Adv. Math. Fin. App., 2024, 9(3), P.873-899

Advances in Mathematical Finance & Applications

www.amfa.iau-arak.ac.ir

Print ISSN: 2538-5569

Online ISSN: 2645-4610

Original Research

Early Warning Model for Solvency of Insurance Companies Using Machine Learning: Case Study of Iranian Insurance Companies

Saeed Naseri Khezerlou, Atousa Goodarzi*

Department of Insurance, ECO College of Insurance, Allameh Tabatabai University, Tehran, Iran

ARTICLE INFO

Article history:
Received 2023-04-25
Accepted 2024-02-23

Keywords:
Insurance
Solvency
Early Warning Model
Machine Learning
Financial Ratio Analysis

ABSTRACT

Stakeholders of an organization avoid undesirable outcomes caused by ignoring the risks. Various models and tools can be used to predict future outcomes, aiming to avoid the undesirable ones. Early warning models are one of the approaches that could help them in doing so. This study focuses on developing an early warning system using machine learning algorithms for predicting solvency in the insurance industry. This study analyses 23 financial ratios from Iranian general insurance companies listed on the Tehran Stock Exchange between 2015 and 2020. The model uses Decision Tree, Random Forest, Artificial Neural Networks, Gradient Boosting Machine and XGBoost algorithms, with Boruta as a feature selection method. The dependent variable is the solvency margin ratio, and the other 22 ratios are the independent variables, which Boruta reduces to 7 variables. Firstly, the performance of the machine learning models on two datasets, one with 22 independent variables and one with 7, is compared based on RMSE values. The XGBoost algorithm performs the best on both data sets. Additionally, the study predicts the 2020 values for 19 insurance companies, performs stage classifications, and compares actual stages to predicted stages. In this analysis, Random Forest has the best estimate accuracy on both data sets, while Gradient Boosting Machine has the best estimate accuracy on the Boruta data set. Finally, the study compares the machine learning models' results in terms of capital adequacy classification, where Random Forest performs the best on both data sets, and Gradient Boosting Machine on the Boruta data set.

1 Introduction

Insurance, as one of the components of financial sector, plays an important role in economic growth and development of countries and a strong insurance industry is a guarantee for the strength of the financial system. Keeping this in mind, we can conclude that failure of insurance companies will result in undesirable outcomes for the economy of a country. Hence, avoiding such a failure is one of the most important concerns of government as well as insurance companies' stakeholders, whose well-being is somehow dependent on intact operation of these companies. One of the most popular factors that measure financial strength of insurance companies and ensures their intact operation is "solvency". It is the ability of a company to meet its long-term debts and financial obligations [1]. By continuous measuring of solvency of insurance companies, regulators and executives of companies can ensure financial health of them and take timely interventions at the case of lower solvency levels (insolvency). Therefore, this

* *Corresponding author. Tel.: 09124949526*
E-mail address: atousagoodarzi@atu.ac.ir

can avoid financial failure of insurance companies and ensures a strong insurance sector, which will lead to economic growth. Different approaches are used for life and non-life insurance companies in the calculation of capital adequacy, i.e., solvency. Solvency II is one the most famous frameworks which is for insurance companies in the European Union that has come into effect on 1 January 2016. This directive consists of 3 Pillars that includes Pillar 1 that is quantitative requirements, Pillar 2 that is requirements for risk management and Pillar 3 that is transparency. The framework used in Iran for solvency of insurance companies is regulation No. 69 of High Council of the Insurance Industry of Iran which is a risk-based system. It identifies the risks facing insurance companies in 4 groups: total underwriting risk, total market risk, total credit risk, and total liquidity risk. These risks can be combined with each other and determine the total risk of an insurance company. Then, for calculating the existing capital of insurance companies, the terms of eligible assets are also determined and by dividing the existing capital by Risk Based Capital (RBC), the solvency margin of insurance companies is calculated. Based on the resulting value of the solvency margin ratio, Iranian insurance companies are classified on five solvency stages, which is an implication for required measures to be taken. All of the approaches for calculation of solvency are used with the aim of continuous monitoring of insurance companies' financial strength. Prediction is one of the best approaches that can assist regulators and insurance companies in knowing insolvencies before they happen and play an important role in having a healthy insurance industry. Many approaches for predicting company failure have been presented. However, because of the unique peculiarities of the insurance business, most of them are not feasible, and only a few have been adopted in this sector. The majority of methods used in insurance company failure prediction are statistical methods such as Discriminant Analysis or Logistic Regression, which use financial ratios as explicative factors. In most circumstances, this type of variable does not meet statistical assumptions. To avoid these issues, a variety of non-parametric algorithms, most of which are part of Machine Learning, such as neural networks, have been developed and effectively applied to these types of situations [2]."Early warning models", which are based on prediction, use statistical and mathematical models to predict the future and alert about possible anomalies such as financial crises. Early warning models form early warning systems that can be useful in continuous monitoring of financial systems and avoid losses by reporting undesirable outcomes based on its predictions. Warning systems were first raised after the currency crisis of European countries in 1992-1993, the crisis of Latin American countries in 1994-1995, and more seriously after the crisis of East Asian countries in 1997-1998. Most of the researches that have been done on the design of the warning system are usually in the field of currency crises and there are few researches in the field of financial crisis. However, after the recent global financial crisis of 2008, researchers and policy makers in all economic sectors around the world, including insurance industry, have devoted all their efforts and attention on diagnosing, understanding and predicting systematic crises. In doing this, literature related to crisis forecasting focused on the expansion and development of early warning systems, systems that somehow seek to predict future financial crises [3].The majority of previous research on the application of early warning systems in financial sector has focused on establishing these systems based on econometric methods, such as the panel-logit model. However, in response to the advancement of data mining, early warning system research is rapidly shifting away from the econometric discipline and towards data-oriented rather than causality-oriented research. Machine learning is one of the most effective data mining methods that can be used in this field, offering the capability to make generalizable predictions on previously unseen data based on given data. Machine learning offers significant advantages over econometric methods. It excels in its flexibility to handle non-stationarity and outliers, thanks to its adaptive nature. Additionally, machine learning algorithms employ advanced estimation techniques like non-linear kernel methods and artificial neural networks, enabling stronger fitting capabilities compared to traditional econometric approaches. Moreover, machine learning models prioritize generalization ability, making them more adept at predicting outcomes on unseen data. These advantages empower machine learning to provide more accurate and robust results, making it a powerful tool for formulating an early warning model [4].

On the other hand, the predominant emphasis in previous research pertaining to the utilization of machine learning for business failure prediction has revolved around comparing a specific method with traditional statistical approaches. Comparisons between two or more machine learning techniques have been relatively limited within this domain [2]. The main aim of this research is to develop an early warning model utilizing machine learning techniques. The model will leverage the ratios derived from the financial statements of Iranian insurance companies to predict their future financial status, specifically focusing on solvency. By employing this early warning model, it becomes possible to generate advance warnings concerning potential instances of insolvency that may emerge or evolve in the future. This, in turn, facilitates timely and accurate interventions by stakeholders and regulators. Furthermore, an additional objective of this study is to ascertain the ratios that demonstrate enhanced effectiveness in discerning early warning indicators.

2 Literature Review

In the review of the conducted research, it was observed that only a few studies have investigated the solvency risk of insurance companies. In this regard, Caporale et al. [5] have stated that the reason for this is that the insurance industry is less exposed to financial market turbulence than other industries, such as banking. There are several possible reasons for this difference. Unlike banks, insurers do not take deposits from customers and therefore are not exposed to the risk of unexpected liquidity shortages that can overwhelm banks. Therefore, given the fact that insurance companies have become riskier, it is necessary to examine the risk factors of insolvency. Furthermore, due to the close relationship between insurance companies and banking crises, insurers' inability can have a significant impact on the stability of financial markets. In this section, we will briefly review the most relevant studies about prediction of financial institutions' failure and utilization of early warning systems and machine learning in this field. Financial ratios were first used in the literature by Beaver [6] to estimate the failures of companies. In this study, 79 failed and 79 non-failed companies were examined. Data were obtained from Moody’s Industrial Manual and cover the period between the years 1954 and 1964. In this study, he used 30 financial ratios which were collected under six groups. These groups are cash flow ratios, net income ratios, debt to total asset ratios, liquid asset to total asset ratios, liquid asset to current debt ratios and turnover ratios. He divided the data into three sections while analyzing. He described the comparison of mean values, which was the first section, as profile analysis. Beaver, who indicates that profile analysis shows the difference between failed and non-failed companies, stated that the lack of this analysis not being able to respond to the magnitude of the difference. The profile analysis showed that the average asset size of non-failed companies is greater than that of failed companies. In the second part, the dichotomous test, a predictive test unlike the profile analysis, is mentioned. This test estimates the failure status of companies. As a result of the study, he classified the bankrupted companies with 78% accuracy, 5 years before their bankruptcy .In the study conducted by Brockett et al. [7], the artificial intelligence neural network model is used as an early warning model to estimate the insolvency status of insurance companies. In order to measure the susceptibility of an insurance company to bankruptcy, the data from two years ago of the companies that had failed in 1991 and 1992 and the existing data of the existing firms were taken into consideration. Early Warning Model is established with regulatory annual statements. The model is first established with 24 variables and then reduced to 8 variables by stepwise Logistic Regression. The results of artificial neural networks are compared with discriminant analysis, the National Association of Insurance Commissioners' Insurance Regulatory Information System ratings and A. M. Best ratings. The findings of the neural network show elevated predictability and generalizability, indicating that this technique is useful in anticipating potential insurance insolvency. The total percentage correctly classified of the Neural Network is 89.3%. Segovia-Vargas et al. [8] propose an approach to predict insolvency of Spanish non-life insurance companies. The approach consists of a Support Vector Machine (SVM) which classifies a firm as healthy or failed, depending on the value of a set of financial ratios which characterize every firm. The SVM is hybridized with a Genetic Algorithm (GA) and a Simulated Annealing (SA) in order to perform on-line feature selection in the space of financial ratios. Their resulting approach allows a very accurate classification of firms into healthy or failed, based on very few financial ratios. They have tested their approach in a real problem of prediction of insolvency of Spanish non-life insurance companies, formed by 72 firms, described by 21 financial ratios. Their approach, using GA and SA search algorithms, has achieved very good results, obtaining the lowest probability of error with only 3 ratios.In a study conducted in Australia on insurance companies whose financial situation is deteriorating, an early warning model has been developed to analyze to the situation of companies by Australian Prudential Regulation Authority. The data is between 1999 and 2001. The model is tried to be estimated by logistic regression. As a result, it is seen that insurers living financially distressed are small in size, have low profitability and have low cession rates. Moreover, these companies have reinsurance assets and properties rather than liquid assets [9].

Another study to predict the insolvency of insurance companies belongs to Rustam and Yaurita [10]. The aim of this study which was conducted in 2018 is to propose an approach to avoid insolvency in insurance companies. In this study, they have predicted the insolvency using two different methods: Support Vector Machines (SVM) and Fuzzy Kernel C-Means (FKCM). Based on their analysis of the experimental results, they have concluded that the use of discrete type of input have a significant effect on both SVM and FKCM method. The highest average accuracy of 71.93% is obtained by SVM with discrete input data types using feature selection.Wang et al. [4] address the issue of constructing effective early warning systems (EWSs) to predict and prevent systemic banking crises. They argue that the conventional EWSs based on panel logit models may face limitations in accurately extracting information during periods of crisis due to changes in economic indicators. To overcome these limitations, the authors propose an alternative framework called the "experts voting EWS," which harnesses the properties of machine learning algorithms. The authors find that among the various machine learning classifiers tested, the random forest classifier, which emulates the experts voting process, exhibits the highest efficiency, with a generalization rate exceeding 80% in terms of the area under the receiver operating characteristic curve. This indicates that the experts voting EWS, which synthesizes multivariate information, shows promise in providing alerts for systemic banking crises in diverse contexts, presenting a departure from the conventional system.In the study of Hagh Verdilou et al. [3] designing of an early warning system for solvency of insurance companies in the insurance market of Iran has been discussed. The empirical model of this study was estimated by econometric methods from panel data of 18 Iranian companies from 1387 to 1396. The findings show that interest rate and board changes with a one-period lag have the largest and smallest impacts on the solvency of Iranian insurers. Due to the strength of the dice, the loss ratio impact also varies with size. In addition, all hypotheses based on the significant impact of variables on the solvency of Iranian insurance companies are tested. These include macroeconomics (inflation rate with one delay), interest rates (with one delay), economic growth (with one delay), corporate variables (ratio of investment in risky assets to total assets), loss ratio, Herfindahl-Hirschman index and corporate governance (changes in major shareholder ownership and board of directors) and international economic sanctions.

3 Theoretical Background

3.1 Early Warning System

Early warning systems refer to a structure that, taking into account various economic, financial and managerial components, identifies and monitors the smallest changes that may lead to a crisis in an insurance institution in the future. The basis of these systems is based on the estimation of the probability of the reduction of the ability to fulfil the obligations and risks accepted by the insurance company. In other words, the early warning system is an evaluation mechanism to monitor the stability and financial health of insurance companies before it is too late to take appropriate measures. The goals of setting up an early warning system can be listed as identifying the crisis before the incident, ensuring the insurer's ability to pay obligations, informing about potential problems and providing the right solution at the right time. The early warning system is actually a system that provides information about some indicators to show the potential adverse performance of the insurance company in advance so that the necessary action can be taken before the occurrence of an adverse situation [3].

3.2 Machine Learning Algorithms

Machine learning algorithms are organized into taxonomies based on the desired outcome of the algorithms. Common algorithm types include supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, transduction and learning to learn [11]. In this study we have used Decision Tree, Random Forest, Artificial Neural Networks, Gradient Boosting Machine and Extreme Gradient Boosting algorithms to predict solvency stage of insurance companies in Iran and Boruta for feature selection. All these algorithms are types of supervised learning algorithms. In supervised learning, algorithms generate functions that map inputs to desired outputs. A standard formulation of supervised learning tasks is the classification problem: the learner should learn a function that maps a vector to one of several classes (to approximate its behaviour) by looking at some example inputs and outputs of the function. Features, operation and background of the utilized algorithms are described in the following.

3.2.1 Decision Tree

The decision tree algorithm is an information classification algorithm. The decision tree has many advantages as an inductive induction algorithm, including the ability to independently select feature variables, fast classification speed, and the ability to effectively filter information. The decision tree algorithm is therefore regarded as one of the statistically optimal algorithms [12]. Fig. 1 depicts a simple decision tree model with a single binary target variable Y (0 or 1) and two continuous variables $X1$ and $X2$, both of which range from 0 to 1. A decision tree model's main components are nodes and branches, and the most important steps in building a model are splitting, stopping, and pruning [13].

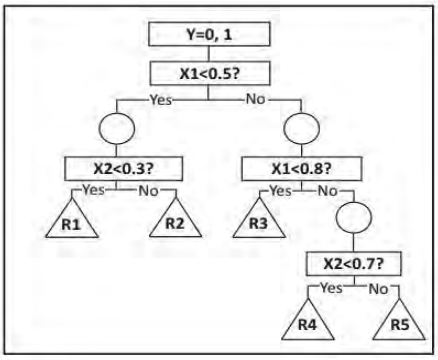

Fig. 1: Sample Decision Tree Based on Binary Target Variable Y [13]

Nodes: Nodes are classified into three types. (a) A root node, also known as a decision node, denotes a decision that will result in the division of all records into two or more mutually exclusive subsets. (b) Internal nodes, also known as chance nodes, represent one of the options available at that point in the tree structure; the node's top edge is connected to its parent node, and the node's bottom edge is connected to its child nodes or leaf nodes. (c) Leaf nodes, also known as end nodes, represent the outcome of a series of decisions or events [12].

Branches: Branches are random outcomes or occurrences that arise from root nodes and internal nodes. A decision tree model is built using a branch hierarchy. Each path from the root node to the leaf node represents a classification decision rule. These decision tree paths are also known as 'if-then' rules. "If condition 1 and condition 2 and condition... and condition k occur, then outcome j occurs," for example [13].The decision tree algorithm consists of two steps. The initial step is to create a decision tree. The second step is to create decision trees. In general, the overall construction process of the decision tree is a continuous judging and classification of information. The feature variable with the greatest difference is left after each construction based on the decision tree algorithm's characteristics. Different decision trees' differences are measured differently. Construction is also a form of pruning. The goal is to improve the decision tree's fit to the data [12].

3.2.2 Random Forest

Random Forest is a decision tree algorithm with a relatively small number of parameters that works well for large data sets. There are black boxes for artificial neural networks and flow charts for decision trees separated from linear model formulas. After building a number of decision trees, the most popular classes are selected. These methods are called random forests.A random forest regression method is generated by evolving the tree according to random vectors such that the numbers are taken from the prediction tree $h(x, \Theta)$ instead of the class markers. The output values are quantitative and the training set is intended to be separated from the random vector distribution Y, X. The mean squared generalization error for each quantitative predictor is $h(x)$ [14].

$$h(x) = E_{X,Y}\big(Y - h(X)\big)^2 \tag{1}$$

The random forest predictor consists of getting the mean over k of all trees $h(x, \Theta_k)$.

When the number of trees in a forest becomes infinite, we can say that this is a forest generalization error [14]:

$$E_{X,Y}\big(Y - av_k h(X, \Theta_k)\big)^2 \to PE^*(forest) = E_{X,Y}\big(Y - E_\Theta h(X, \Theta)\big)^2 \tag{2}$$

If we define the average generalization error of only one tree as $PE^*(tree)$

$$PE^*(tree) = E_\Theta E_{X,Y}\big(Y - h(X,\Theta)\big)^2 \tag{3}$$

As described by [14], assume that $\forall\Theta, EY = E_X h(X,\Theta)$. Then,

$$PE^*(forest) \le \bar{\rho} PE^*(tree) \tag{4}$$

where $\bar{\rho}$ is the weighted correlation of residues between $Y - h(X,\Theta)$ and $Y - h(X,\Theta')$ that Θ and Θ' are independent. $PE^*(forest)$ equals to

$$E_{X,Y}\big[E_\Theta\big(Y - h(X,\Theta)\big)\big]^2 = E_\Theta E_{\Theta'} E_{X,Y}\big(Y - h(X,\Theta)\big)\big(Y - h(X,\Theta')\big) \tag{5}$$

and, covariance is $E_\Theta E_{\Theta'}\big(\rho(\Theta,\Theta')\big)\, sd(\Theta)\, sd(\Theta'), sd(\Theta) = \sqrt{E_{X,Y}\big(Y - h(X,\Theta)\big)^2}$ and weighted correlation can be written as:

$$\bar{\rho} = \frac{E_\Theta E_{\Theta'} \rho(\Theta,\Theta')\, sd(\Theta)\, sd(\Theta')}{\big(E_\Theta sd(\Theta)\big)^2} \tag{6}$$

3.2.3 Artificial Neural Network (ANN)

Neural networks are a well-established idea in machine learning, attempting to identify hidden patterns between input and output by constructing a structure resembling a brain. Each neural network is made up of neurons, which are the smallest processing units. Each neuron receives certain inputs and, after initial processing, generates outputs, each of which can be the output of another layer or the input of another layer of neurons [15]. Each individual neuron calculates the sum of incoming signals, incorporates a bias term, and applies a non-linear function for transformation. The activation function, often chosen as a logistic function, hyperbolic tangent, or ReLu type, is a monotonically increasing function. Using a linear function would not adequately capture the non-linear characteristics of the output data. The transformed signal from one neuron is then transmitted to another neuron in a different layer, and this process is repeated. The output values of the neurons are compared to the target values of the data, and an error or cost function is computed. This error is propagated back through the network, leading to adjustments in the weights in order to minimize the cost function. This iterative process continues until the network accurately estimates the output data within an acceptable range of accuracy. Fig. 2 shows a Neural Network with only one hidden layer [16].

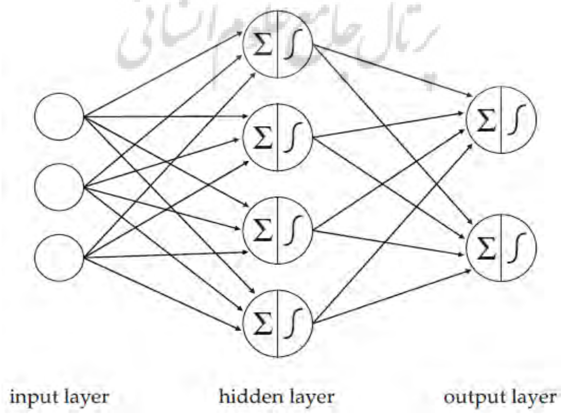

Fig. 2: Illustration of A Multilayer ANN [17]

Let $\boldsymbol{y_{u,k}}$ be the signal received by neuron $\boldsymbol{l}$ in layer $\boldsymbol{v}$ from neuron $\boldsymbol{k}$ in the previous layer $\boldsymbol{u}$, $\boldsymbol{w_{l,k}}$ be the weight $\boldsymbol{l}$ applies to $\boldsymbol{y_{u,k}}$, and $\boldsymbol{b_{v,l}}$ be the bias term to calculate a weighted sum $\boldsymbol{z_{v,l}}$. The signal $\boldsymbol{y_{v,l}}$ in $\boldsymbol{l}$ is generated by applying an activation function σ to $\boldsymbol{z_{v,l}}$:

$$\boldsymbol{z_{v,l}} = \left(\sum_k \boldsymbol{w_{l,k} y_{u,k}}\right) + \boldsymbol{b_{v,l}};\ \boldsymbol{y_{v,l}} = \boldsymbol{\sigma}\left(\boldsymbol{z_{v,l}}\right) \tag{7}$$

For classification problems, the number of neurons in the output layer equals the cardinality of the target attribute's domain. During training, the target function $\boldsymbol{E}$ measures the (quadratic) error between the output signals $\boldsymbol{y_{Output,i}}$ of the output layer and the actual target value $\boldsymbol{y_i}$ for each record:

$$\boldsymbol{E} = \sum_i \frac{1}{2}\left(\boldsymbol{y_{Output,i}} - \boldsymbol{y_i}\right)^2;\ \boldsymbol{y_i} = \boldsymbol{1} \text{ for target domain } \boldsymbol{i}, \text{ otherwise } \boldsymbol{y_i} = \boldsymbol{0}. \tag{8}$$

Since each layer's signal is a function of the previous layer's weights, biases, and signals, $\boldsymbol{E}$ is finally a function of the (averaged) weights and biases from all training records and all layers of the neural network. The gradient of $\boldsymbol{E}$ indicates the sensitivity of the objective function to changes in these parameters:

$$\boldsymbol{\nabla E} = \begin{bmatrix} \vdots \\ \frac{\partial \boldsymbol{E}}{\partial \boldsymbol{w_{l,k}}} \\ \frac{\partial \boldsymbol{E}}{\partial \boldsymbol{b_{v,l}}} \\ \vdots \end{bmatrix}. \tag{9}$$

The larger the partial derivative of $\boldsymbol{E}$, the more the objective function benefits from its manipulation and descends to a minimum. Therefore, at each step of the training, the weight and bias are adjusted simultaneously in proportion to the negative partial derivative. This process is repeated until the cost function improvement falls below a predefined threshold. When using a trained neural network for prediction, the learned rules are applied to new data and the resulting output values are used as prediction values [17].

3.2.4 Gradient Boosting Machine

Another decision tree algorithm, like random forest, is the gradient boosting machine. This is an ensemble technique for regression and classification tree models. Gradient boosting machines are machine learning methods that include two powerful tools: gradient-based optimization and boosting. Gradient-based optimization uses gradient computation to minimize the model's loss function with respect to the training data. Additive boosting collects an ensemble of weak systems to produce a robust learning scheme for prediction tasks.It has more parameters than Random Forest and requires a little more effort to tune, but gives slightly stronger results. It has been observed to give better results, especially in regression studies. The main risk is that it can easily overfit as the number of trees increases [18].

GBM setting for the regression, the connection between $L\left(y, f(x)\right) = \left(y - f(x)\right)^2$ squared-error loss and $L\left(y, f(x)\right) = |y - f(x)|$ absolute loss is similar to the connection between exponential loss and binomial log-likelihood where y is actual data and x is predicted values. The population solutions are for both $f(x) = E(Y|x)$ for squared-error loss, and $f(x) = median(Y|x)$ for absolute loss. They are therefore the same as for the symmetric error distribution. One of the loss functions is the Huber loss for regression and calculation is [19]:

$$L\left(y, f(x)\right) = \begin{cases} \left(y - f(x)\right)^2 & for\ \ |y - f(x)| \le \delta \\ 2\delta|y - f(x)| - \delta^2 & otherwise. \end{cases} \tag{10}$$

3.2.5 Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) includes an efficient linear model solver and decision tree learning algorithm, supporting various objective functions such as regression, ranking, and classification. It brings parallel tree boosting to solve many data science problems quickly and accurately. It is one of the best gradient extension frameworks today for many problems .The XGBoost algorithm defines a two-component objective function instead of optimizing a simple second-order error loss. A loss function is defined on the training data, and a regularization term is defined to penalize the complexity of the model:

$$\mathcal{L}(\phi) = \sum_i L(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \tag{11}$$

$L(y_i, \hat{y}_i)$ can be any differential convex loss function that estimates predictive distinction using the real labels of the given training data.

$$\Omega(f_k) = \gamma T + 0.5\lambda\omega^2 \tag{12}$$

where $\Omega(f_k)$ describes the f_k tree complexity. The number of f_k tree leaves is T and the leaf weight is ω. If $\Omega(f_k)$ is in the objective function, we need to optimize a small tree that minimizes $L(y_i, \hat{y}_i)$ at the same time. This helps reduce overfitting. γT Normally provides a constant charge for each additional leaf, and $\lambda\omega^2$ penalizes extreme weights. The user configuration parameters are γ and λ.Boosting proceeds is repeated and newest objective function is:

$$\mathcal{L}^t = \sum_i L\left(y_i, \hat{y}_i^{(t-1)} + f_k(x_i)\right) + \sum_k \Omega(f_k) \tag{13}$$

and you can find the minimal objective function [20].

3.3 Boruta

Boruta is a feature selection method. Using this method, you can determine which variables in your dataset are important and which are not. Boruta's algorithm is intended as a wrapper around the random forest classification algorithm. In Slavic mythology, Boruta is known as the forest god. Iteratively remove features that statistical tests show to be less important than the sample. This algorithm uses a wrapper approach built around a random forest classifier. This algorithm determines actual property relevance by comparing the relevance of random probes.

Boruta feature selection algorithm follows these steps: Make copies of all independent variables. All copy variables are merged with the original data, but their values are merged to remove their relationship to the target variable. This is called a permuted copy or shadow feature. The random forest classifier is run on the combined data and performs variable importance measures to rank the importance of each variable. The higher the meaning, the more important it is. Then the average precision loss divided by the standard deviation of accuracy loss (Z_{score}) is calculated. Maximum Z_{score} is found among shadow attributes (MZSA). If the variable is too low below the MZSA, the variable is marked as unimportant. And this variable is permanently removed from the process. If a variable is higher than MZSA, it is marked as important. The random forest is updated to a predefined number so that all variables are marked as either important or unimportant [21].

3.4 Loss Functions

One of the necessary tools in machine learning problems is the loss function. Machine learning algorithms are based on the loss or performance function, and the parameters are adjusted to minimize the loss function or maximize the performance function [22]. The principle of the system focuses on whether regression or classification is performed.

$$\boldsymbol{Loss\ Function} = \boldsymbol{L}(\boldsymbol{W}, \boldsymbol{B}|\boldsymbol{j}) \tag{14}$$

where $\boldsymbol{i} \in [\boldsymbol{1}, \boldsymbol{N}-\boldsymbol{1}]$, $\boldsymbol{W_i} \subset \boldsymbol{W}$ denotes the weight matrix connecting layers $\boldsymbol{i}$ and $\boldsymbol{i}+\boldsymbol{1}$ for a network of the number of N layers. Likewise, $\boldsymbol{i} \in [\boldsymbol{1}, \boldsymbol{N}-\boldsymbol{1}]$, $\boldsymbol{b_i} \subset \boldsymbol{B}$ indicates the column vector of biases for layer $\boldsymbol{i}+\boldsymbol{1}$ [23].

3.4.1 Mean Squared Error (MSE)

The Mean Squared Error measures the performance of the predictor for regression. $\boldsymbol{y_j}$ is the observed output data of $\boldsymbol{j_{th}}$ and $\boldsymbol{\hat{y}_j}$ represent the predicted output value [14].

$$\boldsymbol{L}(\boldsymbol{W}, \boldsymbol{B}|\boldsymbol{j})_{\boldsymbol{MSE}} = \frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \left(\boldsymbol{y_j} - \boldsymbol{\hat{y}_j}\right)^{\boldsymbol{2}} \tag{15}$$

3.4.2 Root Mean Squared Error (RMSE)

The basic hypothesis in expressing the Root Mean Squared Error is that the error is unbiased and normally distributed [14].

$$L(W, B|j)_{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}{n}} \tag{16}$$

3.4.3 Mean Absolute Error (MAE)

Mean Absolute Error is a good representation of evenly distributed errors [14].

$$\boldsymbol{L}(\boldsymbol{W}, \boldsymbol{B}|\boldsymbol{j})_{\boldsymbol{MAE}} = \frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \left|\boldsymbol{y_j} - \boldsymbol{\hat{y}_j}\right| \tag{17}$$

4 Data Description

In our analysis, we have used 23 financial ratios. These ratios have been calculated using the financial and technical statements of Iranian general insurance companies, which are active in all insurance lines of business. As it is clear, in this study, other insurance institutions such as life insurance companies, reinsurance companies, etc. are not considered. The period of study is from 2015 to 2020. The time range for the data extraction in this research is from 2015 onwards. This range was selected due to the decreasing number of insurance companies available as we move further into the past, which makes it increasingly challenging to obtain data for existing companies. Furthermore, as machine learning models are known to perform better with larger datasets, the time range was chosen accordingly. Specifically, this range was selected to avoid the exclusion of companies that were not established prior to 2015 or whose data was not available in the dataset. Consequently, the 2015 to 2020 timeframe was deemed to be the most appropriate choice for ensuring the availability of adequate data for the analysis while avoiding potential issues related to data incompleteness or exclusion of some companies. Considering that all Iranian general insurance companies were not accepted in the Tehran Stock Exchange market during the mentioned time period, and on the other hand, some companies are state-owned or private joint stock companies, access to all of the financial and technical data of these companies was not possible. Accordingly, 19 public joint stock companies were selected from among all general insurance companies operating in the insurance market of Iran, and their financial and technical data were extracted from the published financial statements and the Iran's insurance industry statistical yearbooks in order to be used in the analyses.

4.1 Financial Ratios

Following an extensive review of the existing literature, we have identified a set of 22 financial ratios to be utilized as independent variables, while designating a single financial ratio as the dependent variable to be predicted in our study. The selected ratios are described in Table 1. The "(-)" sign in the equations indicates that the account is a negative balance sheet item. These accounts are recorded as negative.

Table 1: Variables of the Study

Ratio	Name	Definition	Type of Variable
X1	Liquid Assets / Total Assets	$\frac{(Cash\ and\ cash\ equivalents + Short-term\ investments)}{(Current\ assets + Non\ current\ assets)}$	Independent
X2	Net Premium Receivables / Total Assets	$\frac{(Receivables\ from\ operating\ activities)}{(Current\ assets + Non\ current\ assets)}$	Independent
X3	Profit / Paid Capital	$\frac{(Net\ profit\ or\ loss\ of\ the\ period)}{Paid\ capital}$	Independent
X4	Payables on Reinsurance Operation / Equity	$\frac{(Debts\ to\ reinsurance\ companies)}{(Equity)}$	Independent
X5	Gross Premiums Written to Equity	$\frac{(Written\ premiums\ (Gross))}{(Equity)}$	Independent
X6	Net Premiums Written to Equity	$\frac{(Written\ premiums\ (Net))}{Equity}$	Independent
X7	Share of Reinsurance from Provisions / Equity	$\frac{\begin{pmatrix} Provision\ for\ unearned\ premiums\ (Share\ of\ Reinsurance)(-) \\ + Provision\ for\ ongoing\ risks\ (Share\ of\ Reinsurance)(-) \\ + Mathematic\ provision\ (Share\ of\ Reinsurance)(-) \\ + Provision\ for\ outstanding\ losses\ (Share\ of\ Reinsurance)(-) \\ + Provisions\ for\ returned\ premiums\ (Share\ of\ Reinsurance)(-) \\ + Provisions\ for\ profit\ sharing\ with\ policyholders\ (Share\ of Reinsurance)(-) \\ + natural\ hazards\ and \\ technical\ complementary\ provisions\ (Share\ of Reinsurance)\ (-) \end{pmatrix}}{Equity}$	Independent

Table 1: Continue

X8	Cash Ratio	$\frac{(Cash\ and\ cash\ equivalents + Short - term\ investments)}{(Short\ term\ liabilities)}$	Independent
X9	Return of Assets	$\frac{(Profit\ or\ loss\ of\ the\ period)}{(Current\ assets\ +\ Non - current\ Assets)}$	Independent
X10	Technical Profit / Gross Written Premium	$\frac{\left(\begin{array}{l} Technical\ income\ of\ Non - Life\ LOB \\ + Technical\ expense\ of\ Non - Life\ LOB \end{array}\right)}{(Written\ premiums\ (Gross))}$	Independent
X11	Equity / Total Payables	$\frac{(Equity)}{(Short\ term\ liabilities\ +\ Long\ term\ liabilities)}$	Independent
X12	Changes in Equity	$\frac{(Equity\ (Current\ Year)\ -\ Equity\ (Previous\ Year))}{(Equity\ (Previous\ Year))}$	Independent
X13	Gross Provision for Outstanding Losses / Equity	$\frac{(Provision\ for\ outstanding\ losses\ (Gross))}{(Equity)}$	Independent
X14	Net Provision for Outstanding Losses / Equity	$\frac{(Provision\ for\ outstanding\ losses\ (Net))}{(Equity)}$	Independent
X15	Gross Paid Losses / Gross Written Premiums	$-\frac{(Paid\ losses\ (Gross)\ (-))}{\big(Written\ premiums(Gross)\big)}$	Independent
X16	Operating Expenses / Gross Written Premiums	$-\frac{(Operating\ expenses\ (-))}{\big(Written\ premiums(Gross)\big)}$	Independent
X17	Financial Leverage Ratio	$\frac{(Short\ term\ liabilities\ +\ Long\ term\ liabilities)}{(Current\ assets\ +\ Non - current\ assets)}$	Independent
X18	Current Ratio	$\frac{(Current\ assets)}{(Short\ term\ liabilities)}$	Independent

Table 1: Continue

X19	Tangible Assets / Equity	$\frac{(Tangible\ assets)}{(Equity)}$	Independent
X20	Non-Current Assets / Long Term Liabilities and Equity	$\frac{(Non-current\ assets)}{(Long\ term\ liabilities\ +\ Equity)}$	Independent
X21	Financial Profitability	$\frac{(Net\ profit\ or\ loss\ of\ the\ period\ (Current\ Year))}{((Equity\ (Current\ Year)\ +\ Equity\ (Previous\ Year))\ /\ 2)}$	Independent
X22	Return on Equity	$\frac{(Profit\ or\ loss\ of\ the\ period)}{(Equity)}$	Independent
X23	Solvency Margin Ratio	$\frac{(Value\ of\ \text{existing}\ capital)}{(Risk-based\ capital)}$	Dependent

5 Discussing the Analyses

As mentioned, 23 financial ratios have been analysed in this study. Among these ratios, the solvency margin ratio is selected as the dependent variable. The reason for this is that the Central Insurance of Iran, as the regulatory body of the insurance industry, uses this ratio as the basis for taking necessary measures regarding insurance companies. The remaining 22 ratios are considered as independent variables in the analysis. The research data set includes the value of 23 ratios for 19 insurance companies over 6 years, which form a total of 2,622 data points.We employed machine learning approaches because of the substantial correlation between several of the ratios used. Because of this dependency, machine learning techniques like neural networks and random forests can handle this strong positive or negative relationship, while multivariate or linear regression models suffer from it. Because the time effect on the ratios is not taken into account, autocorrelation control is not carried out. As analysis techniques, we used Random Forest, Decision Tree, Artificial Neural Networks (Deep Learning), Gradient Boosting Machine, and Extreme Gradient Boosting (XGBoost). Grid approaches are used in the initial stage to determine which model is most appropriate. The best model's parameters are then used to execute the analyses. The Boruta feature selection approach confirms 7 out of the 22 ratios in the data set utilized for this study, and the new data set is acquired as 114 rows (companies × years) and 7 columns (ratios). Boruta is derived from a decision tree technique that works with the methods chosen in this study, hence it does not implement another feature selection method like principal component analysis or factor analysis. The train data set is 2015–2019, and the test data set is 2020. The train data set has 95 rows, while the test data set has 19 rows. Using the values from 2015 to 2019 as a training set, we try to estimate the 2020 values in all analyses.

5.1 Feature Selection with Boruta

The results of implementing the Boruta feature selection on the training data set are shown in the table below. As shown in Table 2, 7 variables out of a total of 22 independent variables have been identified as important variables and 14 variables are unimportant. Also, 1 variable has been identified as tentative.

Table 2: The Output of Boruta Algorithm

Boruta performed 1000 iterations	
Variables confirmed as important	X1, X5, X6, X9, X11, X12, X17
Variables confirmed as unimportant	X2, X3, X4, X7, X8, X10, X13, X14, X15, X16, X18, X20, X21, X22
Tentative variables left	X19

The new data set is made up of the ratios that were determined to be important. Both the data set (114×23) and the data set (114×8) that are collected from the Boruta are used in this study's analyses. As a result, the following analyses make use of two data sets. All data are separated from the 2020 data. The 2020 data, which makes up about 17% of all our data, is meant to be the test data sets. In brief, the data points in the Train, Test, Boruta Train, and Boruta Test data sets are (95×23), (19×23), (95×8), and (19×8), respectively (Observation × Feature).

5.2 Prediction with Decision Tree

The first step is to apply the Grid Search for Decision Tree to the training data. In order to achieve the best results, 45 alternative models have been developed and are arranged from small to large based on RMSE values. Once the optimum model parameters have been found, the Decision Tree is then executed. The train data set for the first Decision Tree model consists of the 23 ratios between 2015 and 2019, and the test data set consists of the 23 ratios in 2020. Fig. 3 shows a graph of the RMSE values of training based on max_depth.

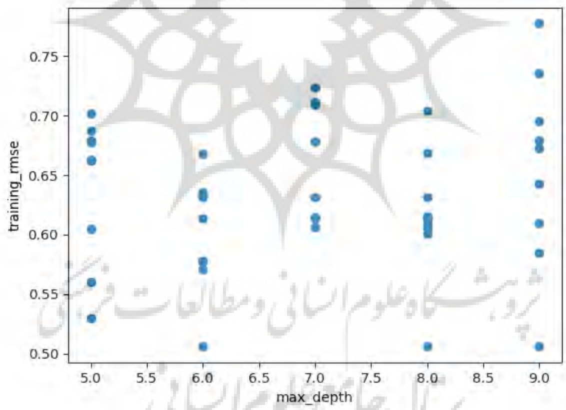

Fig. 3: Training Scoring History of Decision Tree

Table 3 shows the train's and the test's performances. The RMSE value of the training data is 0.2593, while the RMSE value of the test data is 0.8427 as a result of the first Decision Tree analysis.

Table 3: Train and Test Performance of Decision Tree

Measure	DT Train	DT Test
MSE	0.0672	0.7101
RMSE	0.2593	0.8427
MAE	0.1791	0.5296

Sets of data with 7 ratios acquired with the Boruta in the same Decision Tree model are labelled as the 2015-2019 train and 2020 test data. The model is then run. Table 4 compares the performance of test and train data obtained with Boruta data. Fig. 4 depicts a graph of the RMSE values of training based on max_depth.

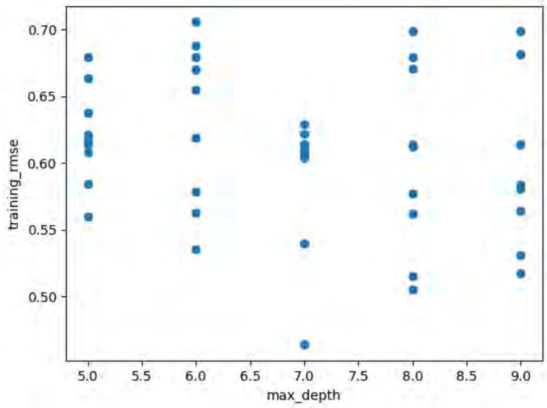

Fig. 4: Training Scoring History of Decision Tree with Boruta Data Set

Table 4: Train and Test Performance of Decision Tree with Boruta Data Set

Measure	DT Train	DT Test
MSE	0.0941	0.6802
RMSE	0.3068	0.8247
MAE	0.2181	0.5949

As a result, when the number of independent variables in the Decision Tree method is reduced from 22 to 7, the RMSE of the train predictions increases while the model's test predictions decrease.

5.3 Prediction with Random Forest

In the second step, Grid Search for Random Forest was performed on the training data set. 200 alternative models have been developed to achieve the best results and are arranged from small to large based on RMSE values. The Random Forest is then executed once the optimal model parameters have been determined. The train data set for the first Random Forest model contains the 23 ratios from 2015 to 2019, while the test data set contains the 23 ratios from 2020. A graph of the RMSE values of training based on n_estimators is also shown in Fig. 5.

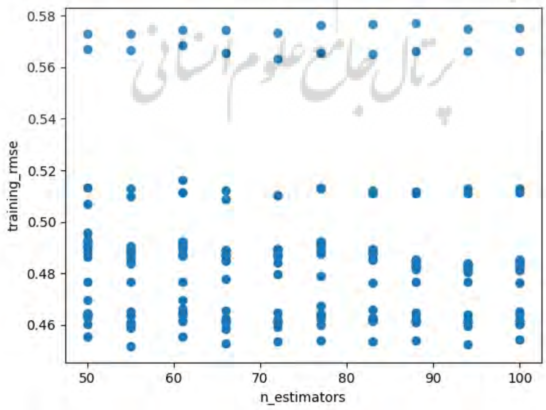

Fig. 5: Training Scoring History of Random Forest

Table 5 show the results of the train and the test. As a result of the first Random Forest analysis, the RMSE value of the training data is 0.2027, while the RMSE value of the test data is 0.6023.

Table 5: Train and Test Performance of Random Forest

Measure	RF Train	RF Test
MSE	0.0411	0.3628
RMSE	0.2027	0.6023
MAE	0.1381	0.4676

The 2015-2019 train and 2020 test data are sets of data with 7 ratios acquired with the Boruta in the same Random Forest model. The model is then run. The performance of test and train data obtained with Boruta data is compared in Table 6. A graph of the RMSE values of training based on n_estimators is shown in Fig. 6.

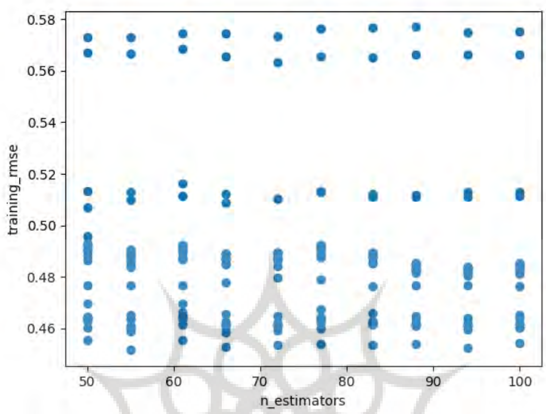

Fig. 6: Training Scoring History of Random Forest with Boruta Data Set

Table 6: Train and Test Performance of Random Forest with Boruta Data Set

Measure	RF Train	RF Test
MSE	0.0321	0.3794
RMSE	0.1792	0.6159
MAE	0.1234	0.4693

As a result, decreasing the number of independent variables in the Random Forest method from 22 to 7 decreases the RMSE of the train predictions while increasing the RMSE of the model's test predictions.

5.4 Prediction with Artificial Neural Networks

Thirdly, a deep learning method is used to conduct an analysis. First, the Grid Search is used to choose the parameters for the neural networks. Choosing the best parameters for the data set is the goal. To get the best results, 16 different models with respect to the hyperparameters have been created and are ranked from small to large based on RMSE values. These are run for the 2015-2019 train and 2020 test data, and then for the Boruta data set in the same manner. Fig. 7 depicts a graph of the MAE values of training based on the number of epochs with respect to both the initial parameters chosen and the Grid Search best parameters.

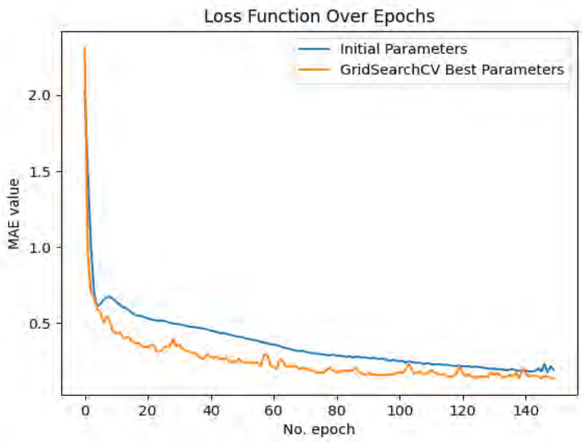

Fig. 7: Training Scoring History of Artificial Neural Networks

The Artificial Neural Networks' RMSE value after completing machine learning using training data is 0.2714. Details about test and train performance are listed in Table 7.

Table 7: Train and Test Performance of Artificial Neural Networks

Measure	ANN Train	ANN Test
MSE	0.0737	0.6957
RMSE	0.2714	0.8341
MAE	0.1336	0.7208

The Artificial Neural Networks method is repeated with the parameters from the first Artificial Neural Networks analysis on the data set obtained with Boruta. Table 8 compares the performance of test and train data obtained with Boruta data. Fig. 8 depicts a graph of the MAE values of training based on the number of epochs with respect to both the initial parameters chosen and the Grid Search best parameters.

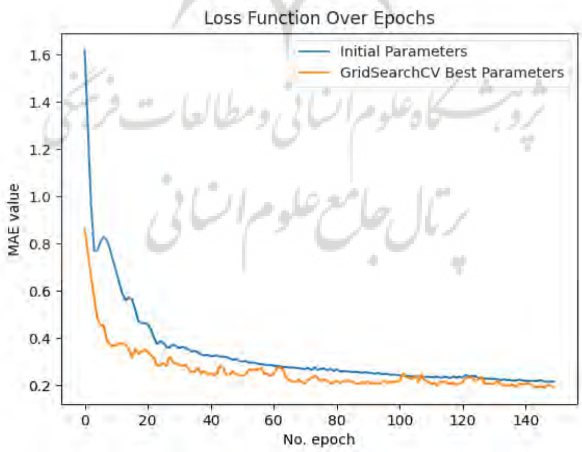

Fig. 8: Training Scoring History of Artificial Neural Networks with Boruta Data Set

Table 8: Train and Test Performance of Artificial Neural Networks with Boruta Data Set

Measure	ANN Train	ANN Test
MSE	0.1197	0.4504
RMSE	0.3460	0.6711
MAE	0.2168	0.5591

The Deep Learning study found that the RMSE value in training is 0.2714 and 0.8341 in testing. In the study conducted with the Boruta data, the training RMSE is 0.3460, while the testing RMSE is 0.6711. When the size of the data set used in Artificial Neural Networks analysis is reduced, the test performance improves while the train performance deteriorates.

5.5 Prediction with Gradient Boosting Machine

Fourth, the Grid search is incorporated into training data for the Gradient Boosting Machine. To get the best possible outcomes, 256 different models have been created and are ranked from small to large based on RMSE values. After selecting the best model parameters, the Gradient Boosting Machine is then used. The 23 ratios from 2015 to 2019 are included in the train data set for the first Gradient Boosting Machine model, whereas the 23 ratios from 2020 are included in the test data set. Fig. 9 illustrates a graph of the RMSE values for training with n_estimators.

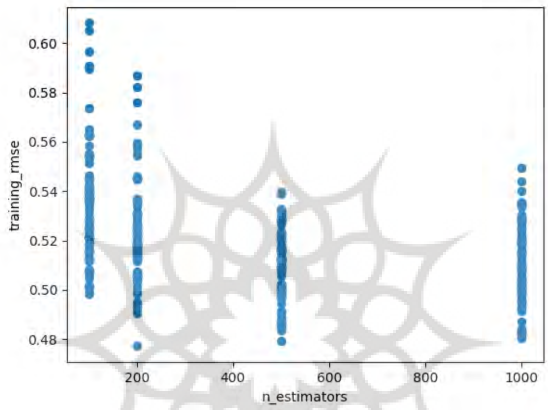

Fig. 9: Training Scoring History of Gradient Boosting Machine

Tables 9 present the train and test results. The RMSE value of the training data is 0.0203 as a result of the first Gradient Boosting Machine analysis, while the RMSE value of the test data is 0.6743.

Table 9: Train and Test Performance of Gradient Boosting Machine

Measure	GBM Train	GBM Test
MSE	0.0004	0.4546
RMSE	0.0203	0.6743
MAE	0.0142	0.4956

The analysis is repeated with the same parameters and splits for Boruta's data set. Table 10 compares the performance of test and train data obtained with Boruta data. Fig. 10 depicts a graph of the RMSE values of training based on n_estimators.

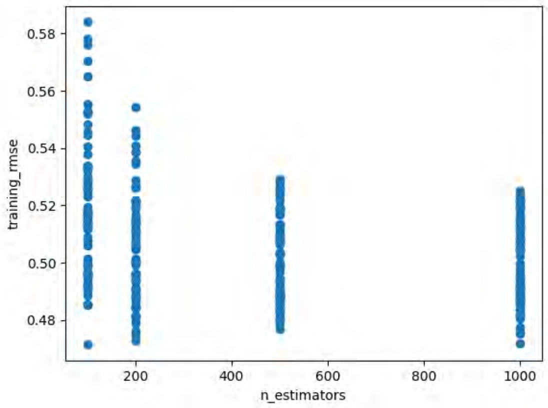

Fig. 10: Training Scoring History of Gradient Boosting Machine with Boruta Data Set

Table 10: Train and Test Performance of Gradient Boosting Machine with Boruta Data Set

Measure	GBM Train	GBM Test
MSE	0.0296	0.3840
RMSE	0.1721	0.6197
MAE	0.1353	0.4865

This causes the RMSE of the train predictions to increase while the RMSE of the model's test predictions to decrease when the Gradient Boosting Machine method's independent variable count is reduced from 22 to 7.

5.6 Prediction with Extreme Gradient Boosting

Finally, the XGBoost method's parameters are chosen using the Grid search. 600 alternative models have been developed and are ordered from small to large based on RMSE values in order to achieve the best results. The XGBoost is subsequently employed following the selection of the ideal model parameters. The train data set for the first XGBoost model has the 23 ratios from 2015 to 2019, whereas the test data set contains the 23 ratios from 2020. A graph of the RMSE values for training with n_estimators is also shown in Fig. 11.

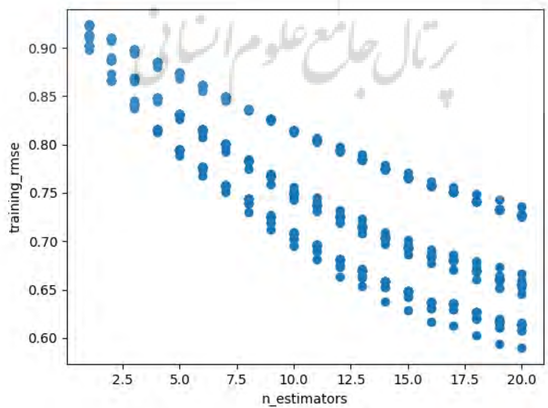

Fig. 11: Training Scoring History of XGBoost

Tables 11 present the train and test results. The RMSE value of the training data is 0.4420 as a result of the first XGBoost analysis, while the RMSE value of the test data is 0.5710.

Table 11: Train and Test Performance of XGBoost

Measure	XGBoost Train	XGBoost Test
MSE	0.1954	0.3260
RMSE	0.4420	0.5710
MAE	0.2915	0.4080

For Boruta's data set, the analysis is performed once more using the same parameters and splits. The performance of test and train data derived using Boruta data is compared in Table 12. A graph of the RMSE values for training using n_estimators is shown in Fig. 12.

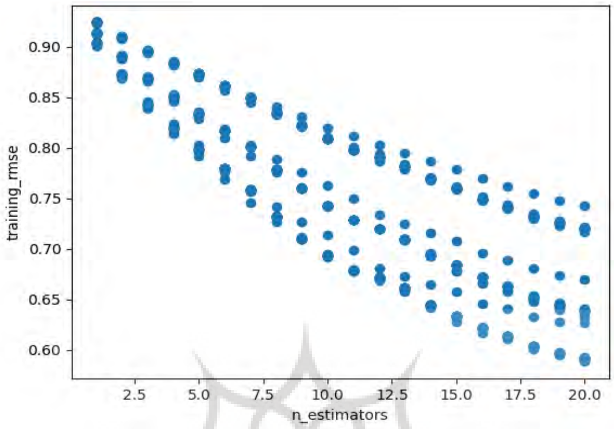

Fig. 12: Training Scoring History of XGBoost with Boruta Data Set

Table 12: Train and Test Performance of XGBoost with Boruta Data Set

Measure	XGBoost Train	XGBoost Test
MSE	0.1537	0.3521
RMSE	0.3921	0.5934
MAE	0.2468	0.4526

When the XGBoost method's independent variable count is decreased from 22 to 7, this results in a decrease in the RMSE of the train predictions and an increase in the RMSE of the model's test predictions.

5.7 Comparison of Results

The results of the MSE values from all analyses may be seen in Table 13 for both train and test, in summary. Values for RMSE and MAE are shown in Tables 14 and 15, respectively.

Table 13: Comparison of MSE Values for the Applied Machine Learning Techniques

Method	Measure	
	Train MSE	**Test MSE**
Decision Tree	0.0672	0.7101
Random Forest	0.0411	0.3628
Artificial Neural Networks	0.0737	0.6957
Gradient Boosting Machine	0.0004	0.4546
XGBoost	0.1954	0.3260
Decision Tree with Boruta	0.0941	0.6802
Random Forest with Boruta	0.0321	0.3794
Artificial Neural Networks with Boruta	0.1197	0.4504
Gradient Boosting Machine with Boruta	0.0296	0.3840
XGBoost with Boruta	0.1537	0.3521

Table 14: Comparison of RMSE Values for the Applied Machine Learning Techniques

Method	Measure	
	Train RMSE	Test RMSE
Decision Tree	0.2593	0.8427
Random Forest	0.2027	0.6023
Artificial Neural Networks	0.2714	0.8341
Gradient Boosting Machine	0.0203	0.6743
XGBoost	0.4420	0.5710
Decision Tree with Boruta	0.3068	0.8247
Random Forest with Boruta	0.1792	0.6159
Artificial Neural Networks with Boruta	0.3460	0.6711
Gradient Boosting Machine with Boruta	0.1721	0.6197
XGBoost with Boruta	0.3921	0.5934

Table 15: Comparison of MAE Values for the Applied Machine Learning Techniques

Method	Measure	
	Train MAE	Test MAE
Decision Tree	0.1791	0.5296
Random Forest	0.1381	0.4676
Artificial Neural Networks	0.1336	0.7208
Gradient Boosting Machine	0.0142	0.4956
XGBoost	0.2915	0.4080
Decision Tree with Boruta	0.2181	0.5949
Random Forest with Boruta	0.1234	0.4693
Artificial Neural Networks with Boruta	0.2168	0.4693
Gradient Boosting Machine with Boruta	0.1353	0.4865
XGBoost with Boruta	0.2468	0.4526

According to the test MSE, RMSE, and MAE values in the tables above, the XGBoost method is the best machine learning model because it has the smallest values when compared to the other techniques. Very low RMSE values, however, could be a sign of overfitting.

A model starts to learn from the noise and false information in the data set when it is given such a large amount of information. The model then incorrectly classifies the data as a result of too much detail and noise. Extremely high training efficiency can result in extremely poor test efficiency.

5.8 Analysis of Regulatory Measures Stages

As stated before, the Central Insurance of Iran, according to Regulation No. 69 of the High Council of the Insurance Industry of Iran, uses the values obtained from the solvency margin ratio as the basis for judging the financial health of insurance companies, and based on these values, it defines 5 stages of solvency. According to the regulation mentioned above, if the insurance institution's solvency margin ratio is at stage 2 according to the opinion of the Central Insurance of Iran, the institution is obliged to prepare a plan to restore its financial status for the next three financial years (divided annually) and submit it to the Central Insurance of Iran.

In this plan, the insurance company must show how it will improve its solvency margin ratio to at least stage 1 within three years. If the solvency margin ratio of the insurance institution is at stage 3 according to the opinion of the Central Insurance of Iran, the institution is obliged to prepare its capital increase plan for the next two financial years (divided annually) in addition to the plan to restore the financial situation and submit it to the Central Insurance of Iran for approval. In these plans, the insurance company must show how it will improve its solvency margin ratio to at least stage 2 within two years. If the solvency margin ratio of the insurance institution is at stage 4 according to the opinion of the Central Insurance of Iran, the institution is obliged to prepare a plan to restore its financial status and increase its capital for the next fiscal year and submit it to the Central Insurance of Iran. In these plans, the insurance company must show how it will improve its solvency margin ratio to at least stage 3 within one year. Finally, if the solvency margin ratio of the insurance company is at stage 5 according to the announcement of the Central Insurance of Iran, the regulatory body is allowed to suspend or cancel the business license of the insurance company in one or more lines of insurance business. In this case, our objective was to determine whether the predicted numbers actually fall inside the specified range.

The observed SMR values for 2020 and values estimated using all methods are numbered in accordance with the solvency stages mentioned above4According to this, values greater than 100% take ‘1’, values between 70% to 100% take ‘2’, values between 52% to 70% take ‘3’, values between 10% to 50% take ‘4’ and values less than 10% take ‘5’.

Table 16 shows the values obtained from the 2020 forecasts in the analysis where the years 2015-2019 are utilized as training data. It is deemed "True" if the projected value and the actual value fall within the same range, and "False" otherwise. The accuracy rate is derived by dividing the total number of observations by the number of observations that were accurately estimated.

Table 16: Predictions by Solvency Stages

Company	y	**DT** $\hat{y}$	**DT_b** $\hat{y}$	**RF** $\hat{y}$	**RF_b** $\hat{y}$	**ANN** $\hat{y}$	**ANN_b** $\hat{y}$	**GBM** $\hat{y}$	**GBM_b** $\hat{y}$	**XGB** $\hat{y}$	**XGB_b** $\hat{y}$
1	1	2	1	1	1	5	4	2	1	1	2
2	1	1	3	1	1	1	1	1	1	1	1
3	1	2	2	1	2	1	2	2	2	2	2
4	1	2	1	1	1	1	2	1	1	2	2
5	1	1	3	1	1	5	3	1	1	1	1
6	1	1	1	1	1	3	2	1	1	1	1
7	1	1	1	1	1	1	2	2	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1
9	2	3	3	1	2	5	4	2	2	2	2
10	1	3	3	1	1	1	1	1	1	2	2
11	1	1	3	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1
14	1	3	3	2	2	4	3	3	4	2	3
15	1	1	3	1	1	1	1	1	1	1	2
16	1	1	1	1	1	1	1	1	1	1	1
17	1	3	2	2	2	2	1	2	2	2	2
18	1	1	1	1	1	1	1	1	1	1	2
19	1	1	1	1	1	1	1	1	1	1	1
False		7	9	3	3	6	8	5	3	5	8
True		12	10	16	16	13	11	14	16	14	11
Percentage		63.16 %	52.63 %	84.21 %	84.21 %	68.42 %	57.89%	73.68 %	84.21%	73.68 %	57.89%

In 2020, there are 19 observations. If the actual and forecast values of these observations differ, we can conclude that our method's predictions for that company fail. Table 16 shows that if the actual and predicted values differ, the predicted value is colored red; otherwise, it is colored green. As a result, the methods with the fewest errors (84%) are Random Forest on both data sets (full data set and Boruta data set) and Gradient Boosting Machine on the Boruta data set. When both algorithms are run on the Boruta data set, the XGBoost, which has the best performance based on RMSE value on test data set, ranks second jointly with the Gradient Boosting Machine in terms of estimation accuracy. Table 17 briefly shows the ranking of each method based on the estimation accuracy.

Table 17: Estimation Accuracy Ranking for Solvency Stage

Rank	Method	Estimation Accuracy
1	Random Forest	84.21%
1	Random Forest with Boruta	84.21%
1	Gradient Boosting Machine with Boruta	84.21%
4	XGBoost	73.68%
4	Gradient Boosting Machine	73.68%
6	Artificial Neural Networks	68.42%
7	Decision Tree	63.16%
8	Artificial Neural Networks with Boruta	57.89%
8	XGBoost with Boruta	57.89%
10	Decision Tree with Boruta	52.63%

5.9 Analysis on SMR

An assessment is made to determine whether or not the company's capital is adequate. The capital of insurance institutions should not be less than the risk-based capital, according to Regulation No. 69 of the High Council of the Insurance Industry of Iran. That is, the solvency margin ratio should be greater than 100%. Despite the fact that the initial range estimates investigation provides more sensitive accuracy, the second examination allows it to be understood to what extent it can be accurately predicted whether the capital of the companies is sufficient or not. If the company's solvency margin ratio is greater than 100%, it has 1; otherwise, it has 0. As a result, we can determine whether the capital adequacy status is sufficient or insufficient. Table 18 also shows the results of the actual and estimated values. If these two values do not correspond, the estimate values are highlighted in red. At the bottom of the table are the false numbers, true numbers, and accuracy percentages for each method.

Random Forest on both data sets (full data set and Boruta data set) was discovered to be the best predictive analysis method, followed by Gradient Boosting Machine on the Boruta data set, similar to the analysis in Section 5.8. With these methods, we can predict whether the capital of the companies is adequate or not with respect to the risks they bear with 84% accuracy. In all algorithms, except Gradient Boosting Machine, the percentage of success with the Boruta data set reduces. At the same time, the Decision Tree and XGBoost methods, which are run with the Boruta data set, have the worst prediction of the dependent variable with an accuracy of 58%. Another interesting point is that all methods except Random Forest predict the insolvency of company No. 9.

Table 18: Predictions of SMR 1-0

Company	y	DT $\hat{y}$	DT_b $\hat{y}$	RF $\hat{y}$	RF_b $\hat{y}$	ANN $\hat{y}$	ANN_b $\hat{y}$	GBM $\hat{y}$	GBM_b $\hat{y}$	XGB $\hat{y}$	XGB_b $\hat{y}$
1	1	0	1	1	1	0	0	0	1	1	0
2	1	1	0	1	1	1	1	1	1	1	1
3	1	0	0	1	0	1	0	0	0	0	0
4	1	0	1	1	1	1	0	1	1	0	0
5	1	1	0	1	1	0	0	1	1	1	1
6	1	1	1	1	1	0	0	1	1	1	1
7	1	1	1	1	1	1	0	0	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1
9	0	0	0	1	0	0	0	0	0	0	0
10	1	0	0	1	1	1	1	1	1	0	0
11	1	1	0	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1
14	1	0	0	0	0	0	0	0	0	0	0
15	1	1	0	1	1	1	1	1	1	1	0
16	1	1	1	1	1	1	1	1	1	1	1
17	1	0	0	0	0	0	1	0	0	0	0
18	1	1	1	1	1	1	1	1	1	1	0
19	1	1	1	1	1	1	1	1	1	1	1
False		6	8	3	3	5	7	5	3	5	8
True		13	11	16	16	14	12	14	16	14	11
Percentage		68.42%	57.89%	84.21%	84.21%	73.68%	63.16%	73.68%	84.21%	73.68%	57.89%

Table 19 also briefly shows the ranking of each method based on the estimation accuracy.

Table 19: Estimation Accuracy Ranking for SMR

Rank	Method	Estimation Accuracy
1	Random Forest	84.21%
1	Random Forest with Boruta	84.21%
1	Gradient Boosting Machine with Boruta	84.21%
4	Artificial Neural Networks	73.68%
4	Gradient Boosting Machine	73.68%
4	XGBoost	73.68%
7	Decision Tree	68.42%
8	Artificial Neural Networks with Boruta	63.16%
9	XGBoost with Boruta	57.89%
9	Decision Tree with Boruta	57.89%

6 Conclusion

In general, the insurance industry's primary concern is whether or not companies can pay their debts. This issue is critical to the company's, industry's, and even the entire economy's continuity and robustness. We can say that the most important indicator is whether the companies' capital is sufficient to pay their liabilities. The use of the solvency margin ratio by regulators to take action against insurance companies is the strongest supporter of this. As a result of this ratio, the company's situation can be clearly seen, and it can be warned or even intervened.The establishment of an early warning model using machine learning methods can be viewed as the starting point of this article. Based on previous research, we intend to develop an early warning model using Machine Learning algorithms and financial data from Iranian general insurance companies. SMR was chosen as the dependent variable. While the 2015-2019 data set is defined as training data, dependent variable in 2020 is wanted to be estimated using twenty-two independent variables. The independent variables encompass a range of financial ratios pertaining to insurance companies, whereby their respective values are derived through meticulous calculations utilizing data gleaned from the financial statements of these insurers.In this research, we used Decision Tree, Random Forest, Artificial Neural Networks, Gradient Boosting Machine, and Extreme Gradient Boosting algorithms to predict the values of the solvency margin ratio of Iranian general insurance companies, so that based on the results, we can identify the best machine learning algorithm to establish an early warning model. But before doing this, we used the Boruta feature selection method in order to identify the independent variables that have the highest importance in the estimation of the dependent variable. This is to increase the predictive power of the models as the information we teach to the machine increases. As a result of this work, the Boruta method introduced seven variables of Liquid Assets / Total Assets, Gross Premiums Written to Equity, Net Premiums Written to Equity, Return of Assets, Equity / Total Payables, Changes in Equity and Financial Leverage Ratio as the most important variables, and the observations related to these variables were separated from others and formed a new data set. As a result, we were left with two data sets (full data set and Boruta data set) that we ran the selected algorithms on both of them separately.After using the grid search method, we were able to identify the best parameters for each of the algorithms and choose the best model for prediction. We then proceeded to predict the values of 2020. Based on the Root Mean Squared Error (RMSE) values obtained for the test data set in each method, it was found that the XGBoost algorithm has the best performance both on the full data and on the Boruta data compared to the others. However, XGBoost performed the best on the full data set in comparison to the Boruta data set.In the next step, the predicted and actual values are compared to see if they fall within the same range. Random Forest on both data sets (full data set and Boruta data set) and Gradient Boosting Machine on the Boruta data set have the highest accuracy percentages (84%). The estimation and actual values are then compared to determine whether capital is sufficient in relation to the risks that insurance companies encounter, and the best models, once again, are the Random Forest on both data sets (full data set and Boruta data set) and the Gradient Boosting Machine on the Boruta data set with 84% accuracy.Based on the findings, this model can be used to monitor the current state of insurance companies and serve as a guide for future actions. It will support supervisory and regulatory authorities' ability to conduct risk-focused supervision.

References

[1] Kasbi, P., Aghaei, M. A., Fair Value Accounting for Liabilities and Own Credit Risk, *Advances in Mathematical Finance and Applications*, 2017; 2(1):55-68. Doi: 10.22034/amfa.2017.529061

[2] Díaz, Z., Segovia, M.J., Fernández, J., del Pozo, E., Machine learning and statistical techniques. An application

to the prediction of insolvency in Spanish non-life insurance companies, *The International Journal of Digital Accounting Research*, 2005; 5(9):1-45. Doi: 10.4192/1577-8517-v5_1

[3] Haghverdilou, M., Peykarjou, K. and ZomorodianS, G.R., Introducing Early Warning System for Solvency of Iranian Insurance Companies, Using Pane data method, *Journal of Investment Knowledge*, 2022; 11(44):417-452, (in Persian).

[4] Wang, T., Zhao, S., Zhu G., Zheng, H., A machine learning-based early warning system for systemic banking crises, *Applied economics*, 2021; 53(26):2974-2992. Doi: 10.1080/00036846.2020.1870657

[5] Caporale, G.M., Cerrato, M. and Zhang, X., Analysing the determinants of insolvency risk for general insurance firms in the UK, *Journal of Banking & Finance*, 2017; 84: 107-122. Doi: 10.1016/j.jbankfin.2017.07.011

[6] Beaver, W. H., Financial Ratios as Predictors of Failure, *Journal of Accounting Research*, 1966; 4:71–111. Doi: 10.2307/2490171

[7] Brockett, P. L., Cooper, W. W., Golden, L. L., & Pitaktong, U., A Neural Network Method for Obtaining an Early Warning of Insurer Insolvency, *The Journal of Risk and Insurance*, 1994; 61(3):402–424. Doi: 10.2307/253568

[8] Segovia-Vargas, M.J., Salcedo-Sanz, S. and Bousono-Calzon, C., Prediction of Insolvency in Non-life Insurance Companies Using Support-vector Machines, *Genetic Algorithms and Simulated Annealing*, *Fuzzy Economic Review*, 2004; 9(1):79-94. Doi: 10.25102/fer.2004.01.05

[9] Sharpe, I.G. and Stadnik, A., Financial distress in Australian general insurers, *Journal of Risk and Insurance*, 2007; 74(2):377-399. Doi: 10.1111/j.1539-6975.2007.00217.x

[10] Rustam, Z. and Yaurita, F., Insolvency prediction in insurance companies using support vector machines and fuzzy kernel c-means, *Paper presented at 2nd International Conference on Statistics, Mathematics, Teaching, and Research*, 2018.

[11] Ayodele, T.O., "Types of machine learning algorithms", in Zhang, Y., *New advances in machine learning, Croatia, InTech*, 2010:19-48.

[12] Zhao, Y., Research on personal credit evaluation of internet finance based on blockchain and decision tree algorithm, *EURASIP Journal on Wireless Communications and Networking*, 2020; :1-12. Doi: 10.1186/s13638-020-01819-w

[13] Song, Y.Y. and Ying, L.U., Decision tree methods: applications for classification and prediction, *Shanghai archives of psychiatry*, 2015; 27(2):130-135. Doi: 10.11919/j.issn.1002-0829.215044

[14] Koçer, G. B., Early warning model with machine learning for Turkish insurance sector, *M.Sc. Thesis, Middle East Technical University, Ankara, Turkey*, 2019.

[15] Tehrani, R., Souri, A., Zohrabi, A., Sadeghi Sharif, S. J., Provide an improved factor pricing model using neural networks and the gray wolf optimization algorithm, *Advances in Mathematical Finance and Applications*, 2023; 8(1):23-45. Doi: 10.22034/amfa.2022.1942547.1641

[16] Codrut Florin, I., Option pricing using Machine Learning, *Expert Systems with Applications*, 2021; 163:113799. Doi: 10.1016/j.eswa.2020.113799

[17] Bärtl, M. and Krummaker, S., Prediction of Claims in Export Credit Finance: *A Comparison of Four Machine Learning Techniques, Risks*, 2020; 8(1): Doi:10.3390/risks8010022

[18] Cook, D., *Practical machine learning with H2O: powerful, scalable techniques for deep learning and AI*, United States of America, O'Reilly Media, Inc., 2016.

[19] Hastie, T., Tibshirani, R. and Friedman, J., The elements of statistical learning, *United States of America*,

Springer series in statistics, 2001.

[20] Mitchell, R. and Frank, E., Accelerating the XGBoost algorithm using GPU computing, *PeerJ Computer Science*, 2017; 3:127. Doi: 10.7717/peerj-cs.127

[21] Kursa, M.B. and Rudnicki, W.R., Feature selection with the Boruta package, *Journal of statistical software*, 2010; 36(11):1-13. Doi: 10.18637/jss.v036.i11

[22] Tavakoli, M., Doosti, H., Forecasting the Tehran Stock market by Machine Learning Methods using a New Loss Function, *Advances in Mathematical Finance and Applications*, 2021; 6(2):194-205.
Doi: 10.22034/amfa.2020.1896273.1399

[23] Candel, A., Parmar, V., LeDell, E. and Arora, A., Deep learning with H2O, *United States of America, H2O.ai, Inc*, 2016.