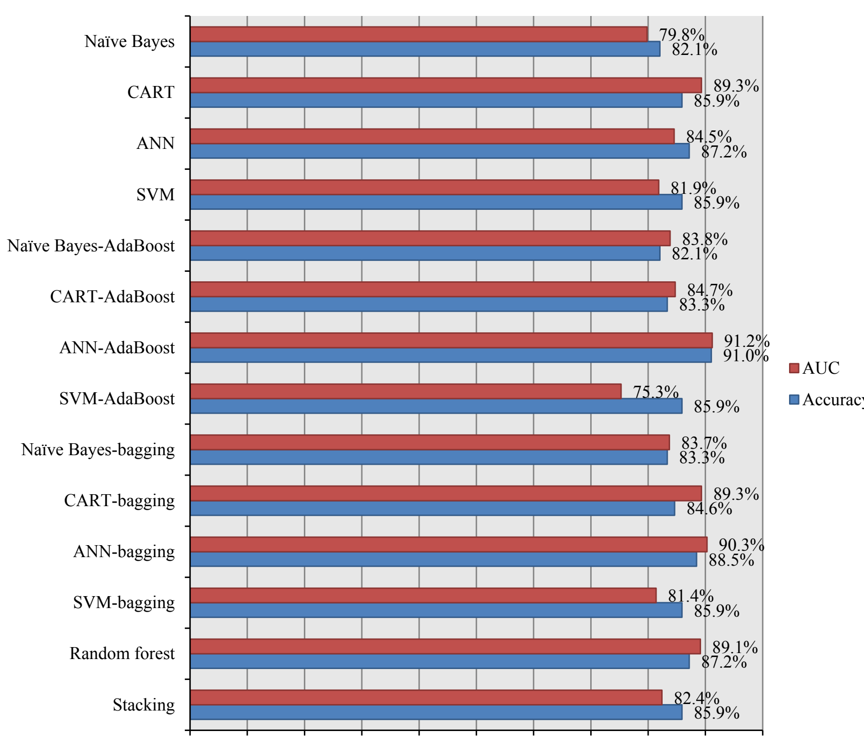

Fig. 5. The results of classification algorithms considering the accuracy and AUC measures.

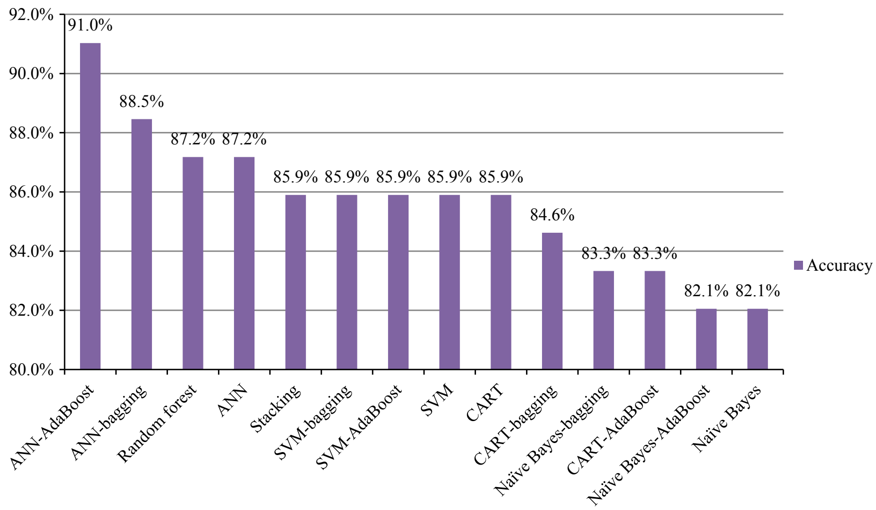

Fig. 6. The results of classification algorithms considering the accuracy measure.

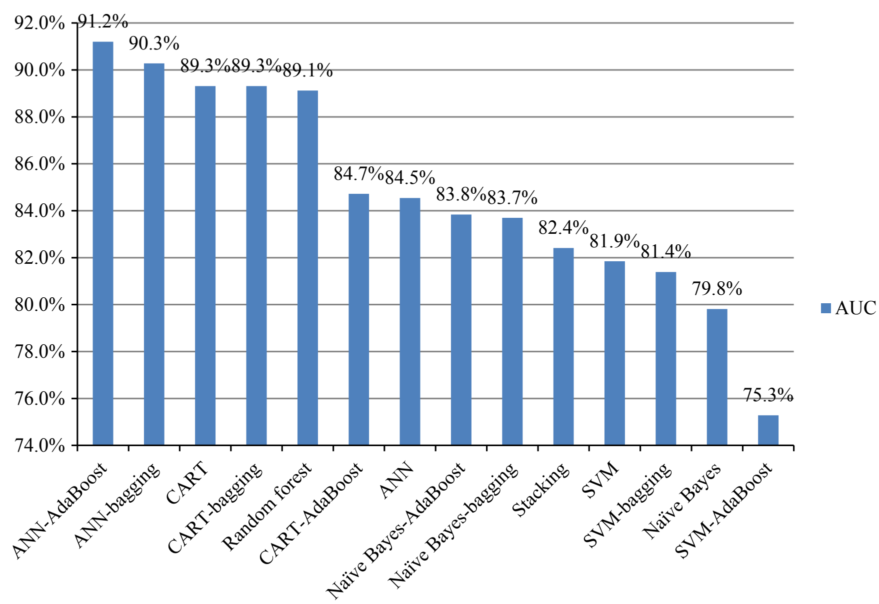

Fig. 7. The results of classification algorithms considering the AUC measure.

4. As displayed in Fig. 8, Naïve Bayes, Naïve Bayes–AdaBoost, and Naïve Bayes-bagging are the three methods with the worst classification performance and it has been shown that Naïve Bayes algorithm cannot play an appropriate role in the ensemble learning strategy.

5. Conclusions and future recommendations

In this paper, a new hybrid credit scoring model of FS algorithms and base/ ensemble learning classifiers was employed in the 'Export Development Bank of Iran' dataset. In the proposed model, four FS algorithms were firstly applied in order to attain the suitable feature subset with higher classification performance. FS algorithm that creates better SVM classification accuracy was selected to be considered in the model. In addition, parameters setting was used to achieve better results in the FS algorithms. Among the GA, information gain ratio, PCA, and relief methods of FS, PCA was selected as the best choice. After choosing the best FS algorithm, the classification algorithms with their parameters setting were implemented and the best model for each classification algorithm was selected. For weaker models, with respect to an iterative process, the best FS method was again selected and used for modeling. As the final result, Figs. 6–8 showed that the ANN–AdaBoost classification algorithm is the best classification algorithm for credit scoring. Moreover, SVM–AdaBoost, Naïve Bayes, and Naïve Bayes–AdaBoost were the worst algorithms in the evaluation measure.

The proposed hybrid model can be used for customer credit scoring in order to present financial facilities to good customers. In addition, it can solve the problems mentioned in this paper, which previous studies had not considered them. In order to solve these problems, this study proposed a comprehensive approach for

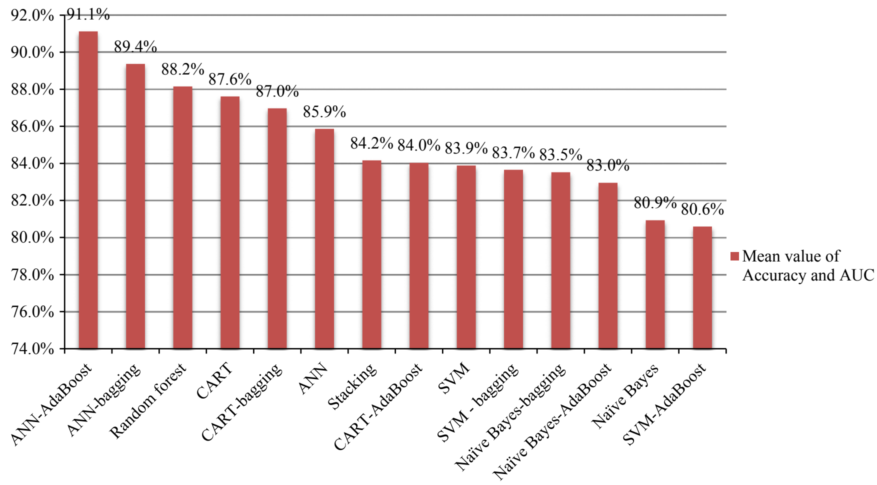

Fig. 8. The results of classification algorithms by mean value of the accuracy and AUC measures (ranking from left to right).

simultaneous application of several FS and classification algorithms with a parameters setting procedure in a hybrid data mining framework in order to increase the confidence level of the final credit scoring models.

In the other direction, this study verifies the result of previous studies that multiple and ensemble learning classification algorithms have better performance than the single ones in credit scoring applications of data mining. Furthermore, former studies have considered hybrid approaches of credit scoring in their researches. The current study presents a hybrid model for simultaneous use of FS algorithms and base and ensemble classification algorithms and compares it with the single and base learners. The results demonstrated that hybrid models are better than single ones and this is an approving outcome in the literature. From another point of view, it can be considerable that parameters setting for FS and classification algorithms should be done in order to obtain a higher performance. At last, the results showed that non-parametric (data mining) methods of credit scoring have better presentation in contrast to parametric ones, which is similar to the literature.

In future researches, the other FS algorithms can be used, such as simulated annealing (SA), PSO, ant colony, F-score, LDA and rough sets. Moreover, the other classification algorithms can be used (including base algorithms, such as LR, CHAID, C4.5, KNN, discriminant analysis, and ensembles), and the comparison can be made with the results of the model. In addition, it should be declared that the parameters setting in the paper was constrained. In future, a comprehensive parameters setting can be performed in order to search and find better results. As another recommendation, the results can be compared in (or with) the other datasets. Furthermore, the ensemble learning algorithms used in the paper can be compared with the other algorithms (e.g., random subspace, DECORATE, and rotation forest). As the last recommendation, among the other types of hybrid credit scoring models, two types can be applied: (1) Using the classification techniques in each cluster, which has been resulted from the clustering algorithm implemented in the dataset; (2) Two-stage hybrid approaches by two classification algorithms.

Acknowledgements

This work was supported by a grant from University of Tehran with number 29999/01/01.

Appendix A. Variables description

- Return on sales: This ratio shows the profit that the company has earned from each Dollar of sales.
- Total quality score (out of 130): The total scores achieved by a legal customer, in terms of the qualitative variables in the dataset, is stored in this variable.
- Claims cycle: This variable represents the ratio between credit sales of one period and the average balance of debtors. This indicator shows how many cycles the debtors have in one year. It is achieved by dividing sales by debts.
- Operation cycle: It includes the investment period in raw materials inventories, their conversion into goods inventories, changing inventories via selling to business debtors, and then changing debtors into cash money via collecting the claims, which are used to pay the obligations created from the current costs and also extension of inventories.
- Inventory cycle: The number of times the inventories renewed in a period is called inventory cycle. The index of inventory cycle is usually stated in terms of the number of times per year. However, the inventory cycle period can be presented according to the sales in a number of days. It can be attained by division of sales by inventories.
- Risk of target markets: This variable demonstrates the political, economic, commercial, and other risks of the target markets of legal customers.
- Experience with the bank: The years, in which the customer has worked with the bank, which reveals the experience and loyalty of the commercial entity.
- History of the company/Background activity: This variable indicates the years of activity of the commercial entity (customer), which is a sign of the continuity of the commercial entity's activities.

- Records of senior managers: The number of working years of senior managers of the commercial entities (customers) is recorded in this variable.
- Legal personality: Various legal characters are stored in this variable.
- Seasonal factors: This variable evaluates the extent of influence of the environmental and seasonal factors and also the corruption of produced goods on production and sales.
- Activity in the domestic market: This variables indicates the legal customers' activity in the domestic market.
- Territory of the foreign market: This variable signifies the legal customer communications with the foreign market.
- Working capital/Circulating capital: This variable indicates the degree of efficiency and adequacy, at which the special properties have been employed in the company's operations, and through comparing this ratio with the standard ratio of the respective field of industry, it will expose the intensity, at which the capital has been utilized.
- Current assets flow: This ratio is obtained by division of sales by current assets.
- Non-current assets flow: This variable represents the ratio of sales to non-current assets.
- Interest rate of fixed capital: This variable calculates the profit rate to fixed capital (or the invested amount in the company), which is representative of the attraction power of income from the investors' point of view.
- Interest rate of investments: It is the result of dividing the gross profit by the total assets of the company. Determination of the investment interest rate is useful and important from several aspects: Evaluation and control of investment plans, profit planning, determining the net profit of shareholders, determining the profit achieved from any kind of goods, determining the profits achieved from each part of the company and pricing of new products.
- Equity dividend rate: This variable represents the ratio of the remaining profit for shareholders to their rights. It is, in fact, a portion of profit that can be either re-invested in the company, or divided among shareholders.
- Current liabilities to equity ratio: This ratio shows that in case of danger, what coverage is available to satisfy the claims of current debtors, in the terms of equity. The smaller the ratio is, it would be better from the viewpoint of short-term creditors.
- Non-current liabilities to equity ratio: This ratio presents the percentage of non-current liabilities out of the shareholders' properties. The higher the ratio is, the long-term creditors will be less secured and this results in a higher risk for them, so that if the company gets bankrupt, the company maybe even unable to achieve its original sources.
- Percentage of covering the financing costs: This ratio exhibits the extent, at which the profit, prior to payment of financing expenses, can cover the financing costs. If this ratio is low, it will be a warning for the creditors that by reducing the interest, the payment of financing costs will subject to the risk of non-payment.
- Current ratio: It is obtained by dividing current assets by current liabilities. This ratio is the most commonly used means to measure the payment power for short-term liabilities, because through this ratio, it can be realized that how many times larger are the asserts, which are converted into cash during the financial year, than the liabilities that will mature during the financial year.
- Non-current assets to equity ratio: This ratio shows that how much of non-current assets are financed from equity. The higher the ratio is, the financial strength of the company will be greater.
- Quick ratio: It is measured by subtracting inventories from current assets and then dividing the remaining part by current liabilities. In fact, it is the result of division of quick assets by current liabilities. Inventories are generally less likely to be liquidated, compared to the other items of current assets, and this fact causes more loss during the settlement. According to this explanation, it is important to measure the payment power for short-term liabilities without relying on the sales of inventories.
- Equity ratio: This ratio specifies the portion of equity out of the company's total assets. In other words, it represents that how many percent of the company's assets can be attributed to the shareholders rights (including capital, savings, and accumulated profit). Subtracting this number from one will disclose the share of liabilities in the company's total assets. If the equity ratio is negative, it means that due to the loss of the company during the activity period, the accumulated loss of the company is higher than the total capital and savings (i.e., liabilities have supplied assets and compensated a part of the company's losses).
- Debt ratio: The less this ratio is, the more willing the financing firms will be to grant long-term facilities.
- The credit expert 1's verdict: This variable reflects the credit expert's opinion in the related bank.
- The credit expert 2's verdict: This variable reflects the credit expert's opinion in the related bank.
- Fulfillment of obligations (Target variable): This variable expresses whether if the legal customer has been able to fulfill his/her commitments in due time, or not. According to the opinions of the bank experts, this variable is ranked into the ranks of zero (able to re-pay up to maximum two months) and one (unable to re-pay).

References

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.
Brown, I., Mues, Ch., 2012. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Syst. Appl. 39 (3), 3446–3453.
Chen, F.L., Li, F.Ch., 2010. Combination of feature selection approaches with SVM in credit scoring. Expert Syst. Appl. 37, 4902–4909.
Chen, Y. Sh, 2012. Classifying credit ratings for Asian banks using integrating feature selection and the CPDA-based rough sets approach. Knowl.-Based Syst. 26, 259–270.
Danenas, P., Garsva, G., 2015. Selection of support vector machines based classifiers for credit risk domain. Expert Syst. Appl. 42, 3194–3204.
Freund, Y., Schapire, R.E., 1996. Experiments with a new boosting algorithm. In: Proceedings of the 13th International Conference on Machine Learning, Bari, Italy, pp. 148–156.
Gray, J.B., Fan, G., 2008. Classification tree analysis using TARGET. Comput. Stat. Data Anal. 52 (3), 1362–1372.
Hajek, P., Michalak, K., 2013. Feature selection in corporate credit rating prediction. Knowl.-Based Syst. 51, 72–84.
Hsieh, N.C., Hung, L.P., 2010. A data driven ensemble classifier for credit scoring analysis. Expert Syst. Appl. 37 (1), 534–545.
Hu, Y.C., Ansell, J., 2007. Measuring retail company performance using credit scoring techniques. Eur. J. Oper. Res. 183 (3), 1595–1606.
Kao, L.J., Chiu, C.C., Chiu, F.Y., 2012. A bayesian latent variable model with classification and regression tree approach for behavior and credit scoring. Knowl.-Based Syst. 36, 245–252.
Larose, D.T., 2005. Discovering Knowledge in Data: An Introduction to Data Mining. WILEY, New Jersey.
Liang, D., Tsai, Ch. F., Wu, H.T., 2015. The effect of feature selection on financial distress prediction. Knowl.-Based Syst. 73, 289–297.
Marques, A.I., Garcia, V., Sanchez, J.S., 2012. Exploring the behaviour of base classifiers in credit scoring ensembles. Expert Syst. Appl. 39 (11), 10244–10250.
Min, J.H., Lee, Y.C., 2008. A practical approach to credit scoring. Expert Syst. Appl. 35 (4), 1762–1770.
Nanni, L., Lumini, A., 2009. An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. Expert Syst. Appl. 36 (2), 3028–3033.
Ong, C.S., Huang, J.J., Tzeng, G.H., 2005. Building credit scoring models using genetic

programming. Expert Syst. Appl. 29 (1), 41–47.
Oreski, S., Oreski, G., 2014. Genetic algorithm-based heuristic for feature selection in credit risk assessment. Expert Syst. Appl. 41, 2052–2064.
Paleologo, G., Elisseeff, A., Antonini, G., 2010. Subagging for credit scoring models. Eur. J. Oper. Res. 201 (2), 490–499.
Ping, Y., Yongheng, L., 2011. Neighborhood rough set and SVM based hybrid credit scoring classifier. Expert Syst. Appl. 38 (9), 11300–11304.
RapidMiner version 4.1 beta 2, 2001–2007. Available at: 〈http://www.rapidminer.com〉.
Salappa, A., Doumpos, M., Zopounidis, C., 2007. Feature selection algorithms in classification problems:an experimental evaluation. Optim. Methods Softw. 22 (1), 199–212.
Setiono, R., Baesens, B., Mues, C., 2009. A note on knowledge discovery using neural networks and its application to credit card screening. Eur. J. Oper. Res. 192 (1), 326–332.
Sustersic, M., Mramor, D., Zupan, J., 2009. Consumer credit scoring models with limited data. Expert Syst. Appl. 36 (3), 4736–4744.
Thomas, L.C., 2000. A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. Int. J. Forecast. 16 (2), 149–172.
Tsai, C.F., 2009. Feature selection in bankruptcy prediction. Knowl.-Based Syst. 22, 120–127.
Tsai, C.F., Wu, J.W., 2008. Using neural network ensembles for bankruptcy prediction and credit scoring. Expert Syst. Appl. 34 (4), 2639–2649.
Twala, B., 2010. Multiple classifier application to credit risk assessment. Expert Syst. Appl. 37 (4), 3326–3336.
Vukovic, S., Delibasic, B., Uzelac, A., Suknovic, M., 2012. A case-based reasoning model that uses preference theory functions for credit scoring. Expert Syst. Appl. 39 (9), 8389–8395.
Wang, C.M., Huang, Y.F., 2009. Evolutionary-based feature selection approaches with new criteria for data mining:A case study of credit approval data. Expert Syst. Appl. 36 (3), 5900–5908.
Wang, G., Ma, J., 2012. A hybrid ensemble approach for enterprise credit risk assessment based on support vector machine. Expert Syst. Appl. 39, 5325–5331.
Wang, G., Hao, J., Ma, J., Jiang, H., 2011. A comparative assessment of ensemble learning for credit scoring. Expert Syst. Appl. 38 (1), 223–230.
Wang, J., Hedar, A.R., Wang, S., Ma, J., 2012. Rough set and scatter search metaheuristic based feature selection for credit scoring. Expert Syst. Appl. 39 (6), 6123–6128.
Wang, Y.Y., Li, J., 2008. Feature-selection ability of the decision-tree algorithm and the impact of feature-selection/extraction on decision-tree results based on hyperspectral data. Int. J. Remote Sens. 22 (10), 2993–3010.
Witten, I.H., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers, Boston.
Xu, X., Zhou, C., Wang, Z., 2009. Credit scoring algorithm based on link analysis ranking with support vector machine. Expert Syst. Appl. 36 (2), 2625–2632.
Yeh, I.C., Lien, C. h, 2009. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Expert Syst. Appl. 36 (2), 2473–2480.
Yu, L., Wang, S., Lai, K.K., 2009. An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: the case of credit scoring. Eur. J. Oper. Res. 195 (3), 942–959.
Zhou, L., Lai, K.K., Yu, L., 2009. Credit scoring using support vector machines with direct search for parameters selection. Soft Comput. 13, 149–155.