

Research on Industrial Clusters and Regional Economic Impact in Underdeveloped Regions Based on ANN-MLP algorithm

Yi Song^{1,#}, Ye Ye^{2,*,#}, Yifei Wang^{3,#}

¹ School of management, Hefei University of Technology, Hefei, China, 230009

² School of Cultural Tourism and Geography, Guangdong University of Finance and Economics, Guangzhou, China, 510320

³ School of Economics and Commerce, Hunan University, Changsha, China, 410082

* Corresponding Author Email: 2783891479@qq.com

These authors contributed equally.

Abstract. The industrial cluster is an important source of regional competitiveness and a powerful carrier of regional economic growth. This paper firstly constructs the index system of regional industrial clusters and regional economy, uses artificial neural network-multi-layer perceptron (ANN-MLP) algorithm for contribution analysis, and uses mean square error (MSE) and coefficient of determination (R2) as the algorithm test index, and finally takes Jiangxi province as an example to empirically study the impact of industrial clusters and regional economy in less developed regions. The research results show that the model has high accuracy, when the parameter MSE is 2259314 and R2 is 0.969. Through this research, this paper can help managers make correct decisions and promote economic development. Industrial clusters can drive economic growth, and regional economies can bring more benefits to industrial clusters.

Keywords: ANN-MLP algorithm, regional industrial clusters, regional economy.

1. Introduction

In recent years, the role of regional industrial clusters in regional development and industrial upgrading has become more and more obvious. Industrial clusters are a universal spatial economic phenomenon that can reduce costs through the benefits of economies of scope and scale, and thus gain additional benefits. In addition, industrial clusters can also enhance the overall comprehensive competitiveness of the cluster area and facilitate knowledge spillover and technology diffusion. By studying the economic benefits of industrial clusters in less developed regions, this paper can grasp the problems of industrial economy in less developed regions more accurately, make industrial development plans less developed regions regions, further improve the market environment, optimize the allocation of innovative resources, provide strong policy guarantee for industrial development, and maximize the contribution of cluster effect regional economic development.

Shen [1] pointed out that improving the innovation capacity and competitiveness of industrial clusters can promote the sustainable development of regional economy. Fan [2] pointed out that localities should try to take advantage of the localized characteristics of increasing returns to scale in manufacturing or service industries to improve regional labor productivity through regional industrial agglomeration. Bian and Liu [3] showed that the coupling coordination between industrial agglomeration level and regional economic development has a large overall gap. Zheng [4] pointed out that industrial cluster development cannot only stand on the status quo of a certain industry to analyze and view the problem, but needs to aim at the foresight of technology and grasp the direction of industrial change. Deng [5] used principal component analysis to quantify the economic effects of industrial clusters in Jilin Province, and the results showed that the cluster effects of major industrial clusters in Jilin Province showed a steady growth in general. Gao and Yao [6] studied the construction of government-led regional branding based on the perspective of industrial clusters, and proposed the corresponding countermeasures of government "guidance as the main focus and intervention as the

supplement". Zhou [7] constructed a panel data regression model of economic effects of high-tech industry clusters and economic growth from two dimensions: economic development level and sustained economic growth, the conclusion is that if the regional economy wants to grow sustainably, it needs to expand the cluster size and increase the innovation input.

Although previous studies have done a lot of exploration on the relationship between the two, the research for data mining needs to be deepened urgently. Therefore, based on previous studies, this paper firstly constructs an indicator system of industrial clusters and regional economy in less developed regions, secondly constructs an artificial neural network-multi-layer perceptron (ANN-MLP) algorithm, uses mean square error (MSE) and coefficient of determination (R^2) as algorithm test indicators, and selects the impact of industrial clusters and regional economy in Jiangxi Province, China for empirical analysis. It is conducive to the formulation, government guidance and supervision of regional industrial development plans, the optimization of innovation resource allocation, the improvement of the comprehensive competitiveness of cluster areas, the upgrading of industrial transformation and regional synergistic development, and the cluster effect.

2. Construction of Regional Industrial Clusters and Regional Economic Indicators System in less developed regions Areas

2.1. Regional industrial clusters

Geographical concentration of industrial clusters is a universal spatial economic phenomenon. It is a cooperative enterprise in which competitive enterprises in an industry come together and interact with enterprises, professional suppliers, service providers, relevant industrial manufacturers and relevant institutions (such as universities, research institutions, standard-setting bodies, trade associations, etc.) in a particular region. Industrial clusters generate spillover effects, competition effects, agglomeration effects, market effects and division of labor effects, thus contributing to the development of regional economies. This paper intend to adopt the regional industrial cluster output value (Y) to quantify the development of regional industrial clusters.

2.2. Regional Economy

A regional economy is also called a "regional economy". It refers to the part of the national economy that is distributed in each administrative region. It is formed as a result of the geographical division of labor. In the long-term socio-economic activities, due to historical, geographical, political, economic, religious and other factors, a number of economically connected settlements have gradually formed unique economic zones. The regional economy is a microcosm of the national economy and is comprehensive and regional in nature.

The influence of regional industrial clusters on the regional economy is mainly reflected in three aspects: (1) integration. According to the original industry, resource endowment and other factors, the characteristic industries of regional development should be determined so as to promote local economic development. (2) Cost reduction. Industrial clusters can enhance the competitiveness of the region's industries, giving them a stronger voice and bargaining power in the acquisition of raw materials and the pricing process of commodities, reducing production costs and transaction costs. (3) Promoting technological innovation. The similarity of products and technologies among enterprises in industrial clusters makes it possible for them to imitate and learn from each other. In order for enterprises to gain a competitive advantage, they must continuously innovate technology to provide impetus for regional economic development.

Based on the above, this paper combines the literature to construct a regional industry cluster and regional economy index system as shown in Table 1.

Table 1. Industrial clusters and regional economic indicators

Primary indicators	Secondary indicators	Indicator units	References
Regional Industry Clusters	Regional industry cluster output value (Y)	hundred million yuan	[3]
	GDP (X ₁)	billion yuan	[3]
	Imports and Exports (X ₂)	million yuan	[8]
	Fixed Asset Investment (X ₃)	million yuan	[8]
Regional Economy	Per capita disposable income (X ₄)	yuan	[2]
	RMB deposit balance of financial institutions (X ₅)	million yuan	[2]
	Regional finance General public budget revenue (X ₆)	million yuan	[6]
	General public budget expenditures (X ₇)	million yuan	[6]
Regional Science and Technology	Number of Patent Applications (X ₈)	Pieces	[7]
	Number of patent licenses (X ₉)	Pieces	[7]
	Number of important scientific and technological achievements (X ₁₀)	items	[7]

3. ANN-BRF

3.1. ANN Method

Artificial neural network (ANN) is an information processing system based on imitating the structure and function of brain neural network. ANN shows obvious advantages in fuzzy data, nonlinear data and random data processing, and is more suitable for systems with fuzzy information, complex structure and large scale. ANN has the ability of self-learning, associative storage and high-speed searching for optimal solution. Artificial neural network computer can provide economic forecast, market forecast, benefit forecast; it can also find an optimal solution to a complex problem. Using a feedback artificial neural network designed for a problem, the high-speed computing power of the computer can be used to find the optimal solution quickly. The ANN diagram constructed in this paper is shown in Figure 1:

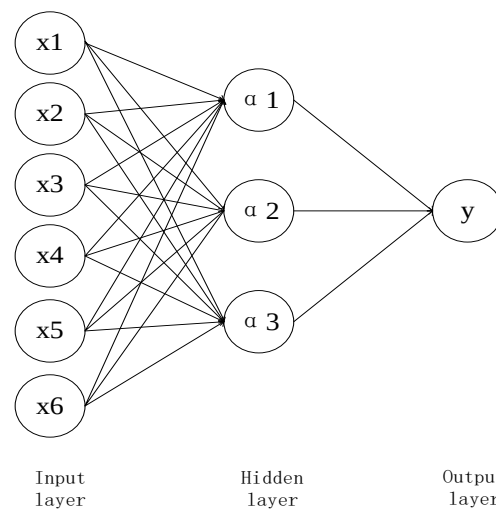


Figure 1. ANN diagram

3.2. MLP model construction

A multilayer perceptron (MLP) is called a multilayer feedforward neural network model. MLP is a forward-structured ANN artificial neural network. MLP can deal with nonlinear separable problems. Based on the biological neuron model, the basic structure of the multi-layer perceptron MLP can be

obtained. The most typical MLP includes three layers: input layer, hidden layer and output layer. The different layers of the MLP neural network are fully connected. MLP can be viewed as a directed graph consisting of multiple node layers, each of which is fully connected to the next layer. In addition to the input node, each node is a neuron with a nonlinear activation function. The supervised learning method of BP back propagation algorithm is used to train MLP. The advantages of MLP are high nonlinear global effect, good fault tolerance, very strong adaptive and self-learning function.

The activation function used in the MLP model is $g(h) = \sigma(h) = \frac{1}{1+e^{-h}}$. First, the input data is weighted and summed $h_j = \sum_{i=0}^M w_{ij} x_j$, then it is substituted into the activation function as an independent variable $a_j = g(h_j) = g(\sum_{i=0}^M w_{ij} x_{ij})$, the final result is $y = a_k = g(h_k) = g(\sum_{i=0}^M w_{ij} x_{ij})$, h_j represents the weighted sum of all inputs for the current node; a_j represents the output value of the hidden layer neuron, w is the weight, and x is the input value. $w_{0j}x_{0j}$ represents offset node, $a_j = x_{jk}$ is the output value of the current layer neuron, equal to the input value of the next layer of neurons; represents the input weighted sum of the output layer neuron k .

4. Setting test indicators

Before conducting empirical analysis, the first thing to do is to check the quality of the data set and have a basic understanding of the data in order to perform data processing. At the same time, the index test is helpful to determine the accuracy of the model, thus ensuring the effectiveness of the algorithm and the confidence of the results. The determination coefficient (R^2) and mean square error (MSE) were used to test the accuracy of the model.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

$$R^2 = \frac{(y - \bar{y})^2 - (y - \hat{y})^2}{(y - \bar{y})^2} \quad (2)$$

MSE is the performance function of the network, the mean square error of the network, such as n pairs of input and output data, each pair is $[P_i, T_i]$, $i = 1, 2, \dots, n$. MSE is the expected value of the square of the difference between the predicted value and the true value. The smaller the MSE value, the better the accuracy of the prediction model in describing the experimental data. R^2 refers to the goodness of fit, which can quantify the relationship between the model response variable and the dependent variable. R^2 is between 0 and 1, and the closer to 1, the better the regression fitting effect.

5. Empirical analysis

5.1. Research object selection and data collection

Underdeveloped areas refer to areas with a certain gap between the level of economic development and developed areas, the level of scientific and technological innovation is not advanced, and the development of productive forces is unbalanced but has development potential. This paper selects China's Jiangxi Province as the research object of underdeveloped regions, Jiangxi Province is located in the southeast of China, the capital of Nanchang. In 2021, the GDP of Jiangxi Province will be 29679.7 billion yuan, and the GDP of Jiangxi Province ranks 16 th in the country. There are 6669 high-tech enterprises in the province, and the comprehensive scientific and technological innovation level index is also ranked 16 th in the country. The province 's economic and social productivity development is still unbalanced, urban and rural development is unbalanced, and regional construction is unbalanced. However, in recent years, Jiangxi Province has great development potential in terms of GDP, scientific and technological innovation level and other indicators. Therefore, Jiangxi Province is selected as the research object of underdeveloped areas.

Based on the four databases of National Bureau of Statistics, Jiangxi Provincial Bureau of Statistics, Jiangxi Statistical Yearbook, Jiangxi Statistical Bulletin of National Economic and Social Development and EPS database, this paper collects and sorts out the data of Jiangxi Province in the past 20 years from 2002 to 2021 based on the index system constructed in this paper. Among them, there are two missing values in per capita disposable income index and one missing value in patent application index. In this paper, the average processing of adjacent points and the FORCAST prediction of Excel are used to fill and predict the missing values of the two variables respectively. This paper describes the indicators statistics are shown in Table 2:

Table 2. Indicators Descriptive Statistics

Indicators	Min	Max	Mean	Mid
X ₁ (billion yuan)	2450.5	29679.7	12980.5	13554.0
X ₂ (million yuan)	1267519.0	49784421.0	19239727.1	21965651.0
X ₃ (million yuan)	9246027.0	320057132.0	124299735.3	118122053.0
X ₄ (yuan)	2002.0	30610.0	14160.5	11434.0
X ₅ (million yuan)	17066270.0	447560405.0	178035996.3	181493099.0
X ₆ (million yuan)	1405457.0	28123369.0	12991926.5	14966149.0
X ₇ (million yuan)	3413843.0	67785284.0	30296937.1	32447629.0
X ₈ (number)	2037.0	124740.0	34058.0	14698.0
X ₉ (number)	1044.0	97372.0	21677.0	8978.0
X ₁₀ (number)	71.0	145.0	112.0	108.0
Y (hundred million yuan)	316.0	29949.0	9076.0	5718.0

5.2. Accuracy calculation of ANN

Based on SPSS 26.0 software, this paper uses neural network algorithm to analyze the 20-year data of Jiangxi Province from 2002 to 2021. The parameter range of the number of hidden layer neurons is set from 6 to 15, and MSE and R^2 are used as discriminant indicators to calculate the model fitting degree. The accuracy results are shown in Figure 2:

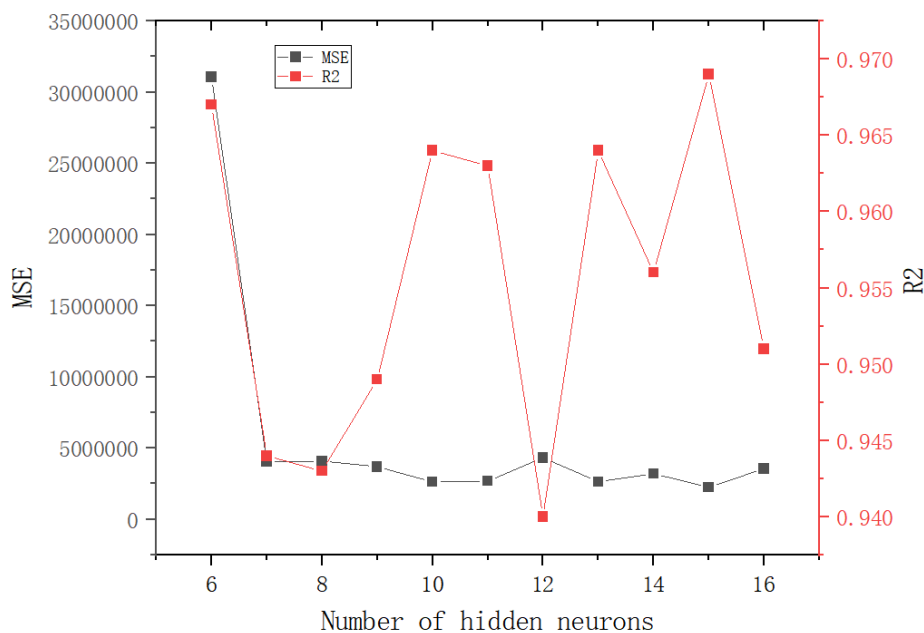


Figure 2. Accuracy results

After the above training, when the number of hidden neurons is 15, the highest accuracy. The square sum error of the model training is 0.01, the relative error is 0.002, the square sum error of the test is 0.03, and the relative error is 0.009. Among them, MSE is the smallest, reaching 2259314, R^2

is the largest, reaching 0.969, indicating that the model converges and the confidence of the results is high.

5.3. Contribution analysis

Based on 5.2 training, it is found that when the parameters X1 is 0.09, X2 is 0.116, X3 is 0.148, X4 is 0.185, X5 is 0.06, X6 is 0.039, X7 is 0.03, X8 is 0.079, X9 is 0.233 and X10 is 0.020, the results are the best. Contribution analysis is performed to obtain the normalized importance results of each parameter, as shown in Figure 3:

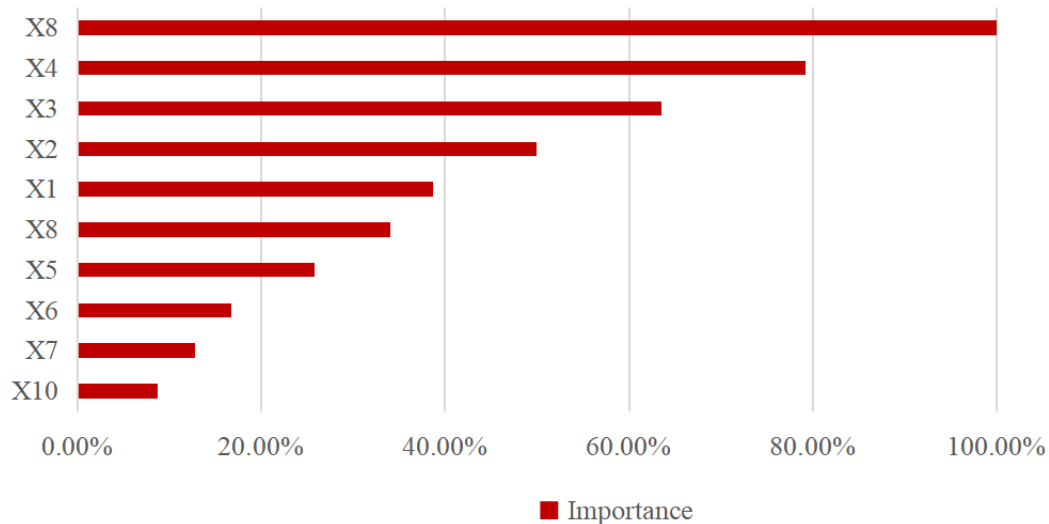


Figure 3. Normalization importance

As shown in Figure 3, the number of patent authorizations is of the highest importance to industrial clusters, reaching 100% after normalization. The importance of per capita disposable income, fixed asset investment and import and export to industrial clusters is second, and the importance degree is greater than or equal to 50%, which is 79.2%, 63.5% and 50% respectively. The importance of general public budget revenue, general public budget expenditure and the number of important scientific and technological achievements to industrial clusters is less than 20%.

It can be concluded that regional science and technology and regional finance have a high and positive impact on industrial clusters [9], while regional finance has a low impact on industrial clusters. Technological innovation is the productivity of development. The rise of regional technological level can improve the synergy between interrelated enterprises and overcome technical obstacles, thus promoting industrial clusters. Enterprises in industrial clusters influence and connect with each other, and each factory and company has interrelated interests [10]. When the level of science and technology rises, these enterprises get cluster innovation and technological progress. Industrial clusters can coordinate and integrate innovative resources, reduce costs, achieve innovative economies of scale and increase the output value of industrial clusters. Regional finance has played a driving role in the development of industrial clusters. The improvement of regional financial level can expand the supply of financing and provide diversified funds to support the development of industrial clusters. The involvement of financial capital can effectively expand the scale of clusters, optimize the industrial structure of clusters, stimulate the vitality of industrial upgrading, and enhance the profitability of clusters [11].

6. Conclusion

Based on the review of relevant literature, on the basis of previous studies, this paper constructs the index system of industrial clusters and regional economy in underdeveloped regions. By constructing ANN-MLP neural network, MSE and R^2 are set as test indexes to measure the accuracy of the algorithm. Finally, based on the calculation of SPSS software, the influence of industrial

clusters and regional economy in Jiangxi Province of China is selected for empirical analysis. The main conclusions are as follows:

(1) When the number of hidden neurons trained in this paper is 15, the accuracy is the highest. At this time, MSE is 2259314 and R2 is 0.969, indicating that ANN-MLP is more suitable for this research object.

(2) The three factors of patent authorization (importance is 0.233), per capita disposable income (importance is 0.185), and fixed asset investment (importance is 0.148) have the greatest impact. The number of important scientific and technological achievements has the least impact.

In summary, this paper predicts the impact of industrial clusters and regional economy in Jiangxi Province of China. The results show that the model has high accuracy and can help managers make correct decisions to promote economic development. Jiangxi Province of China can promote regional economic growth from three aspects: patent authorization, per capita disposable income and fixed asset investment. By stimulating innovation vitality, supporting and encouraging patent inventions, increasing per capita disposable income, and increasing investment in fixed assets, industrial clusters can better stimulate economic growth. Regional economy can also bring more benefits to industrial clusters.

References

- [1] Shen Zhengping, Liu Haiyang, Jiang Tao. Exploration of industrial clusters and regional economic development [J]. China Soft Science, 2004 (02): 120-124.
- [2] Fan Jianyong. Industrial agglomeration and labor productivity differences between regions [J]. Economic Research, 2006(11): 72-81.
- [3] Bian Henran, Liu Yi. Study on the coordination of industrial clusters and regional economic coupling in Guangdong-Hong Kong-Macao Greater Bay Area [J]. Technology Economics and Management Research, 2022, (07): 124-128.
- [4] Zheng Xiyu. Research on the promotion role of industrial cluster building on regional economic development [C] // Docking to Beijing and Tianjin--Industrial Development Education Proceedings. 2022: 85-98. DOI: 10.26914/c.cnkihy.2022.012380.
- [5] Deng Xiaobei. Study on the economic effects of industrial clusters in Jilin Province [D]. Northeast Normal University, 2021. DOI: 10.27011/d.cnki.gdbsu.2021.000590.
- [6] Gao Shiyong, Yao Jiawan, Ouyang Youquan. Research on the construction of government-led regional industries based on industrial clusters [J]. Economic Geography, 2015, 35(04): 108-113. DOI: 10.15957/j.cnki.jjdl.2015.04.015.
- [7] Zhou WJ. Research on the economic growth effect of high-tech industry clusters [D]. Harbin Engineering University, 2020. DOI: 10.27060/d.cnki.ghbcu.2020.001526.
- [8] Niu Huan. Research on the coupling relationship between industrial clusters and spatial pattern of towns in Jiangxi Province [D]. Jiangxi University of Finance and Economics, 2021. DOI: 10.27175/d.cnki.gjxcu.2021.001681.
- [9] Yang, C. C.. High-tech industrial clusters in China promoted by technological innovation competition [J]. Natural Dialectics Letters. 2022, 44(05).
- [10] Yang Zhicheng. China 's high-tech industrial clusters promoted by technological innovation competition [J]. Dialectics of Nature Newsletter. 2022, 44 (05).
- [11] Cheng Pengfei, Guan Xiaodong, Su Changgui and so on. The Impact of Science and Technology Finance on the Cultivation of World-class Manufacturing Clusters in China - A Case Study of Hunan Rail Transit Industry Cluster [J]. Hunan Social Sciences. 2021, (05).