Journal of Data Analytics and Intelligent Decision-making **Vol. 1 Issue. 1 (2025) 33-41**

Journal of Data Analytics and Intelligent Decision-making

Journal homepage: jdaid.qom.ac.ir

University of Qom

Financial Forecasting Using an Intelligent Model Based on Reliability

Sepideh Etemadi[a], Mehdi Khashei*[b]

a. Assistant Professor, Industrial Engineering Department, Kosar university of Bojnord.
b. Assistant Professor, Department of Industrial and Systems Engineering, Isfahan University of Technology.

A R T I C L E I N F O

Keywords:
Multilayer perceptron
Classification
Financial decision making
Generalizability
Accuracy-based methodology
Reliability-based methodology

A B S T R A C T

The functional logic of classifier models is based on the principle that, to maximize their ability to generalize—an essential factor affecting decision quality in real-world problems—it is crucial to minimize the classification error rate of available historical data. In other words, accuracy is considered the only factor affecting the generalizability of classification methods. However, due to fluctuations in financial variables, stable and reliable forecasts are also necessary for correct and profitable decision-making. Despite the importance of the reliability factor in creating stable and robust results, it has been neglected in the literature on modeling and classification. To address this research gap and enhance decision-making processes in financial applications, a modeling method based on reliability maximization is presented. This paper develops a multilayer perceptron model with the aim of maximizing reliability rather than accuracy. To evaluate the performance of the proposed model, five different financial datasets are selected from the UCI database, and its classification error rate is compared with that of the conventional multilayer perceptron model. The findings show that the reliability factor has a greater impact than the accuracy factor on the generalizability and performance of classification models. The results indicate that the proposed reliability-based multilayer perceptron model demonstrates superior efficiency and performance compared to the conventional multilayer perceptron model and can serve as a viable alternative in financial applications.

1. Introduction

Classification is one of the fundamental data mining methods for assigning a target variable in studied systems to a specific class among existing classes. It is effectively used across a wide spectrum of applications in the financial domain. In recent decades, classification methods have been increasingly employed for decision-making objectives in various financial areas, such as

* Corresponding author.
E-mail addresses: etemadi.sepideh@yahoo.com (S. Etemadi), khashei@cc.iut.ac.ir (M. Khashei)

Received 3 March 2025; Received in revised form 12 April 2025; Accepted 20 April 2025
Available online 15 June 2025

Cite this article: Etemadi, S., Khashei, M. (2025). Financial Forecasting Using an Intelligent Model Based on Reliability . *Journal of Data Analytics and Intelligent Decision-making*, 1(1), 33-41. https://doi.org/10.22091/JDAID.2025.14091.1006

risk assessment and management, stock management, stock market prediction, credit scoring, and more. Given the increasing complexity, rapid changes, and multiple factors affecting issues, forecasting in the financial domain has become one of the most challenging areas for researchers (Choudhary & Thenmozhi, 2024). Various statistical and intelligent data mining methods with different characteristics, such as artificial neural networks, multilayer perceptrons, support vector machines, long short-term memory, decision trees, logistic regression, random forests, etc., have been developed for classification purposes in various financial fields.

Shrivastava et al. (2024) proposed a new cost function that allocates loss values based on the importance of samples. Comparative analysis confirmed the superiority of the proposed support vector classifier trained using the new loss function compared to other models for credit scoring of clients. Moral-García and Abellán (2023) focused on increasing diversity in classification groups to improve the accuracy of credit scoring. They increased the diversity of the ensemble by changing the values of hyperparameters for each classifier, which led to better credit scoring results. Gicić et al. (2023) proposed a hybrid architecture consisting of a stacked long short-term memory layer and a bidirectional long short-term memory for enhancing credit scoring performance. Khalili and Rasaatgar (2023) introduced a hybrid model that utilized a genetic algorithm for feature selection and optimized parameters for each algorithm using a cost-sensitive objective function based on the probability of correctly classifying samples. Nazemi et al. (2022) evaluated the performance of deep learning classifiers, support vector machines, and hybrid models in predicting the collection rate of non-compliant customer credits. The findings of this study emphasized the potential of deep learning as a tool that can improve performance in credit risk management and provide higher accuracy compared to other classifiers.

Among various classification approaches, multilayer perceptron is a commonly used intelligent classification method that is widely utilized in modeling and data mining. Zhao et al. (2015) presented a credit scoring method based on multilayer perceptron networks where model parameters are determined using the backpropagation algorithm. Mabu et al. (2015) introduced an optimized multilayer perceptron with a swarm learning algorithm in the context of stock trading. Golbayani et al. (2020) compared the accuracy of multilayer perceptron, support vector machine, random forest, and decision tree for credit scoring of firms in financial domains. Zhang and Luo (2021) compared the accuracy of neural networks with a deep fuzzy learning algorithm for classifying stock price patterns. Tsai et al. (2021) primarily focused on evaluating the performance of several widely used classifiers, such as artificial neural networks, support vector machines, logistic regression, and decision trees, as well as combining these methods for predicting bankruptcy and credit scoring. Vrontos et al. (2021) predicted bankruptcy and credit risk through an ant colony optimization-based algorithm that quantitatively and qualitatively extracted decision-making rules from expert decisions. The proposed model outperformed the baseline functions, multilayer perceptron, logistic regression, and random forest model. Vrontos et al. (2021) compared and evaluated the performance of nearest neighbor methods, generalized linear Bayesian approaches, and penalized logistic regression for predicting the probability of a recession in the United States. Jurgovsky et al. (2018) designed a long short-term memory network for identifying credit card fraud. The accuracy of the proposed model was higher than that of the random forest model.

Despite the various classification approaches, they have all been developed based on the same assumption that minimizing the classification error rate on the training dataset leads to greater generalizability on the test datasets. In other words, generalizability in these types of models is solely dependent on performance accuracy. Although accuracy is one of the most important factors influencing the generalization power of the model, it is not the unique explanatory factor for how the generalizability of the model changes. It seems that the generalizability of a classification model depends simultaneously on the model's accuracy as well as the level of confidence in the results. In fact, given the fluctuations and variations in financial parameters, achieving stable and reliable results is as important as generating accurate outcomes. Therefore, reliability is another important factor that must be considered to achieve generalizable models. Hence, this paper proposes a reliability-based multilayer perceptron model to achieve more stable and reliable results. The main idea of the proposed modeling approach is to develop a process that minimizes the performance fluctuations of the multilayer perceptron model and overcomes the limitations of the classical multilayer perceptron, which is developed solely based on maximizing accuracy. To demonstrate the efficiency and effectiveness of the proposed model, the reliability-based modeling performance is evaluated using five financial datasets from the UCI repository, and the results are compared with the conventional multilayer perceptron model. The results indicate the superiority of the proposed method over the classical model in financial applications. The rest of the paper is organized as follows: The methodology section presents the concepts and formulas of the proposed reliability-based multilayer perceptron classifier. The data description section elaborates on the financial datasets used to evaluate the proposed model. In the results analysis section, the performance of the proposed multilayer perceptron model is compared with that of the classical multilayer perceptron model across five benchmark financial datasets. Finally, the conclusion is presented.

2. Research Methodology

In conventional multilayer perceptron models, the modeling logic has always been based on minimizing the classification error rate on the training data. Accordingly, the basis for achieving maximum accuracy on the test data or the model's generalization power is assumed to be exclusively through maximizing accuracy on the training data. Since the quality of decisions is closely related to the model's ability to generalize in real-world problems, considering performance accuracy as the only effective factor on the model's generalization ability and neglecting the potential for increased generalization by improving other influential factors is not correct. Consequently, improving the generalization capacity of models has become one of the most challenging research areas in prediction and modeling literature. In this regard, it seems that confidence in performance accuracy is one of the influential factors on the model's generalizability, which has not been sufficiently addressed in the conventional multilayer perceptron modeling process. Stability in modeling a system with different data refers to the reliability of the results. Additionally, the ability to generalize the model means achieving accurate and reliable predictions. Therefore, confidence levels and accuracy are effective factors on the model's generalization ability, and modeling uncertainty is considered. In this paper, a multilayer perceptron classification modeling idea is presented that maximizes the model's reliability instead of accuracy. The proposed operational logic is inspired by calculating the variations in the performance accuracy of models executed on validation data. Thus, the basis of the proposed approach is that multilayer perceptron models with less variation in accuracy against validation data are expected to yield more reliable results on test data. In other words, the less the performance fluctuations are in the validation data, the greater the stability of the performances in the out-of-sample or test data. In this way, the logic for creating a reliability-based multilayer perceptron model is described.

The multilayer perceptron is one of the most popular and widely used artificial neural networks for modeling and prediction. The multilayer perceptron is typically used to model the complex nonlinear relationship between a dependent variable and several independent variables. Thus, a multilayer perceptron model with k variables, including the dependent variable Y and m explanatory variables $X_1, X_2, \ldots X_m$, can generally be represented as follows:

$$Y_t = \beta_0 + \sum_{j=1}^{p} \beta_j . g\left(\beta_{0j} + \sum_{i=1}^{m} \beta_{ij} . X_{ti} \right) + u_t \qquad t = 1,2,3,\ldots,N \tag{1}$$

here β_{ij} and β_j are the model weights, p is the number of hidden neurons, g is the activation function of the hidden layer, u_t is the random error term, and N is the sample size. Estimating the model weights using a gradient optimization method based on a cost function, namely the sum of squared classification errors, is one of the most popular accuracy-based techniques. In the optimization algorithm, the unknown parameters and weights are estimated in such a way that the sum of squared classification errors, which is the difference between the actual values and the fitted values, is minimized. However, like other accuracy-based estimators, this method focuses on accuracy as the primary influencing factor on generalization ability. In contrast, the proposed least squares method based on reliability minimizes the variations in classification errors.

According to the proposed method, a portion of the training data is initially considered as validation datasets. Subsequently, in this research, the accuracy is calculated as the sum of squared classification errors for the training data and also for the training data along with each data point from the validation datasets, as described below:

$$\sum_{t=1}^{N} \hat{u}_{0t}^{\,2} = \sum_{t=1}^{N} \left(Y_t - \left(\hat{\beta}_{00} + \sum_{j=1}^{p} \hat{\beta}_{0j} . g\left(\hat{\beta}_{00j} + \sum_{i=1}^{m} \hat{\beta}_{0ij} . X_{ti} \right) \right) \right)^2 \tag{2}$$

and similarly, for each validation dataset:

$$\begin{gathered} \sum_{t=1}^{N+1} \hat{u}_{1t}^{\,2} = \sum_{t=1}^{N+1} \left(Y_t - \left(\hat{\beta}_{10} + \sum_{j=1}^{p} \hat{\beta}_{1j} . g\left(\hat{\beta}_{10j} + \sum_{i=1}^{m} \hat{\beta}_{1ij} . X_{ti} \right) \right) \right)^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \sum_{t=1}^{N+n} \hat{u}_{nt}^{\,2} = \sum_{t=1}^{N+n} \left(Y_t - \left(\hat{\beta}_{n0} + \sum_{j=1}^{p} \hat{\beta}_{nj} . g\left(\hat{\beta}_{n0j} + \sum_{i=1}^{m} \hat{\beta}_{nij} . X_{ti} \right) \right) \right)^2 \end{gathered} \tag{3}$$

Where

$$\sum_{t=1}^{N+k} \hat{u}_{kt}^{\,2} \qquad k = 0,1,2,\ldots,n$$

is the sum of the squared residuals, and n is the size of the validation dataset. Now, to find the parameters and unknown weights of the proposed multilayer perceptron model with minimal changes in the squared classification errors across all validation data points, we have the following expression:

$$Min \quad \sum_{k=0}^{n}\sum_{t=1}^{N+n}\hat{u}_{nt}{}^{2}-\sum_{k=0}^{n}\sum_{t=1}^{N+k}\hat{u}_{kt}{}^{2}=\sum_{k=0}^{n}\left(\sum_{t=1}^{N+n}\hat{u}_{nt}{}^{2}-\sum_{t=1}^{N+k}\hat{u}_{kt}{}^{2}\right) \tag{4}$$

So that:

$$Min \quad \sum_{k=0}^{n}\left(\sum_{t=1}^{N+n}\left(Y_t-\left(\sum_{j=0}^{p}\hat{\beta}e_j.g\left(\sum_{i=0}^{m}\hat{\beta}e_{ij}.X_{ti}\right)\right)\right)^2-\sum_{t=1}^{N+k}\left(Y_t-\left(\sum_{j=0}^{p}\hat{\beta}e_j.g\left(\sum_{i=1}^{m}\hat{\beta}e_{ij}.X_{ti}\right)\right)\right)^2\right) \tag{5}$$

Where

$$\hat{\beta}e_j,\hat{\beta}e_{ij} \quad i=0,1,2,\ldots,m \quad j=0,1,2,\ldots,p$$

are the weights and biases of the proposed multilayer perceptron model. Ultimately, by considering the cost function according to the relation above, the unknown parameters can be estimated. In general, the main advantage of the reliability-based multilayer perceptron model compared to other nonlinear models is its consideration of the impact of reliability on the model's generalization ability. Moreover, by maximizing the reliability of the proposed model, the uncertainty of the proposed model is minimized compared to other nonlinear models. The proposed model is applicable to various decision-making problems and provides more efficient and accurate results in cases with greater uncertainty. Therefore, the quality of decisions derived from the reliability-based multilayer perceptron model is superior to that of other nonlinear models.

3. Data Description

In this paper, five benchmark datasets selected from the UCI repository are considered to evaluate the performance of the proposed multilayer perceptron compared to the conventional multilayer perceptron version. These datasets include real and simulated examples in various financial fields such as auditing, credit scoring, bid evaluation, stock index prediction, and bank account classification. The sample sizes of these datasets range from 536 to 45,211 data points. Additionally, the number of variables in the models varies from 5 to 23 explanatory variables. The characteristics of the five benchmark datasets are summarized in Table 1.

Table 1. General Information of Benchmark Datasets from the UCI Repository

Application Area	Variable Type	Sample/Variable	Year of Release	Title
Fraud Detection in Bidding	Categorical	11.6,321	2020	Shill Bidding
Auditing	Continuous	5.776	2018	Audit Data
Credit Scoring	Continuous/Categorical	23.30,000	2016	Default of Credit Card Clients
Stock Index Prediction	Categorical	7.536	2013	Istanbul Stock Exchange
Deposit Classification	Categorical	16.45,211	2012	Banking Marketing

4. Research Findings

In this section, the "Shill Bidding" dataset is initially selected as a sample from the entire data, and the modeling process of the proposed method, along with the related analyses, is presented. Subsequently, the results obtained from the application of the proposed multilayer perceptron model on five benchmark datasets are analyzed and discussed. This dataset includes information related to bid evaluation to detect normal behavior in bidding (class zero) versus abnormal behavior in bidding (class one). The explanatory variables include:

1. A unique identifier for a record in the dataset,
2. A unique identifier for the bid,
3. A bidder participating in a limited number of bids,
4. A bidder who increases the reserve price and attracts higher bids from legitimate participants,
5. A bidder who consistently outbids themselves even if they are the current winner to gradually raise the price with small incremental increases,
6. A bidder who becomes inactive in the final stage of bidding (more than 90% of the bidding duration) to avoid winning the auction,
7. Bids with specific activities tend to have significantly more bids than the average number of bids in concurrent auctions,
8. A bidder typically offers a low initial price to attract legitimate bidders to the auction,
9. A bidder tends to bid early in the auction (less than 25% of the auction duration) to attract the attention of auction participants,
10. A bidder competes in many auctions but rarely wins all of them,
11. The duration of the auction.

This dataset is initially divided into two subsets: training and testing datasets. According to the proposed model, a portion of the training datasets is considered as validation data. In this study, 85% of the raw data is randomly selected as the training set, while the remaining 15% is chosen as the testing set. Additionally, 10% of the training data is randomly selected as validation datasets. To eliminate the effect of stochastic processes on model performance, each model is repeated 100 times. Performance indicators, including classification error rate and improvement of the reliability-based multilayer perceptron model compared to the classical accuracy-based version, are obtained as 0.005274, 0.021097, and 75%, respectively. The results from this case study indicate that the proposed reliability-based multilayer perceptron model can significantly improve the performance of the classical accuracy-based multilayer perceptron model. However, it is generally shown in the modeling literature that the characteristics of the data significantly affect model performance, and the type and extent of this impact vary across different models. Therefore, to eliminate the influence of data characteristics on model performance, the proposed and classical modeling processes previously described are repeated for four other benchmark datasets. The results obtained for each dataset are summarized in Table 2.

Table 2. Performance Metrics in the Proposed and Classical Multilayer Perceptron Models

Improvement Rate (%) Classification error	Accuracy-Based Model Classification error	Reliability-Based Model Classification error	Title
75	0.021097	0.005274	Shill Bidding
20	0.129310	0.103448	Audit Data
0.792602	0.168222	0.166889	Default of Credit Card Clients

-5.882353	0.212500	0.225000	Istanbul Stock Exchange
0.557769	0.370097	0.368033	Banking Marketing
18.093604			Average

The results of the proposed and classical perceptron models are compared in terms of quality, specifically the number and extent of superiority, as well as accuracy, i.e., the percentage improvement in generalization ability. The results obtained from implementing the proposed and classical models on the five benchmark datasets indicate that in 4 out of 5 cases (80% of the time), the proposed model outperforms the classical multilayer perceptron model in terms of the classification error rate evaluation metric. Quantitatively, these results clearly demonstrate the impact of reliability as an influential factor, in addition to accuracy, on the generalization ability of the multilayer perceptron model. Furthermore, the results indicate that when generalization is a concern, the relative importance of reliability is, on average, greater than that of accuracy. Similarly, regarding the improvement in accuracy results, the proposed model enhances the classical multilayer perceptron model by 18.094% in the classification error rate metric. These results indicate the superiority of the proposed model over the classical multilayer perceptron model. Based on these findings, it can be inferred that reliability is a more effective factor than accuracy in improving the generalizability of the model.

5. Conclusion and Recommendations

Financial predictions play a significant role in business planning and investment decision-making. Various classification methods are widely used for predictions in different financial fields. All methods are developed based on a common assumption that achieving more accurate results in the training dataset guarantees greater generalization power for the testing dataset. Although maximizing accuracy seems to be a logical and common procedure in classification models, it is not the only approach to achieve generalizable results. In fact, due to fluctuations and changes in various financial patterns, stable and reliable predictions are as important as precise predictions. Therefore, the reliability of results is another crucial factor that should be considered in modeling processes. However, this factor has been overlooked in the literature on modeling and classification. The classical multilayer perceptron model is one of the popular intelligent methods successfully used in a wide range of financial predictions. It is a model based on accuracy, like other conventional classification models. Accordingly, in this paper, a multilayer perceptron model that considers reliability in the modeling process is proposed based on maximizing the reliability of results. Additionally, to evaluate the performance of the proposed reliability-based multilayer perceptron model, it was implemented on five financial benchmark datasets from the UCI repository. The comparison of results based on the evaluation metric of improvement in classification error rate supports the superiority of the proposed reliability-based model over the classical multilayer perceptron in financial predictions. The results clearly indicate the importance of reliability as an effective factor on the quality of financial predictions. Therefore, the proposed reliability-based multilayer perceptron model can serve as a suitable alternative to the conventional multilayer perceptron in financial predictions. It is also suggested that the proposed model be applied in other practical fields and that the idea of reliability-based modeling be implemented on various models as future research recommendations.

References

Choudhary, P., & Thenmozhi, M. (2024). Fintech and financial sector: ADO analysis and future research agenda. *International Review of Financial Analysis, 93,* 103201. https://doi.org/10.1016/j.irfa.2024.103201

Dua, D., & Graff, C. (2019). UCI machine learning repository [http://archive. ics. uci. edu/ml]. Irvine, CA: University of California, School of Information and Computer Science. *IEEE transactions on pattern analysis and machine intelligence*, *1*(1), 1-29.

Etemadi, S., & Khashei, M. (2021). Multiple linear regression. *Measurement*, *186,* 110080. https://doi.org/10.1016/j.measurement.2021.110080

Etemadi, S., Khashei, M., & Tamizi, S. (2023). Reliability-based multi-layer perceptrons for classification and forecasting. *Information Sciences, 651,* 119716. https://doi.org/10.1016/j.ins.2023.119716.

Gicić, A., Đonko, D., & Subasi, A. (2023). Intelligent credit scoring using deep learning methods. *Concurrency and Computation: Practice and Experience, 35*(9), e7637. https://doi.org/10.1002/cpe.7637.

Golbayani, P., Florescu, I., & Chatterjee, R. (2020). A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. *The North American Journal of Economics and Finance, 54,* 101251. https://doi.org/10.1016/j.najef.2020.101251.

Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, *100*, 234–245. https://doi.org/10.1016/j.eswa.2018.01.037.

Khalili, N., & Rastegar, M. A. (2023). Optimal cost-sensitive credit scoring using a new hybrid performance metric. *Expert Systems with Applications, 213,* 119232. https://doi.org/10.1016/j.eswa.2022.119232.

Mabu, S., Obayashi, M., & Kuremoto, T. (2015). Ensemble learning of rule-based evolutionary algorithm using multi-layer perceptron for supporting decisions in stock trading problems. *Applied Soft Computing*, *36,* 357–367. https://doi.org/10.1016/j.asoc.2015.07.020

Moral-García, S., & Abellán, J. (2023). Improving the results in credit scoring by increasing diversity in ensembles of classifiers. *IEEE Access, 11,* 58451–58461. https://doi.org/10.1109/ACCESS.2023.3284137

Nazemi, A., Rezazadeh, H., Fabozzi, F. J., & Höchstötter, M. (2022). Deep learning for modeling the collection rate for third-party buyers. *International Journal of Forecasting*, *38*(10), 240–252. https://doi.org/10.1016/j.ijforecast.2021.03.013.

Shrivastava, S., Shukla, S., & Khare, N. (2024). Support vector machine with eagle loss function. *Expert Systems with Applications, 238,* 122168. https://doi.org/10.1016/j.eswa.2023.122168

Tsai, C. F., Sue, K. L., Hu, Y. H., & Chiu, A. (2021). Combining feature selection, instance selection, and ensemble classification techniques for improved financial distress prediction. *Journal of Business Research*, *130(2),* 200–209. https://doi.org/10.1016/j.jbusres.2021.03.018

Uthayakumar, J., Vengattaraman, T., & Dhavachelvan, P. (2020). Swarm intelligence based classification rule induction (CRI) framework for qualitative and quantitative approach: An application of bankruptcy prediction and credit risk analysis. *Journal of King Saud University-Computer and Information Sciences, 32*(6), 647–657. https://doi.org/10.1016/j.jksuci.2017.10.007.

Vrontos, S. D., Galakis, S., & Vrontos, I. D. (2021). Modeling and predicting US recessions using machine learning techniques. *International Journal of Forecasting*, *37*(2), 647–671. https://doi.org/10.1016/j.ijforecast.2020.08.005.

Zhang, D., & Lou, S. (2021). The application research of neural network and BP algorithm in stock price pattern classification and prediction. *Future Generation Computer Systems, 115*, 872–879. https://doi.org/10.1016/j.future.2020.10.009.

Zhao, Z., Xu, S., Kang, B. H., Kabir, M. M. J., Liu, Y., & Wasinger, R. (2015). Investigation and improvement of multi-layer perceptron neural networks for credit scoring. *Expert Systems with Applications, 42*(7), 3508–3516. https://doi.org/10.1016/j.eswa.2014.12.006.