Application of Big Data technology in Enterprise Management

Chao Zhang*

Department of Finance and Economics,Chengdu Polytechnic,Sichuan,China
*Corresponding author: 56142797@qq.com

ABSTRACT: Big data is quietly changing our world; all walks of life are using big data. Big data can provide decision-making basis for the government, enterprises, research and development, so it is very necessary to master the correct big data analysis method and intelligent, in-depth and valuable information extraction! Shannon, the father of information theory, once said that information is used to eliminate mistrust, such as predicting whether it will rain tomorrow. If you know today's weather, wind speed, clouds, air pressure and other information, it will help to draw more accurate conclusions. Therefore, big data is used to eliminate uncertainty and master more effective data, which can drive enterprises to make scientific and objective decisions.

KEYWORDS: big data; enterprise management; informatization

1. Definition of Big Data

With the advent of the era of big data, the hot concept of AI, people's cognition has improved. Why is big data valuable? Is this just a virtual concept? How do you think about data driven? Why is it more effective to have more data? These questions are hard to answer, but big data is by no means big and empty [1].

Data acquisition follows the four characters rule of "big", "complete", "fine" and "timely".

"Big" emphasizes macro "big", not physical "big". Big data is not blindly pursuing the "big" data volume. For example, the apple price data of each prefecture level city is only 2MB every day, but an apple intelligent scheduling system based on this is a big data application. Although some data are very large, their value is limited.

"All" emphasizes multiple data sources. Big data collection focuses on total quantity rather than sampling. In addition to collecting client data, it also needs to collect data such as server logs, business databases, and third-party services, such as the poll before the US election [2]. Hilary has more than 70% odds, but trump

became president of the United States, because the data are biased, and the bottom people of trump support will not reply online.

"Fine" emphasizes multi-dimensional data collection, that is, the dimension, attribute and field of events are collected [3]. For example, for the event of "joining the shopping cart" in the e-commerce industry, in addition to collecting the click data of users, it is also necessary to collect the data of which commodity the user clicks and the corresponding merchant, so as to facilitate subsequent cross analysis.

Time emphasizes the timeliness of data. Obviously, the data with timeliness has reference value. For example, national index and CPI index, the value of information collected at the beginning of the month is obviously different from that obtained in the middle of the month. The data need to be obtained in real time and analyzed in real time.

2. Analysis Method of Big Data

2.1 Formula Method

The so-called formula method is to use formula to decompose the influencing factors of an index. In Bi tools, "drill down" is also an application of this method. For example, analyze the reasons for the low sales volume of products, decompose the sales level by layer, and refine the granularity of evaluation and analysis.

2.2 Comparison Method

Comparative method is to compare two or more data, which is the most common method, and it needs to be mastered in rapid data analysis [4]. For example, compared with competitors, year-on-year and month on month, growth rate and fixed ratio in time dimension.

2.3 Quadrant Method

It can be expressed by two dimensions. Quadrant method is a strategy driven idea, which is usually used in product analysis, market analysis, customer management and commodity management.

For example, the classic RFM model divides customers into eight quadrants according to the recent consumption (proximate cause), consumption frequency (frequency), and consumption amount (currency).

The launch quadrant analysis method can establish the optimization strategy of different quadrants. For example, provide more resources for key development customers. It is essential in the rapid data analysis [5].

2.4 Pareto Analysis

In terms of personal wealth, it can be said that 20 percent of the world's people have 80 per cent of their wealth. It can be understood in data analysis that 20% of data has 80% effect, and it needs to be mined around 20%.

For example, it is used for product classification to measure and build ABC models. For example, 20 percent of customers who contribute 80 percent of the company's profits, then focus on maintaining these 20 percent of the customers with limited resources. This point needs to be paid attention to when analyzing the fast-growing data.

2.5 Funnel Method

Funnel method is a funnel diagram, a little like inverted pyramid. It's a streamlined way of thinking [6]. It is often used to develop new users, shopping conversion rates, which have changes and certain processes in the analysis.

The core idea of the overall funnel model can be divided into decomposition and quantification. For example, monitor user transitions at each level and find optimized points for each level.

3. Specific Strategies of Big Data in Enterprise Management

3.1 Formula Method

In an enterprise, for sales personnel, data itself is not the result, but should serve the process of business decision-making. However, many people are eager to collect all the information available and find the most interesting findings, instead of finding the decisions they need to make at the moment from the business itself, and deduce backward what data they really need, how to analyze it, and what insights they can come up with. This leads to misunderstanding of data. Decision making should be made before data. The process of data analysis is the support point of decision-making. For sales personnel, what they need may not be a series of data, but face-to-face communication with customers is the only way to find a breakthrough point to reach the final transaction. In this way, it is very important to improve the creative thinking of front-line employees.

3.2. Making Good Use of Data for Decision Making

Data analysis is not comprehensive and detailed. The use of data may be mixed with feelings. If the decision-makers only rely on customer analysis to make decisions, it will lead to one-sided data. The process of data analysis is a comprehensive process, which needs to integrate multiple factors, market trend prediction, the executive power of front-line staff, etc. At the same time, the

selection of data also needs to be adjusted by many parties. Good at using data to make multiple decisions. To deal with all kinds of sudden situations.

3.3. Joint Development with Users

The implementation of enterprise strategy is closely around a central point "user". For enterprises, the employees who contact users are front-line employees. Through continuous contact with customers, front-line staff gradually find out the needs of users. For decision makers, this is where data analysis can be used for reference. Through the feedback of users, we constantly adjust the strategy according to the needs of users. Develop together with users, and finally serve users.

3.4. Establish Analytical Ability and Culture Extensively

In the process of data analysis and decision-making, enterprises should always communicate with front-line staff in a timely manner. And the feedback of the front-line staff is solved in time, and a set of strong analysis system and strong corporate culture atmosphere are established by combining the strength of various parties in the enterprise. Only when the right front-line personnel get the right and insightful information, can the enterprise achieve a truly gorgeous transformation.

The construction of the system framework needs practical tools to complete. In terms of data analysis and information feedback, visual analysis comprehensively displays the data, which improves the efficiency of data analysis. The application of intelligent permissions facilitates the front-line staff to grasp the trend of data changes in time, and the general-purpose function can display the information more conveniently. Deeply loved by all walks of life [7].

4. Four Key Links of Data Analysis in Enterprises

Data analysis is divided into four parts: data acquisition, data modeling, data analysis and indicators. He put forward a point of view, to do a good job in data analysis, we must have a bottom-up concept. Data analysis in many companies is driven from top to bottom, and business analysis indicators are used to determine what data to collect. This is a demand driven engineer mode, which is not conducive to long-term data collection. A healthy bottom-up model can help companies to establish a data flow and data analysis system in line with their own business.

4.1. Data Acquisition

To really do a good job in big data analysis, we must first build a good data base, and the core is "comprehensive" and "detailed". Data collection cannot be done only through app or the client collects data. The server data and database data should be collected at the same time to collect full data instead of sampling data. At the same

time, relevant dimensions should be recorded. Otherwise, it may be found that the historical data is insufficient when analyzing the business. Therefore, do not care about the excessive amount of data. The cost of disk storage is very cheap compared with the value of data accumulation.

4.2 Data Modeling

Many companies have business databases, which store user registration information, transaction information, etc., and then product managers and operators seek help from technical personnel, and use business databases to support business data analysis. But this maintenance cost is very high, and tens of millions, hundreds of millions of data cannot be operated well. Therefore, data analysis and normal business operation have two kinds of analysis. Data analysis models and solves problems separately.

There are two standards for data modeling: easy to understand and good performance.

Data driven is not the patent of data analysts and database administrators, so that every business personnel of the company can use data for data analysis in their work, and can obtain second level response, verify their new ideas, new ideas and try new methods, which is the healthy state of data-driven for all employees.

Multidimensional data analysis model (OLAP) is the most effective model in user data analysis. It classifies the user's access data into dimensions and indicators. City is the dimension, operating system is also the dimension, sales and user volume are indicators. A good multidimensional data analysis model is established to solve the problem of some business index analysis, and users can flexibly combine to meet various needs.

4.3. Data Analysis

Data analysis supports product improvement.

It is unscientific for product managers to improve the functions of products, and then reprocess the primary ideas. Lean entrepreneurship talks about a concept, which introduces data analysis into product iteration, conducts data collection and data analysis on existing functions, and draws useful conclusions, which are introduced into the next iteration, so as to improve products. Big data analysis is crucial in this process.

The founder of Facebook once described how his company determined the direction of product improvement. Facebook uses a mechanism: if every employee has an idea, they can sample hundreds of thousands of users to try it. If the result is not good, they will give up the idea. If the effect is very good, it will be promoted to a larger range. This is a scientific method to introduce data analysis into product iteration.

4.4. Indicators

How to define indicators? For enterprises, there are two methods that are very effective: the first key index method and the pirate index method.

The first key indicator method is that any product at a certain stage has an indicator that needs to be paid attention to most [8]. Other indicators are derivatives of this indicator. This indicator determines the current work focus of the company. For a start-up company, it may start to pay attention to daily life and expand some indicators around daily life. When the company's products mature, realization will become the key Revenue (Gmv) will become the first key indicator.

5. Data Analysis

Data analysis leaders need to act now, but also to think about the future. Balancing these needs requires them to play a leading role in developing a clear, comprehensive strategy.

However, enterprises usually lack proper data and analysis organization structure. In many companies, the data and analysis groups belong to the IT department and are rarely funded in terms of personnel and budget, as they are often seen as cost centers, which makes it difficult to justify the high spending on analytical tools and skills.

To overcome budget constraints, managers in these groups should first understand specific organizational needs to establish their unique data management strategies. In this way, they avoid investing millions of dollars in a complex data management infrastructure, only to find that they need far more resources. In addition, analysis leaders can ensure that the tools and skills groups budget by measuring the ROI of the system and emphasizing short-term and long-term benefits [9].

As the amount of data an organization and it collects grows over time, data analysis can become increasingly difficult to navigate. Without proper policies, the process of gathering information, collaborating on projects, and generating reports is prone to errors. In order to avoid inaccuracy, the key is to implement a system that can adapt to the fast changing speed and develop with the organization.

According to IDC, the success of big data and analytics can be driven by increased collaboration, especially among it, line of business and analysis teams. That's why risk managers should choose flexible tools that provide a 360 degree view of data and leverage integrated processing and analysis capabilities. Automated collection and sorting, easy sharing and extraction, real-time collaboration, and the ability to compress various data sets into one type of analysis are some of the most important factors to consider.

6. Conclusion

In the era of big data, it provides data reference for enterprise decision-making. However, for the first-line individuals, the decision-making plan at the enterprise level should deal with the first-line problems and carry out structural transformation. The valuable experience gained by time-consuming and laborious work is really applied to front-line staff.

In the future, the amount of data processing will increase geometrically, and the existing processing and analysis technology may not be able to meet the timely analysis of data. Therefore, enhancing the ability of data analysis will be the first priority in the development of data analysis technology in the future. At the same time, strengthening the combination of analysis technology and Bi will continue to promote the development of enterprises.

The huge amount of data will lead to the lack of data management. The development of data analysis technology in the future will also improve the ability of data management. Enhance data management, use ml function and AI engine to develop enterprise information management categories, including data quality, metadata management, master data management, data integration, and self configuration and self adjustment of database management system. It is believed that in the future, manual data management tasks will be greatly reduced by adding ml and automating service level management [10].

The development of intelligence solves the tedious operation process, and the future development of data analysis technology will follow the development of intelligence. Continuous intelligence is a design pattern in which real-time analysis is integrated into business operations to process current and historical data to specify actions that respond to events. It provides decision automation or decision support. It provides a powerful driving force for data analysis and processing. The core concept of blockchain is decentralization. The proposition is to provide decentralized trust in an untrusted network of participants. The future development of data analysis technology also needs to solve the distrust in the network. Through the blockchain to improve the trust between people, data sharing will be powerful.

Blockchain is a data source, not a database, and will not replace the existing data management technology. It will be the driving force of the data.

The new persistent memory technology will help to reduce the cost and complexity of the architecture supporting memory computing. High capacity memory with high cost performance can be provided for high performance workload. It has the potential to improve application performance, availability, startup time, clustering methods and security practices while controlling costs. The new server workload requires not only faster CPU performance, but also larger memory and faster storage space.

References

[1] Souiden Imen, Brahmi Zaki (2017). A survey on outlier detection in the context of stream mining. 16th International Conference on Intelligent Systems Design and Applications,p.363-372.

[2] Tao G, Yan Y G, Liu J, and Zou J (2015). Discrimination of continuous attributes based on improved SOM clustering. Journal of Computer Applications, p.83-87.

[3] He Ming and Ren Wanpeng (2017). Attribute Reduction with Rough Set in Context-Aware Collaborative Filtering, China Journal of Electronics,p.24-25.

[4] Yu Zhanqiu (2019). Big data clustering analysis algorithm for internet of things based on K-means. International Journal of Distributed Systems and Technologies, p.3-7.

[5] Portugal Ivens, Alencar Paulo, Cowan Donald (2018). The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications ,p.212-214.

[6] H Wang, SL Hung (2012). Phylogenetic tree selection by the adjusted k-means approach. Journal of Applied Statistics, p.655-657.

[7] Zhang R G, Hu X H, Zong Y S (2017). Discretization of continuous attributes based on improved discrete particle swarm optimization. Computer Engineering and Applications,p.101-104.

[8] Jonathan L, Shyam V, Vanathi G (2011). Application of an efficient Bayesian discretization method to biomedical data. BMC Bioinformatics, p.1410-1610.

[9] Zang Maolin (2018). Human Resource Management in the Era of Big Data. Journal of Human Resource and Sustainability Studies,p.25-27.

[10] Wang L (2018). Algorithm of continuous attribute discretization based on improved particle swarm. CEA, p.21-30.