Deep research on big data index analysis method in intelligent manufacturing industry

YuXiang Song
Guangdong Peizheng College Guangzhou, China
Emal:2603767@peizheng.edu.cn

Abstract. Big data analysis is widely used in intelligent manufacturing industry. Data analysis refers to analyzing a large amount of data collected by appropriate statistical and analytical methods, summarizing and understanding them, and digesting them, so as to maximize the development of data functions, give full play to the role of data, and improve and strengthen the intelligent level of manufacturing system. Data analysis is to extract useful information and form conclusions, and to study and summarize the data in detail. The data of intelligent manufacturing system is also called observation value, which is the result of experiment, measurement, observation and investigation. The data processed in data analysis is divided into qualitative data and quantitative data. The results of data analysis are of great significance in the intelligent manufacturing industry.

1. Data analysis type description
In the field of intelligent manufacturing and statistics, data analysis is divided into descriptive statistical analysis, exploratory data analysis and confirmatory data analysis. Exploratory data analysis focuses on the discovery of new features in the data, while confirmatory data analysis focuses on the confirmation of existing hypotheses.

Exploratory data analysis is a method to analyze the data in order to form a hypothesis worthy test. It is a supplement to the traditional statistical hypothesis testing method and an innovative data analysis process. Qualitative data analysis, also known as qualitative data analysis, refers to the analysis of non numerical data such as phrases, photos, data and observations. Offline data analysis is used for more complex and time-consuming data analysis and processing, and is usually built on the cloud computing platform. Online data analysis, also known as online analytical processing, is used to process users' online requests, which requires high response time.

Index data statistics is a method to sort out and analyze the data through charts or mathematical methods, and to estimate and describe the distribution state, digital characteristics and the relationship between random variables of the data. Descriptive statistics is divided into trend analysis and correlation analysis. Trend analysis mainly relies on the average, median and mode statistical indicators to express the trend of data. Trend analysis is used to measure the relationship between two random variables statistics, standard deviation and other statistical indicators to study the trend of data.

2. Data analysis method

2.1. List method
It is the most common method to record and process the data in the form of list according to certain rules. The design of the table requires clear corresponding relationship, which is conducive to find the...
correlation between the relevant quantities. In addition, it is also required to indicate the name, symbol, order of magnitude and unit of each quantity in the title block. Calculation columns and statistics columns other than original data can also be listed as needed.

2.2. Drawing method
The mapping method can express the changing relationship between various physical quantities intuitively and vividly. Some experimental results can be obtained from the graph line, and some complex functional relations can be expressed graphically through certain transformation.

Charts and graphs are mainly generated automatically by programs. Tabulation by programs is to input survey data into programs through corresponding software. Through the operation of these software, the final results can be obtained, and the results can be displayed by charts or graphs. Graphs and charts can directly reflect the results of system analysis.

2.3. Analysis software
Using data analysis tools can complete many professional software data statistics and analysis. It includes: histogram, correlation coefficient, covariance, probability distribution, sampling and dynamic simulation, overall mean judgment, mean inference, linear, nonlinear regression, multiple regression analysis, moving average and so on.

3. Index analysis algorithm description
In practical work, the index analysis method is the most widely used, but also in the use of other methods for analysis at the same time with the use of prominent key points of the method, the analysis directly uses some basic indicators in statistics to do data analysis, such as the average, mode, median, maximum, minimum and so on. When choosing which basic index to use, we need to consider the orientation of the analysis results.

average
It can show the data situation of the same kind of data in different time periods, and can be used to summarize trends and find problems in general laws. We can also compare the differences of similar data in different regions, different environments, different conditions and different situations, which is more convincing than the total or individual value.

median
Also known as median, it refers to the number in the middle of a group of data arranged in order, representing a numerical value in a sample, population or probability distribution, which can divide the numerical set into two equal parts. Because it is obtained by sorting, it is not affected by the maximum and minimum extreme values. For example, in the statistics of market recruitment salary in this quarter, it is more meaningful to use the median because a small part may belong to the maximum or minimum value.

The change of some data has no effect on the median. When the individual data in a group changes greatly, it can often be used to describe the centralized trend of this group of data.

Maximum and minimum
The maximum and minimum values can often be used to show the abnormal conditions in the data. In some data analysis, the abnormal values can be ignored, but some analysis of the maximum and minimum values can study the influencing factors, so as to find breakthrough actions or avoidable methods, so as to promote the growth of the business.

4. Example of index analysis algorithm
This time, 5000 pieces of data are used. After sorting, screening and filtering, 2000 pieces of data are left. After another selection, 256 representative data are retained, and finally 108 data are selected for system data analysis and processing.

Data analysis is a large-scale concept. It is a process of checking, cleaning, transforming and modeling data. The purpose is to find useful information, draw conclusions and promote
decision-making. The process of data analysis can be divided into the following steps.
Clear analysis purpose → form analysis ideas → build analysis index system → collect data →
process data → establish analysis model → visual management.

4.1. Data collection
Open the data file dd77.dat and select 108 representative data. As shown in Figure 1:

![Figure 1 Schematic diagram of a group of 108 data in data file dd77.dat](image1)

Data acquisition refers to the real-time acquisition of any object or process that needs to be
monitored, connected and interacted in the industrial system through a variety of information sensing
devices, and the acquisition of various required information such as sound, light, heat, electricity,
mechanics, chemistry, biology, location, etc., forming a huge network combined with the Internet. Its
purpose is to realize the connection between objects, objects and people, all objects and the network,
so as to facilitate the identification, management and control. The Internet of things breaks through the
traditional thinking of separating physical equipment and information transmission, realizes the
communication between things, embodies the concept of great integration, and has great strategic
significance. The existing communication is mainly interpersonal communication. At present, the
global communication users are close to saturation, and the development space is limited. The
communication objects of the Internet of things are more things. If these so-called things are included
in the communication application category of the Internet of things, the number of communication
connections that may be involved can reach tens of billions, which provides a huge space for the
expansion of the communication field.

4.2. Use the sorting program to sort 108 data. The data file dd88.dat is formed, and the data sorting is
shown in Figure 2

![Figure 2 Schematic diagram of 108 data sorting in data file dd88.dat](image2)

4.3. Run the index data analysis program and call the drawing program. The graphical results of data
analysis are shown in Figure 3

![Figure 3 Schematic diagram of 108 data in data file dd77.dat](image3)
4.4. Analysis on the running results of index data analysis method

Through the analysis of 108 data, the results are as follows:
- The maximum number is 255. As benefit, income, income, gain and profit.
- The minimum value is 1. As cost, input, loss and input-output ratio.
- The average value was 112. As the balance point of data system.
- Median 114. As the value of probability distribution. Wages, conditions, environment and treatment.
- The number of frequent occurrence is 11,28,34. The benefit value of the system, the number of times users buy.
- Other data are mode.

5. Summary

In this paper, in the process of data analysis, there will be many factors affecting the data of the system. Business trend analysis based on data indicators is the most basic way to display data information. In data analysis, we can quickly understand the market trend of enterprises and businesses, the number of orders and the completion of performance through intuitive figures or trend charts, so as to intuitively absorb data information. Data analysis can help enterprises and businesses to optimize the process, reduce costs and improve turnover. This kind of data analysis is defined as index data analysis. The goal of index data analysis is to use big data to make quick, high-quality and efficient decisions for all workplace enterprises and businesses, and to provide scalable solutions. The essence of index data analysis is to create business value and drive the growth of business efficiency of enterprises and businesses.

Reference

[1] Cheng Xueqi. Overview of big data system and analysis technology. Acta Sinica Sinica [J], 2014
[2] Ren Lei. Overview of big data visual analysis. Acta Sinica Sinica [J], 2014
[3] Wu Xiaofang. New thinking of Internet of things and big data [J]. Communication world, 2017
[4] Wei Wei. Learning control and its application in robot control [D]. Zhejiang University, 1994
[5] Jin Yi. Application of fuzzy set theory in production planning and scheduling [D]. Southeast University, 1994
[6] Jiang Qiyou. Research on artificial intelligence theory and technology and its application in fault diagnosis of large centrifugal compressor [D]. Beijing University of chemical technology, 1993
[7] Chen Hui. On the application of big data technology in the Internet of things industry [J]. Jiangxi communication technology, 2017
[8] Wang Zhiliang. Internet of things engineering training course [M]. China Machine Press, 2012