Research on Digital Economy Development Based on Multiple Linear Regression and Random Forest Regression

Rui Yang [1st], Chenghui Wu [2nd], Xinkang Yang[3rd], Jing Liu[4th*], Junxian Wang [5th], Mengjuan Xia [6th], Ling Yuan [7th]

e-mail: yr20010507@163.com, e-mail: w1743432290@163.com, e-mail: yang1744837916@163.com, e-mail: 466181299@qq.com, e-mail: wangjunxian0305@163.com, e-mail: 2419966916@qq.com, e-mail: yl3302666256@163.com

[124567]School of Mathematics and Statistics, Suzhou University, Suzhou, Anhui, China

[3]School of Life Sciences, Bengbu Medical College, Bengbu, Anhui, China

Abstract—The digital economy is highly innovative, highly permeable, and widely covered. It is not only a new economic growth point, but also a fulcrum for transforming and upgrading traditional industries. The digital economy can effectively expand consumer demand, stimulate investment vitality, and provide new jobs. It is an indispensable part of the construction of a modern economic system. This research is based on the data related to the development of digital economy in Zhejiang Province, China from 2011 to 2020,and analyzes the research data through Rstudio to establish a multiple linear regression model and a random forest regression model. The study found that the development of the digital economy requires improving the quality of employment in the industry, deepening the level of industry development, and optimizing the development modelof the industry, rather than blindly expanding jobs. And put forward a plan to promote the development of the digital economy and solve the employment problem.

Keywords-digital economy; regression analysis; random forest

1 INTRODUCTION

Digital economy was originally a concept of economic development, and it was an efficient form in which human beings guided by big data and realized the rapid optimal allocation of resources, so as to achieve high-quality economic development[1].In the process of digital economy development, an economic form that directly or indirectly uses data to guide resources to play their role and promote the development of productive forces can be called digital economy. Because the digital economy covers a wide range of fields, there is room for progress in production, manufacturing and technology. As the scale of the digital economy continues to expand, bringing new kinetic energy to economic development, digital economy can achieve high-quality development by promoting entrepreneurial activities, thus realizing big data and cloud computing [2]. The birth of new terms such as " block chain " has become an important driving point of economic development [3]. In the era of digital economy, digitalization has become the key factor of production to promote economic development, and various industries including manufacturing industry are showing different global value chain

dynamics [4]. New functions and new forms. Real-time exchange of data and information have become a reality, and the rapid flow of information elements continuously enhances the spatial mobility of elements. China and even other countries in the world attach great importance to the goal of building a numerical power. In order to actively promote their own economic development, relevant implementation strategies and favorable policies are advanced [5]. Through the application and integration of government digital information technology, the integration of government and people, government and market, government and government has been obtained. Information exchanges and resource sharing among governments will provide more efficient and convenient digital government services for enterprises and resident, provide more fairer and effective development environment and institutional conditions for the market, and realize more effective allocation of production factor resources.

In this way, the government governance system and management models in the digital economy era can adapt, integrate and develop   the digital economy, and finally to promote the effective, stable and sustainable development of the social economy. This study takes the digital economic development level of Zhejiang Province in China as an example. It mainly explores the influence of the development of the digital economy, and at the same time look forward to the development of the digital economy in Zhejiang province in China at the same time.

2 MODEL CONSTRUCTION AND RESULT DISPLAY

This research is based on the data related to the development of the digital economy in Zhejiang Province, China from 2011 to 2020. The missing values in the collected data are filled with ARIMA predictions. The Internet penetration rate, employment in related industries and urban units, overall urban employment, and year-end regularity are used to fill in. The resident population, the penetration rate of mobile phones, the digital financial inclusion index of Peking University(In the subsequent study, the independent variables were named x1, x2...x7 in this order),and the added value of the tertiary industry are independent variables, and the total telecommunication business is used as the dependent variable to fit the development level of the digital economy, so as to construct a multiple linear regression model and a random Forest regression model.

As a traditional linear regression model, multiple linear regression is often a common model for studying such problems. In this study, a nonlinear regression model such as random forest regression model is innovatively introduced to study this problem from a nonlinear perspective. class problem, we may get unexpected results.

2.1 Multiple Linear Regression

The research data was analyzed by Rstudio, and the multiple linear regression results were obtained as shown in the following table.

Table 1 Multiple Linear Regression Results Table

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-414900	75410	-5.502	0.0315
x1	-2650	331.6	-7.991	0.0153

x2	795.4	167.5	-4.75	0.0416
x3	-8.124	7.963	-1.02	0.415
x4	111.3	18.4	6.048	0.0263
x5	121.3	52.5	2.311	0.147
x6	21.52	14.28	1.507	0.2708
x7	-3.187	0.6299	-5.059	0.0369

The results of multiple linear regression analysis show that the R-square value of the model is 0.997, which means that the Internet penetration rate, information transmission, software and information technology service industries employed in urban units, employed persons in urban units, permanent resident population at the end of the year, mobile phone penetration rate, and Peking University figures The financial inclusion index and the added value of the tertiary industry can explain 99.7% of the changes in the total telecommunications business, and the overall fitting effect of the multiple linear regression model is excellent. The model fitting effect graph is shown below.
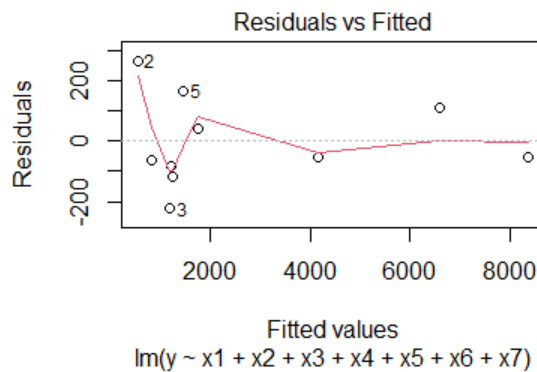

Figure 1. Multiple linear regression fitting effect chart

When the F test was performed on the model, it was found that the model passed the F test (F=98.979, p=0.010<0.05), which means that the Internet penetration rate, the employment of urban units in information transmission, software and information technology services, and the employment of urban units at the end of the year. At least one of the permanent population, the penetration rate of mobile phones, the digital financial inclusion index of Peking University, and the added value of the tertiary industry will have an impact on the total amount of telecommunications services.

It is worth noting that the regression coefficient value of employees in related industries and urban units is -795.439 (t=-4.750, p=0.0416<0.05), which means that employees in related industries and urban units will have a significant positive effect on the total telecommunications business. affect the relationship. The regression coefficient value of employed persons in urban

units is -8.124 (t=-1.020, p=0.415>0.05), which means that employed persons in urban units will not have an impact on the total amount of telecommunication business.

2.2 Random Forest Regression

Through Rstudio's analysis of the research data, the results of random forest regression are obtained. From the figure below, we can see that the error in the random forest regression model is basically stable when the number of decision trees is set at 5000.
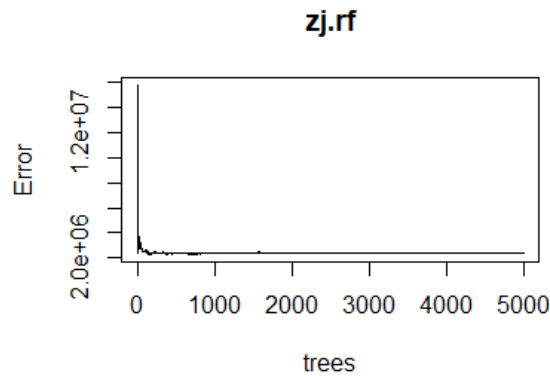

Figure 2. Random Forest Regression Mean Error Chart

The results of random forest regression analysis show that the model variance explanation coefficient is 0.6393,which means that the Internet penetration rate, the employment of urban units in information transmission, software and information technology services, the employment of urban units, the permanent resident population at the end of the year, the penetration rate of mobile phones, the number of Peking University The financial inclusion index and the added value of the tertiary industry can explain 63.93% of the changes in the total telecom business, and the random forest regression model has a good overall fitting effect.

Table 2 Random Forest Regression Variable Importance Table

variable	IncNodePurity
x1	1353590
x2	9833512
x3	4224624
x4	11456630
x5	4081140
x6	12253708
x7	11679784

As can be seen from the above table, in the random forest regression model, the most important variable is the Peking University Digital Financial Inclusion Index. It can be seen that the main

aspects of the development of the digital economy lie in the availability, coverage, and benefits of digital finance, while optimizing emerging emerging industries. to improve the overall environment for the development of the digital economy. We also noticed that in the model, the importance of employment in related industries in urban units is higher than that in urban units in general. Therefore, the development of the digital economy needs to improve the quality of employment in the industry, deepen the level of industry development, and optimize the development model of the industry instead of blindly expanding the number of jobs.

3 CONCLUSIONS

Based on the data related to the development of the digital economy in Zhejiang Province from 2011 to 2020, this research analyzes the impact on the development of the digital economy on the quality of industry employment, industry development level, industry development mode and high-quality employment, and finally draws the following conclusions.

First of all, the digital economy has a profound impact on the transformation of the job market, causing major changes in employment carriers and employment models. However, at the same time, the digital economy has made outstanding contributions to the provision of new jobs, stable employment, and high-quality employment. However, the digital economy contains. The huge potential for expanding total employment in China has not been fully activated, and there are still some shortcomings in the labor system and other aspects. At the same time, the digital economy has achieved remarkable results from optimizing the employment structure, but we still need to be alert to the risks brought by high-quality employment. The digital economy is an important cornerstone of achieving more adequate and higher-quality employment in the future. To break with employment barriers, it is necessary to optimize the industrial system, optimize the market system, stabilize the total number of employment, improve the employment structure, and improve the quality of employment, and strive to build human resources to ensure stability. Contribute to the overall national employment situation.

Second, digital finance is an important pillar of the development of the digital economy. Digital finance uses modern technology to expand the scope of financial services, improve the availability of financial services for a large number of smalls and micro economic entities and disadvantaged economic entities, and promote the efficiency and level of economic development of financial services. At the same time, digital finance provides a broader platform for the development of the digital economy, has a significant impact on the development of the digital economy, and has become an important driving force for the high-quality development of the digital economy. Among them, in the process of digital finance affecting the high-quality development of the digital economy, it has an important impact on the upgrading of the industrial structure, the quality of employment in the industry, the level of industry development, and the development model of the industry. Therefore, it is necessary to continuously promote the digital economy support services of digital finance, improve the inclusiveness and accuracy of digital finance, and continuously increase the depth of use of digital finance, so as to fully release the dividends of digital economic development.

Obviously, if we want to improve the digital economy, it is important to address the issues of high-quality employment and the development of digital finance.

REFERENCES

[1]   Liu Chuanhui. Research on the spatial economic connection and effect of urban agglomerations under the background of digital economy [D]. Southwestern University of Finance and Economics,2019. DOI: 10.27412/d.cnki.gxncu.2019.002463.

[2]   Li Fuyi. China's manufacturing industry and its governance and upgrading in the era of digital economy [D]. University of International Business and Economics,2018.

[3]   Hua Liwei. Research on the conduction path of my country's digital economy development to promote economic growth [D]. Zhejiang University,2021. DOI: 10.27461/d.cnki.gzjdx.2021.001019.

[4]   Xiao Ling. Measurement and Obstacle Factor Diagnosis of Digital Economy Development in the Yangtze River Economic Belt[D]. Jiangxi University of Finance and Economics, 2021. DOI:10.27175/d.cnki.gjxcu.2021.001038.

[5]   Zhao Tao, Zhang Zhi, Liang Shangkun. Digital Economy, Entrepreneurial Activity and High-Quality Development: Empirical Evidence from Chinese Cities [J]. Management World, 2020, 36(10): 65-76. DOI: 10.19744/j. cnki.11-1235/f.2020.0154.