

# ***Research on the Impact of Digitalization on Bank Performance Based on BP Neural Network***

**Minqiao Zheng<sup>1,a,\*</sup>**

<sup>1</sup>*School of Economics, Zhejiang University of Finance & Economics, Hangzhou, Zhejiang, 310018, China*

*a. zmqklein@zufe.edu.cn*

*\*corresponding author*

**Abstract:** With the promotion and development of digital finance, the degree of digitalization has an increasingly significant impact on the banking business, which cannot be ignored in the operation of commercial banks. Therefore, establishing a digitalization bank profitability system model is of great significance for the sustainable development of banks and digital finance. This article selects 6 Chinese state-owned banks from 57 listed banks in the A+H stock market as samples, constructs a digital indicator system and evaluation model for banks through the literature review method, and uses simulation training to explore the effectiveness of BP neural network in evaluating the impact of digitalization on bank profitability. The aim is to provide a reference for the evaluation of financial statement information of other banks and other factors affecting bank digitalization. The innovation of the article lies in the innovative use of the entropy weight coefficient of variation method to establish a digital evaluation system model, which utilizes the data itself to objectively establish an evaluation system, incorporate the established digital system into the black box model of the BP neural network model to study the impact of changes in digitalization on banking operations.

**Keywords:** Bank profitability, BP neural network, Entropy weight coefficient of variation method, The impact of bank digitization.

## **1. Introduction**

With the modernization of Chinese society and the continuous improvement of the scientific and technological levels, China has officially entered the era of a digital economy. The development requirements for digitization and intelligence (i.e. "digitalization") have risen to the national strategic level, which is clearly stated in China's "14th Five Year Plan for Digital Economy Development". The digitalization level of the banking business, as an important link in the development of the banking industry, plays a crucial role in improving bank performance and economic benefits. In depth, research and promotion of digital transformation in the banking industry can not only enhance bank competitiveness, improve risk management capabilities, and optimize customer experience, but also play an important role in the stable development of the financial market [1].

However, to gain a deeper understanding of the digitalization level of the banking industry, it is necessary to consider the diversity and complexity of banking operations. The banking business, in a broad sense, includes many parts and covers a wide range of fields. According to the complexity of

business and the degree of dependence on branches, banking business can be divided into two parts: traditional business and complex business. Traditional businesses mainly rely on branch networks and business volume, while complex businesses have relatively low demand for branch networks and belong to high-tech and high-profit business areas. At the same time, the banking business is mainly divided into three categories: liability business, asset business, and intermediary business, each of which has different digital needs and challenges. Therefore, to evaluate the level of digitalization in the banking industry, it needs to comprehensively consider various factors and evaluate them based on specific business situations and actively respond to the challenges brought by digitalization to ensure the sustainable development of its business.

**Digitization level.** This indicator can reflect the level of development of banks in digital finance. According to the literature review, Yang Leixin's research used factor analysis and principal component analysis to evaluate the degree of digital transformation of banks [2]. However, Wei Shun's research needs to comprehensively consider various factors, including technological infrastructure, data application information, etc., and collect relevant data, such as the comprehensive evaluation of the bank's digital investment, digital business proportion, etc. [3, 4].

For the convenience of research and the need for quantification, it is necessary to select some indicators to measure the effectiveness of digital transformation. Such as return on assets, return on equity, cost-income ratio, net interest margin, non-performing loan ratio (NPL), deposit growth rate, and loan growth rate, and standardize them. These indicators can help analyze the profitability, risk control ability, business growth ability, and operational efficiency of the bank, thereby reflecting the status and performance of the bank's business [5]. This article selects five standard return on equity, return on total assets, and capital adequacy ratios that can reflect profitability and safety levels to reflect the profitability of banks [6].

At present, it has been found that there has been no discussion on the combination of digitalization and bank operating capabilities in the field of research. There is a lack of research based on BP neural networks in the field of listed banks in China [7]. Therefore, this article uses the BP neural network combined with the data characteristics of listed banks, selects financial indicators as comprehensively as possible to construct a bank digitalization and profitability model, explores the effectiveness of digitalization in evaluating bank profitability, and provides relevant decision-making suggestions for listed banks to further develop digitalization. By collecting data, such as selecting indicators to measure the effectiveness of digital transformation [8]. Such as research and development intensity, number of mobile banking customers, patent growth ratio, and proportion of scientific researchers, standardizing them, multiple indicator variables are simplified into comprehensive indices to reflect the effectiveness of digital transformation in banks.

## **2. Method and Data**

### **2.1. Method**

The main idea of the entropy weight coefficient of the variation method is to combine the weights obtained from different weighting methods to form new weights, because for weighting methods, each method is weighted according to a certain rule, and it cannot be said that one method is better than another [9]. There are many ways to combine them for weighting, and the methods in this reference are combined for weighting. The process diagram is shown in Figure 1.

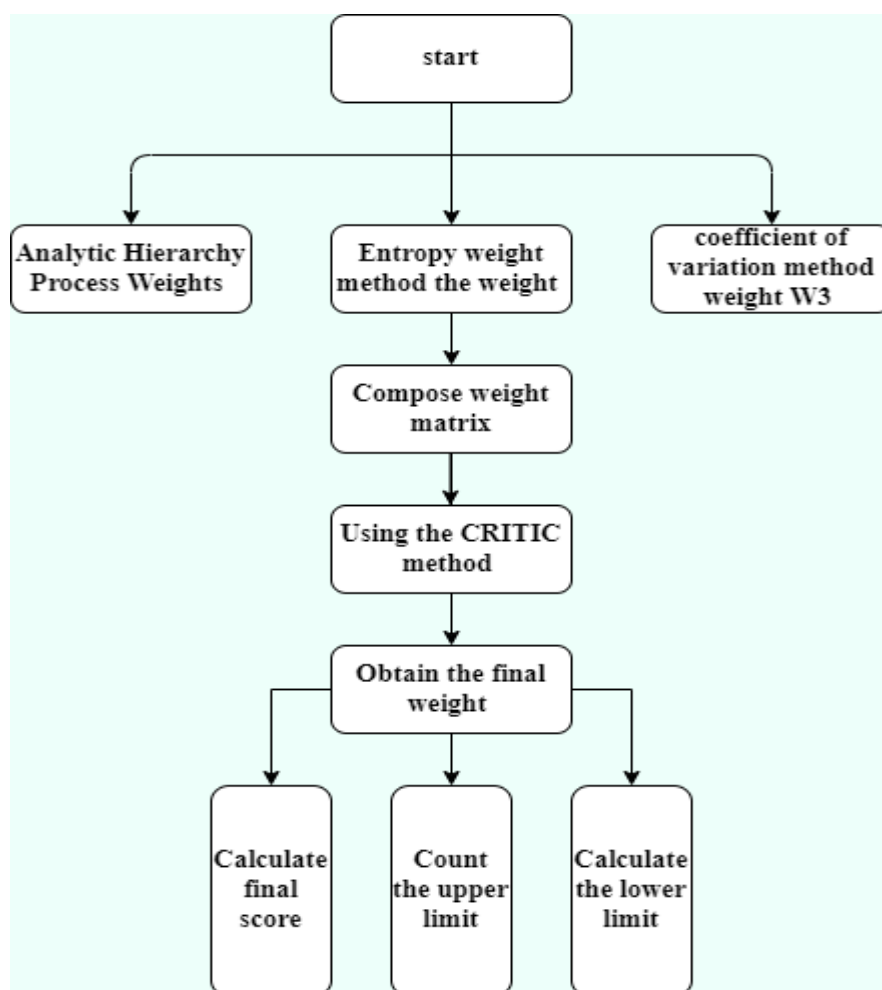


Figure 1: Entropy weight coefficient of variation method (Photo/Picture credit: Original).

According to the research of Zhou Wenhui et al. on digital empowerment, it can be seen that digitalization can indeed increase profits by reducing costs and opening up. The specific explanation of how digitalization reduces costs is as follows [10].

In terms of operating costs, by migrating simple businesses online, operating costs can be greatly reduced.

In terms of human capital, traditional banking services typically require a significant amount of human resources, and through digital transformation, the number of tellers and costs associated with window staff can be reduced.

Optimized processes provide support by introducing intelligent devices such as self-service terminals, which can further reduce the demand for tellers. With the advancement of financial technology, commercial banks can achieve process optimization by directly transferring front-end business to the backend for processing through electronic device distribution systems, thereby reducing unnecessary intermediate links.

BP neural network is a multi-layer feedforward artificial network proposed by scientists such as Rumelhard and McClland in 1986. It consists of an input layer, a hidden layer, and an output layer, with connections between neurons in each layer and neighboring neurons, while there is no connection between neurons in the same layer. The connected neurons all have corresponding weights. This type of network is an autonomous learning algorithm. In neural networks, the mapping and weight equations between layers do not need to be set in advance. Instead, the weights and

thresholds between layers can be continuously adjusted through the autonomous learning of BP neural networks, ultimately achieving the best results.

## 2.2. Data

The first important factor is the profitability level as the main variable and the degree of digitization as the dependent variable. Based on the Digital Inclusive Finance Index constructed by the Digital Finance Research Center of Peking University and the measurement index system released by the G20 Group, starting from the perspective of banks, the final selection includes research and development intensity, patent growth ratio, proportion of scientific researchers, and number of mobile banking customers. There are four dimensions used as evaluation criteria for the degree of digitalization, among which R&D intensity refers to the proportion of financial technology investment to operating income. The second important factor is that this article selects five standards that can reflect profitability and safety levels: return on equity (ROE), asset growth rate, revenue growth rate, capital adequacy ratio, and net interest margin [11].

In terms of data selection for determining the degree of digitization, this article obtained data by consulting the company's published annual report and reading the Wind database. Regarding the number of annual patents, there is no distinction between the status of valid patents and pending patents, and only the number of patents applied for by each bank in the current year is taken as the indicator. The number of mobile banking customers is based on the number of people who use and download mobile banking on an annual basis, as the five major state-owned industries first disclosed their investment in fintech in their 2019 annual reports. So the selected digital data is offline until 2019. Because the annual report of Postal Bank does not mention the number of researchers engaged in financial technology, Postal Bank's data is not included in the construction of the digital system.

## 3. Empirical Research

### 3.1. Digital Index Calculation and Weight Analysis

According to the data from four dimensions of digitalization in 2022, the digitalization indices of five state-owned banks calculated using the entropy weight coefficient of variation method are shown in Table 1 [12].

Table 1: Digitalization level of banks

Bank name	Digitization level
Industrial and Commercial Bank of China	100
Bank of China	73.2538
Construction Bank'	66.1539
Bank of Communications	61.9513
Agricultural Bank of China	44.0888

In terms of building digital indicators for banks, the three dimensions of research and development intensity, patent growth ratio, and proportion of scientific researchers will affect the evaluation of their level of digitalization. In addition, when it comes to the breadth of impact that enterprises have on society, the online business of commercial banks is also an important aspect of evaluation. The number of mobile banking customers reflects the user refinement and digital operation of some excellent banks, continuously improving the data-driven intelligent service experience, which is an important part of digitalization.

Through consistency testing and normalization, the weights of each indicator under the obtainable degree and digitalization degree are finally obtained, as shown in Table 2

Table 2: Digital dimension weight

Digital dimension	Weight
R&D intensity	30.92%
Number of mobile banking customers	22.56%
Patent growth ratio	21.93%
The proportion of scientific researchers	24.58%

## 3.2. Neural Network Model Construction and Simulation

### 3.2.1. BP Neural Network Model Settings

There is currently no unified method for setting hidden layer nodes in BP neural networks, and most scholars use empirical and trial-and-error methods to determine them [13]. This article found through continuous experiments that the model achieved the highest training accuracy when the number of hidden layer nodes was set to 7, which is generally consistent with the screening range of factor analysis before. Therefore, the number of hidden layer nodes was determined to be 7. The number of output layer nodes is determined by the number of output indicators. The fundamental purpose of this study is to study the evaluation of digital banking profitability, and the profitability indicator is the output indicator. Therefore, the number of output layer nodes is determined to be 1.

This article uses a three-layer BP neural network, including an input layer, a hidden layer, and an output layer. The number of nodes in the input layer is determined by the number of selected evaluation indicators, which evaluate the profitability of banking institutions. In the BP neural network, the input and output layers are determined by the problems they need to solve. For studying the profitability evaluation of banking institutions through digitalization, this article selects 5 indicators, so the number of input layer nodes is set to 5. At present, there is no unified standard for setting hidden layer nodes, which is generally determined through trial and error. In this study, after multiple trainings, it was found that the model achieved the highest training accuracy when the number of hidden layer nodes was set to 7. Therefore, the number of hidden layer nodes was determined to be 7. The number of output layer nodes is determined by the number of output indicators. The fundamental purpose of this study is to study the evaluation of bank profitability by digitalization. The profitability indicator is the output indicator. Therefore, the number of output layer nodes is determined to be 1.

After setting the number of layers and structural parameters of the BP neural network, it is also necessary to set the training parameters. This article uses the default training method, Levenberg Marquardt, with Trainlm as the training function. The transfer function from the input layer to the hidden layer is the tansig function, while the transfer function from the hidden layer to the output layer is the purelin function. Meanwhile, set the maximum number of iterations to 1000, the target error to 0.001, and the learning rate to 0.01.

### 3.2.2. Setting up the Training and Testing Sets

This article uses Matlab software for model simulation testing. Before conducting model simulation testing, it is necessary to first set up the training and testing sets using Matlab software. This study selected the profit data of six state-owned banks listed on A-shares and H-shares over five years, with a total of 30 data points. 20 data points were randomly selected as the training set, 5 as the testing set, and the other 5 as the validation set. After setting the parameters, train and simulate the model. Figure

2 shows the mean square error graph, where blue represents the mean square error of the training set, green represents the mean square error of the validation set, and red represents the mean square error of the test set. In the figure, the Best dashed line indicates that when the model is iterated 6 times, as shown in Figure 2, the training error reaches the preset target value, and the training is completed

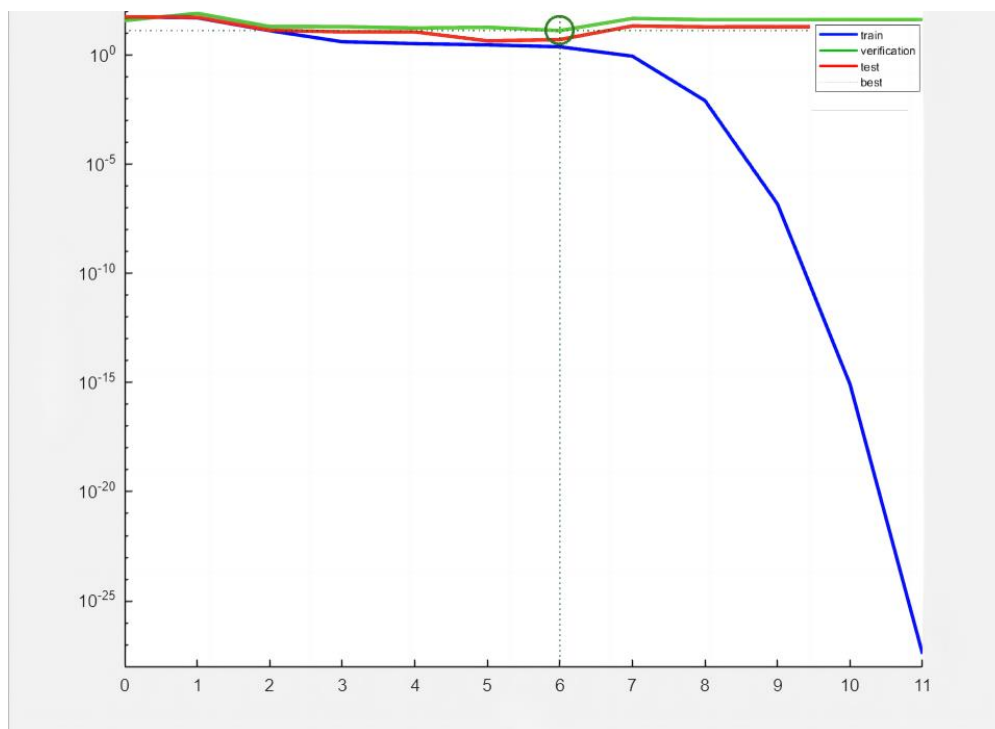


Figure 2: Mean squared error plot (Photo/Picture credit: Original).

### 3.2.3. Model Training and Simulation

During the training process of the BP neural network, samples are automatically divided into three categories: training data, validation data, and test data. When conducting simulation training, only the training data is used to train the model. Through model simulation training, obtain the discrete data distribution of training samples, validation samples, test samples, and all samples. As shown in Figure 3, the correlation coefficient of the training sample is 0.91593, the correlation coefficient of the validation sample is 0.75146, the correlation coefficient of the test sample is 0.95715, and the correlation coefficient of all samples is 0.82287. Because bank digital information was only released in 2018, this study can only fit the degree of digitization over five years. Due to the lack of relevant data available for the query, it can only consider from a statistical perspective that samples with a correlation coefficient greater than 0.8000 meet the requirements. The training results of this article show a high degree of fit and strong regression characteristics, and the training situation is good.



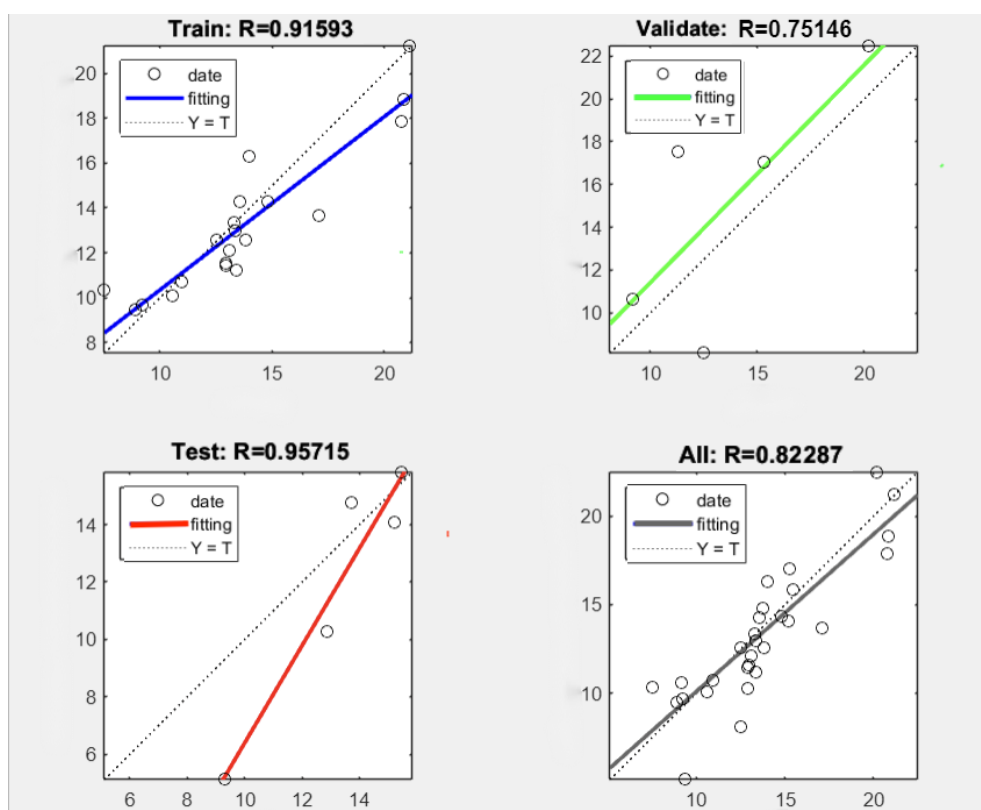


Figure 3: Training results (Photo/Picture credit: Original).

### 3.3. Key Findings and Areas for Future Research

In today's rapidly developing financial technology, banks are placing increasing importance on digitalization, and the factors that need to be considered in the research of establishing evaluative neural networks are becoming increasingly complex and diverse [14]. Due to limitations in research time and conditions, there are still some shortcomings in this study. Firstly, this study did not consider the classification of different banks for judgment. Due to the different focuses of business operations in different banks, the impact of different indicators selected on profitability may vary. Therefore, there may be incomplete selection in the construction of this model. Secondly, the selection of independent variables in this study is based on the existing financial indicators of listed banks, but almost all of these indicators exist in the balance sheet, and there is no good analysis and citation of other data on profitability, which may result in omitted variable bias. Therefore, in future research, in addition to considering the financial indicators used in this article, more data related to profitability, such as EVA, can be considered.

## 4. Conclusion

This article studies the impact of digitalization level on performance evaluation of listed banks. By analyzing the digitalization level of listed banks in China, appropriate indicators were selected, and a performance evaluation system for listed commercial banks was constructed. An appropriate model was used to evaluate the performance of commercial banks. The research conclusions are summarized as follows:

1. This article uses the literature review method to screen suitable indicators and uses the combination method of entropy weight coefficient of variation in the objective weighting method to determine the weights of each evaluation indicator and the digitalization level of each bank. It

overcomes the subjective shortcomings of methods such as expert scoring and balanced scorecard, and can effectively measure the correlation degree between various indicators. The evaluation shows that the digital capability indicators, including research and development intensity, number of mobile banking customers, patent growth ratio, and proportion of scientific researchers, account for 30.92%, 22.56%, 21.93%, and 24.58%, respectively.

2. In this study, after understanding the theoretical basis of BP neural network technology, it constructed a BP neural network performance evaluation model. By training this model with input sample data and continuously adjusting weights and parameters, it successfully achieved the preset target error. It has validated the effectiveness of the model and found that the error between the simulated output and the expected output of the selected test samples is very small, which meets the requirements of the target error. Therefore, it can be concluded that the constructed network model has good generalization ability and can effectively evaluate the performance of listed banks. However, due to limitations in data sources, it did not include qualitative indicators in the selection process. It looks forward to improving this in future research to better achieve the goals of bank performance evaluation.

From this, it can be seen that the evaluation digitalization bank profitability model built using the BP neural network can evaluate the profitability of banks, that is, using the BP neural network to evaluate digitalization for the profitability of banking institutions is effective and feasible. From this, it can be seen that combining the objective weighting method with the BP neural network can effectively evaluate the impact of digitalization on the profitability of listed banks, providing a reference for the evaluation of financial statement information of other banks and other factors affecting bank digitalization.

## References

- [1] Congying, S., Yue, P., Ying, L. (2023). *Digital Transformation and Manufacturing Investment: An Empirical Analysis Based on Listed Company Data*. *China Collective Economy*, (27): 33-35
- [2] Jian, L. (2019). *Analysis of the Reasons and Strategies for the Financial Technology Transformation of Commercial Banks in China*. *Financial Economy*, (18): 48-49
- [3] Leixin, Y. (2024). *Research on the Digital Transformation of China Merchants Bank under the Background of Financial Technology*. *Shanxi University of Finance and Economics*, 2024.
- [4] Wei, S. (2024). *Research on the Impact of Digital Inclusive Finance on the Profitability of Commercial Banks*. *Shanxi University of Finance and Economics*.
- [5] Bowen, Q. (2021). *Research on Financial Risk Warning of Pharmaceutical Listed Companies Based on Neural Network Models*. *Jiangxi University of Finance and Economics*.
- [6] Rong, L. (2020). *Overview of the Basic Principles of Artificial Neural Networks*. *Computer Products and Circulation*, (06).
- [7] Feng, G. Jingyi, W., Fang, W., et al. (2020). *Measuring the Development of Digital Inclusive Finance in China: Index Compilation and Spatial Characteristics*. *Economics (Quarterly)*, 19 (04): 1401-1418.
- [8] Xinmiao, H., Huiying, D., Jia, W. (2024). *Strategy for constructing a BP neural network warning model for operational risks of listed companies based on financial indicators*. *Modern Business*, 2024 (01): 189-192.
- [9] Xuanli, X., Shihui, W. (2022). *Digital Transformation of Commercial Banks in China: Measurement, Process, and Impact*. *Economics (Quarterly)*, 22(06): 1937-1956.
- [10] Zhilin, W., Hui, S. (2020). *Determinants of Chinese Listed Commercial Banks' Profitability—Under A GMM Approach* *Journal of Economics Business and Management*, 2020, 11(5): 691-701.
- [11] Jinzhi, W. (2024). *Pricing Research on A-share Market Based on Multi Factor Model*. *Harbin Normal University*.
- [12] Ladi, W., Jiangxu, W. (2020). *Financial Risk Assessment of Real Estate Listed Companies*. *Friends of Accounting*, (10): 31-36.
- [13] Lijuan, W., Zeyuan, M. (2023). *Research on Green Credit Risk Evaluation of Banking Institutions Based on BP Neural Network*. *Times Economics and Trade*, 20 (11): 72-76.
- [14] Yiping, H., Zhuo, H. (2018). *The Development of Digital Finance in China: Present and Future*. *Economics (Quarterly)*, (04).