Research on Macroeconomic Data Analysis Based on Data Mining Technology

Yingfa Yang, Guangwei Guo, Hui Zhao*
Hebei University of Engineering, Handan, Hebei, 056038, China
**Corresponding Author*

Abstract

Because of its excellent characteristics, big data is widely used in people's lives and work in various industries, and its application can play an important role in related fields. Big data is not only a new national strategic resource, but also a support for judging economic trends. Nowadays, the revolutionary influence of big data on macroeconomic analysis has been paid more and more attention by macroeconomic policy makers and academic circles. In the era of big data, complicated data are available in real time, and the whole social economy has undergone fundamental changes. Compared with the traditional data volume and data source channels generated by macroeconomics, there are more data volume and data source channels generated by macroeconomics under the background of big data era, and the authenticity and accuracy of data are higher. This paper discusses the sources of macroeconomic data mining and macroeconomic forecasting methods under the background of big data, and analyzes how to introduce data mining technology into macroeconomic analysis.

Keywords*: Big data, Macro economy, Data mining*

Ⅰ. Introduction

With the rapid development of the Internet era, the total information has exploded. Big data refers to a large and diverse data set. The processing of these data exceeds the current mainstream software capabilities, so it is necessary to improve the tools for processing data [1]. Big data is in the process of science and technology development, and big data has gradually been promoted and applied. Big data is widely used in people's lives and the work of various industries because of its own excellent characteristics. Its application can play a pivotal role in related fields [2]. Big data is of revolutionary significance to macroeconomic analysis. In the era of big data, the complex data can be obtained in real time, and the whole social economy has undergone fundamental changes [3]. From the short-term demand, the investment demand has risen, and consumption is slightly slower than last year, but the domestic demand is still stable in general [4]. With the development of macroeconomic theory, macroeconomic prediction has become another important aspect of empirical analysis. In an important aspect of economic model analysis and application, the core idea of traditional macroeconomic forecasting method is to find the internal laws of statistical data through specific models and methods. Big data mainly involves the collection, storage, analysis and application of data, which plays an increasingly important role in finance, e-commerce, medical treatment, agriculture, government affairs and other fields [5]. It not only leads the behavior of market subject, but also subverts the traditional mode of economic system by changing the market operation mode [6].

The long-term and stable economic development is one of the important development goals in China. There are many factors involved in economic development. The role of each factor will support or restrict the economic development. Therefore, if the economic development is not affected by various restrictive factors, we need to analyze these problems in advance. Macroeconomic prediction by big data method and technology is more and more valued by macroeconomic policy makers and academia [7]. The new information technology reduces the cost of information collection and information processing. In the era of big data, there are many kinds of data sources. More abundant analysis methods can be adopted to construct macroeconomic analysis model by using the collected massive data, which provides policy reference for the government to formulate policies [8]. Compared with the traditional macro-economic data and data source channels, the macro-economy produces more data and

data sources in the background of big data age, and the data authenticity and accuracy are higher [9]. Macroeconomic analysis can be based on these large amounts of data for scientific analysis, which provides an important reference for guiding China's economic development. This paper discusses the source and the method of macroeconomic prediction of macroeconomic data mining in the background of big data, and analyzes how to introduce data mining technology into macroeconomic analysis.

Ⅱ. Characteristics of "Big Data" and Its Complementarity to Traditional Macroeconomic Forecast

Big data has initiated a great transformation of the times. As far as macroeconomic analysis is concerned, the transformation brought by the era of big data is significant and revolutionary. The data provided by the network platform records the information according to the actual situation when the event occurs, which reduces the manual operation and provides relatively original data, rather than the data information collected artificially and processed, so it is more accurate. Under the background of big data era, a large amount of data is generated all the time, and its impact on macro economy also exists all the time. In macroeconomic analysis, we will be faced with the ever-increasing data, which makes the selection of information more difficult. Traditional economic analysis relies mainly on samples, but in the era of big data, the obtained data may be the whole itself. For example, in terms of prices, every transaction information that is sold on e-commerce websites can be recorded [11]. When some enterprises analyze macroeconomics, in the current era of big data, they can publicize relevant information through some online publicity platforms and channels such as WeChat and Weibo. Compared with the traditional propaganda methods such as newspapers, this kind of propaganda way through the network platform and channels has higher timeliness and wider propaganda scope, which is the best way to improve the propaganda speed for enterprises.

By using the Internet big data information, we can obtain the whole or nearly whole sample information, instead of obtaining the sample information by statistical sampling to infer the whole information. With the support of such big data, the sample size used for calculation is massive, and it can greatly approach the whole sample, so as to directly obtain the most authentic and comprehensive statistical index information. Big data has its own unique characteristics in economics, which constantly promote economic development and have a greater impact on the economy. Many characteristics of big data also provide a more favorable development direction for economic development, such as the effectiveness of data and the timeliness of many data. Under the background of big data, macroeconomic analysis not only needs to analyze the overall structure of the economy, but also needs to produce some influential network information data, which makes the actual working mode more complicated to a certain extent and improves the difficulty of information selection to a certain extent [12]. Large and comprehensive available data is extremely important for macroeconomic analysis, which can accurately understand the macroeconomic situation, correctly predict macroeconomic development and reasonably formulate macroeconomic policies. The arrival of the era of big data brings convenience to the collection and processing of data, but it also threatens the security of data to a certain extent, reduces the guarantee degree of data security, and affects the timeliness of macroeconomic analysis. Figure 1 shows the multivariate analysis process in financial analysis and management.

ISSN: 0010-8189

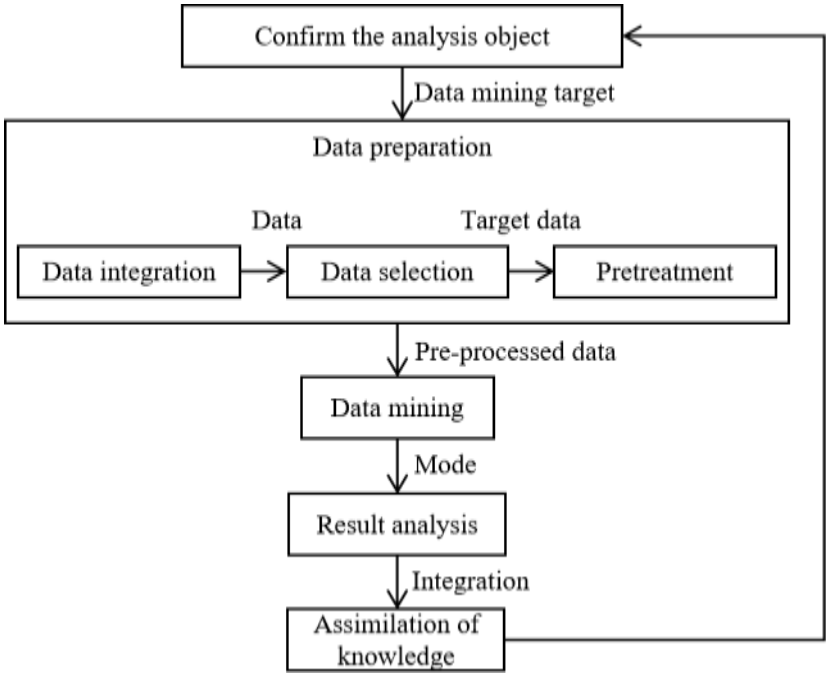

Fig.1 Multivariate Analysis Process in Financial Analysis and Management

Big data is not an innovation to the traditional macro forecasting model, but a supplement and improvement to the traditional macro economic analysis and forecasting method. By improving the data used in traditional statistical analysis methods, we can break through the fundamental limitations of traditional methods and improve the prediction effect and application scope of the model. Big data economic model can make full use of the real-time nature of data, improve the timeliness of analysis or prediction, and provide the fastest information and basis for economic early warning and policy making. In the in-depth analysis of macroeconomic data, we mainly analyze the relevant information database comprehensively. However, the security degree of the database itself needs to be improved at present. Therefore, in-depth detailed analysis will also reduce the guarantee degree of data security. Through big data technology and methods, timely data can be obtained. Combined with the traditional macro forecasting and analysis model, economic theory can be effectively used to explain economic problems, and the data information obtained by big data can break through the problems existing in traditional statistical data, effectively improve the effect of macroeconomic forecasting and analysis, and bring new breakthroughs for macroeconomic forecasting and analysis.

Ⅲ. Application of Big Data Algorithm in Macro Economy

A. Application of Correlation Algorithm

Big data has initiated a great transformation of the times. As far as macroeconomic analysis is concerned, the traditional analysis method mainly collects specified information from fixed channels, obtains structured data through statistical sampling, and forecasts and analyzes macroeconomics by comparing the changes of samples. Cost and risk control need to be realized by analyzing the corresponding data, which is also an important content in the financial industry. There is a huge amount of data information in big data, so it plays an obvious role in the financial industry, especially in the innovation of the current financial industry. Internet is an indispensable tool for people's economic life in the era of big data. People's online life occupies an increasingly important position in daily life, and the way of information acquisition shifts to the network channel. In the era of big data, the speed of information transmission has greatly increased, and people can quickly collect a large amount of data, such as transaction information, through various channels.

Correlation analysis refers to the analysis of the direction and degree of the relationship between two or more phenomena, to determine whether there is a correlation between phenomena and the closeness of the correlation. Correlation algorithm is to judge the correlation between two things by correlation coefficient, and the expression of correlation coefficient is:

$$r = \frac{\sum_{i-1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i-1}^{n}(x_i - \bar{x})^2 \times \sum_{i-1}^{n}(y_i - \bar{y})^2}} = \frac{n\sum_{i-1}^{n} x_i y_i - \sum_{i-1}^{n} x_i \times \sum_{i-1}^{n} y_i}{\sqrt{n\sum_{i-1}^{n} x_i^2 - \left(\sum_{i-1}^{n}(x_i)^2\right)} \times \sqrt{n\sum_{i-1}^{n} y_i^2 - \left(\sum_{i-1}^{n}(y_i)^2\right)}} \quad (1)$$

In the production of national economy, the development of various industries is inextricably linked. By using correlation algorithm, the relationship between industries can be deeply explored. By mining and analyzing a large number of transaction data, decision makers can judge the high and low trend of prices, and then judge the trend of macroeconomics. In the era of big data, the real-time and rapidity of information provides the possibility of timely and efficient prediction, and the huge scale of information provides the opportunity of accurate and reasonable prediction. Under the background of big data, a lot of data information is of low value, even without any analytical value. Therefore, the accuracy and authenticity of this information need to be identified and analyzed by special safety supervisors. Macro-economic analysis is based on large and comprehensive available data. The era of big data improves the speed of obtaining information and broadens the channels of collecting information, which not only allows users to obtain a large amount of useful data more conveniently, but also greatly reduces the time lag in traditional analysis methods, and enables decision-makers to have a more accurate understanding of the macro-economic situation. For some work processes, detailed and unified work steps should be worked out to prevent managers from appearing abnormal operation processes and causing data leakage. When recruiting, enterprises should consider students' data sensitivity and new technology application ability, establish the concept of the importance of big data talents, regularly and irregularly carry out big data training in combination with national forms, write big data teaching materials suitable for enterprises, and build a big data talent system.

B. Application of Clustering Algorithm

In the Internet age, the frequency of updating the business requirements of enterprises has accelerated a lot, so the analysis and processing model of big data must be flexible to meet the new business requirements, which also leads to its complexity. The application of big data in internet finance can effectively promote the optimal allocation of resources. At present, the judgment of macroeconomics depends on the statistical data released by various statistical investigation systems, but one of the biggest difficulties is that the data on macroeconomic statistics lag too much, and the prediction based on this statistics is even considered to contribute to macroeconomic fluctuations. The generation of a large amount of real-time data in the era of big data makes it possible to estimate macroeconomics more quickly, and the development of big data theory and methods provides new conditions for current forecasting. When big data analyzes economic problems, it is more descriptive than structural, and lacks the support of economic theory [13]. Facing the expansion of statistical objects brought by big data, the government should not only pay attention to structured data, but also pay attention to mining unstructured data in order to find appropriate economic statistical indicators. Every enterprise has its own transaction records. If the big data in the enterprise is used to extract the relevant data reasonably, the problems of the enterprise can be judged effectively, such as the financial management system in the enterprise, the current business situation and so on. After analyzing these problems through big data, we can find more effective improvement strategies.

Clustering is to place classified objects in a multi-dimensional space and classify them according to their spatial relationship. R index can be used to evaluate the clustering effect. R index is defined as: calculate the average square of the distances between all elements in each cluster center and its corresponding class, and measure the

cohesion of the class.

$$R = \frac{1}{n}\sum_{i=1}^{k}\sum_{x_j \in s_i}\left\|x_j - u_i\right\|^2 \quad (2)$$

Where n is the number of all elements in the model, u_i is the cluster center of the i-th class, and $\left\|x_j - u_i\right\|^2$ is the square of the distance between the two points.

The big data revolution provides opportunities for the government to make macroeconomic policies. In policy making, the government can improve the quality of public services, increase the types of services and provide better policy guidance for public services through big data analysis system. Through big data technology and methods, timely data can be obtained. Combined with the traditional macro forecasting and analysis model, economic theory can be effectively used to explain economic problems, and the data information obtained by big data can break through the problems existing in traditional statistical data, effectively improve the effect of macroeconomic forecasting and analysis, and bring new breakthroughs for macroeconomic forecasting and analysis. Using clustering algorithm, enterprises can be classified unsupervised, and the basic data set is constructed by using enterprise economic index data, and the data set is trained by clustering algorithm. Through the data characteristics of various centers, the characteristics of various enterprises are analyzed to make the characteristics of enterprises more obvious, so as to have a deeper understanding of the economic characteristics of enterprises and provide data basis for the government's supporting policies.

Ⅳ. Conclusions

Under the background of big data era, macroeconomic analysis will not only face certain opportunities, but also be challenged to a certain extent. The era of big data has greatly broadened the sources of information, improved the timeliness of obtaining information, and provided powerful data resources and technical support for macroeconomic analysis. In order to better develop the economy, we should establish a macro-economic big data analysis model as soon as possible, and collect the data needed by the model through various channels, so as to complete the macro-economic big data prediction and analysis system as soon as possible, and provide more timely and accurate predictions for decision-makers in macroeconomic regulation and control. Facing the expansion of statistical objects brought by big data, the government should not only pay attention to structured data, but also pay attention to mining unstructured data in order to find appropriate economic statistical indicators. In order to improve the efficiency of macroeconomic analysis in the era of big data, we should constantly improve the technology of collecting and processing related information and data, pay attention to environmental supervision and provide corresponding guarantees, so as to give full play to the role of big data technology in macroeconomic analysis and promote the healthy and sustainable development of China's economy.

Acknowledgement

Theresearchresultsofhandansocialscienceplanningproject”researchontheapplicationofintelligenttechnologyandhandancityannotatedwithculturalimplications”(2019108).

References

[1] Wang Yu, Wang Ning. Automatic early warning system of economic management benefit loss risk based on data mining[J]. Automation and Instrumentation, 2020, 246(04):139-142.
[2] Chen Xing. Research on Regional Economic Correlation Development Based on Data Mining[J]. Guide

Journal of Science & Technology Economy, 2020, 733(35):245-246.
[3] Wang Tong. On the application of data mining technology in economic statistics[J]. Industrial Innovation Research, 2019, 26(12):176-177.
[4] Shi Li. Research on the Economic Benefits of Headquarters Economy Enterprises Based on Small Sample Data Mining[J]. Statistics and Management, 2020, 035(004):14-19.
[5] Zhao Junhao, Li Yuhua, Huo Lin, et al. Macroeconomic forecasting method integrating Weibo sentiment analysis and deep learning [J]. Computer Applications, 2018, 038(011): 3057-3062.
[6] Niu Wensheng, Dong Xiaobo, Gao Jiali. The application of big data in macroeconomic analysis[J]. Economic Forum, 2018, 575(06):92-94.
[7] Zhang Tao, Liu Kuanbin. The application of "big data" in macroeconomic forecasting and analysis[J]. Finance Think Tank, 2018, 15(03):67-85+145.
[8] Zhao Junhao, Li Yuhua, Huo Lin, et al. Macroeconomic forecasting method integrating Weibo sentiment analysis and deep learning[J]. Computer Applications, 2018, 38(11):11-16.
[9] Li Xinhua, Wang Yong, Yan Jiajing, Hou Wengang. Multi-source and multi-dimensional government affairs shared data analysis system for macroeconomic analysis [J]. Computer and Modernization, 2020, 301(09): 29-35+40.
[10] Zhang Zhen. Research on the Application of Data Mining Technology in Economic Statistics [J]. Satellite TV and Broadband Multimedia, 2020, 510(05):17-18.
[11] Zhu Zhaojun. Research on the Application of Data Mining Technology in Economic Statistics [J]. Modern Business, 2020, 569(16):120-121.
[12] Li Ye, Lu Shu, Yuan Xiang, et al. Macroeconomic index prediction algorithm based on correlation analysis [J]. Command Information System and Technology, 2020, 011(001): 84-88.
[13] Zhao Jianbo. The application of data mining technology in economic statistics[J]. Economic Research Guide, 2018, 362(12):193-194.
[14] Tian Guanshun. Discuss the application of data mining technology in economic statistics[J]. Industrial Technology Innovation, 2019, 33(33):96-98.