Problems and Solutions of Data Analytics in Business

Liwei Liu[ID][a]

School of Engineering and Applied Science, Gonzaga University, Spokane, U.S.A.

Keywords: Data Quality, Data Breaches, Business Data Analysis, Data Encryption, Interdepartmental Collaboration.

Abstract: Nowadays, business cannot thrive without data, and naturally, data analytics has become an indispensable tool in business. Data analysis can not only intuitively provide decision-making support, but also help enterprises to understand the market and customers more deeply and provide enterprises with more accurate and efficient programs and development direction. But while data analytics in business provides results for companies, there are also not a few problems. This study analyzes the problems of data analytics in business by analyzing them and proposing countermeasures. This paper examines four problems of data analysis in enterprises: data quality problem, data leakage problem, inaccurate data analysis problem, and data application problem. In response to these problems, this paper establishes a data processing mechanism, encrypts data, upgrades the tools and related talents for data analysis, and promotes good communication among departments, so as to ensure the stability of data quality, prevent the leakage of information, improve the accuracy rate, and enable better application in reality.

1 INTRODUCTION

In today's competitive and ever-changing business environment, data analytics has become a critical success factor for organizations. With the advancement of digital transformation, companies can use data analytics to get the results they want to make strategic decisions, optimize operational processes, and enhance customer experience, ultimately determining the direction of the company. Whether it's market trend forecasting, customer information analysis, or supply chain optimization, data analytics is driving companies to achieve more efficient and competitive business models (Wang, 2024). These phenomena show that data analytics has never been more important in business.

With the rapid advancement of data analytics, organizations have undergone a dramatic transformation in their ability to collect, store, process and apply data. For starters, the rise of big data technologies has enabled organizations to handle unprecedented scales of data and extract valuable information from it. For example, Amazon has used its powerful data analytics to boost sales through personalized recommendation systems (See Figure 1) (Huang, Jiang, Wu, & Wang, 2020).

Second, the application of machine learning and deep learning has enabled companies to make more accurate predictions and decisions. These technologies can automate the processing of large amounts of complex data and identify potential patterns and trends, thus helping enterprises gain an advantage in the fierce market competition. For example, an innovative decision support model "AlphaVision" designed for stock price prediction by seamlessly integrating real-time news updates and Return on Investment (ROI) values, utilizing various machine learning and deep learning approaches (Divyashree et al., 2024).

However, while data analytics technologies present tremendous opportunities, they are also accompanied by a new set of challenges and problems. In the face of these rapidly changing technological and market environments, in-depth analysis and optimization of data analytics technology has become inevitable. Enterprises must not only rely on existing technical means, but also need to comprehensively improve the efficiency and accuracy of data analysis from multiple dimensions, such as data quality, data security, analytical tools and talents, as well as cross-departmental cooperation. Only in this way can enterprises maintain a leading position in the future market competition and

[a] https://orcid.org/0009-0007-5621-2765

Figure 1: Example of Amazon.

continue to realize business growth. Therefore, this paper will deeply explore the key issues of data analytics in business applications and put forward corresponding countermeasures, with a view to providing valuable reference and guidance for enterprises in the data-driven era.

2 PROBLEMS

2.1 Data Quality

In this era of information explosion, data and information from a variety of sources, data types, so when this paper is collecting data there will be incomplete data, data duplication, and even incorrect data. Handle and clean the data is one of the perennial challenges in data analytics (Chu, Ilyas, Krishnan, & Wang, 2016). Therefore, the quality of data is a key determinant of the quality of decisions made (Choughri, et al., 2018). When the data is duplicated, it will cause great errors in data analysis. In life it is also very easy to have data entry errors that lead to incorrect data. When this paper unknowingly enters incorrect information, it can make it difficult to find effective ways to refine and modify the model. In daily life, data entry errors are also a major cause of incorrect data. For example, users may double-enter information, or machines may make mistakes during testing and recognition, all of which can make the quality of the data much worse. And when this paper unknowingly enters this incorrect information into the system, this paper may be stuck in a predicament where it is difficult to figure out effective ways to refine and modify the model. These are all factors that contribute to data quality problems, when the results of the analysis may be biased and lead to poor decision-making.

2.2 Data Leakage

As the amount of data increases, the issue of data privacy and security becomes more and more important. All that is needed for data analytics is a lot of data, and the more data is piled up then the greater the risk of leakage. Chen et al. (2017) discussed relationships revealed in this study. "As the volume of data is growing exponentially and data breaches are happening more frequently than ever before, detecting and preventing data loss has become one of the most pressing security concerns for enterprises."
. So, the more organizations use and store data, the greater the risk of data leakage and unauthorized access. Cyber attackers will invade the enterprise network through malware, phishing attacks, DDoS attacks and other means to steal sensitive data. As technology advances, attackers' tactics become more sophisticated, and traditional firewalls and anti-virus software can no longer fully defend against these advanced threats. In addition, corporate insiders may also cause data leakage through intentional or unintentional behavior. Insiders often have higher privileges and direct access to data, so when they are the source of a breach, the damage is often more severe. A data breach can then lead to data analysis results that may become inaccurate for company decisions or even affect the advancement of data analysis. Not only can this lead to financial losses, but it can also have a serious impact on a company's reputation.

2.3 Inaccurate Data Analysis

The business market is evolving rapidly, so the cost of trial and error for companies is remarkably high. So, the accuracy of data analysis is crucial for companies. Due to the differences in the selection of

data analysis tools and the differences in the professionals, so the problem of inaccuracy may occur in the market and in the decision making of companies using data analysis. When data analysts are not technically literate enough, it may be difficult to extract valid values from the data or perform more complex and large data analysis. When the tools do not match the actual needs, there is a high probability of overfitting and inefficient analysis effects. The impact of these data analysis inaccuracies on the company and the market is huge. Once the error is large it may lead to poor decision making and wrong direction of development of the company, resulting in irreversible consequences.

2.4 Application of Data Analytics

Even if the results and accuracy of the data analysis are good, it is useless when the results of the analysis are not well applied to the market and the company's decisions and actions. If companies do not make good use of data processing and utilization in the market, then they will be eliminated by their competitors (Hu, 2018). Just like our human body, when our brain receives the information and completes the instructions, if our nerve center is not able to transfer the instructions accurately and quickly to all the organs of our body, then our body won't be able to function properly. In reality, when a data analytics scientist analyzes market and company data and gets a result, and the business department is unable to understand the result and distribute it to the executive departments, then the result is not well applied in decision making and action.

3 SOLUTIONS

3.1 Reasonable Processing of Data

For the problem of data quality, what this paper has to do is to improve the quality of data, that is, to make reasonable processing of data. This paper can carry out data cleansing to clean up the missing values and duplicate values in the data set. In the enterprise there should be data analysis scientists according to the different requirements of each project, to establish a set of appropriate data cleaning mechanism, to help in data analysis in the data analysis to improve and optimize the data analysis, regular review of the data, fill in the missing values, remove duplicate data, so as to improve the quality of the data. For example, for missing data, data preprocessing and missing value verification are performed to ensure data integrity.

For duplicate data, data cleansing techniques are used to remove duplicate data and ensure that each piece of data appears only once in the data set. For incorrect data, the end can be achieved by multiple data entry and comparing the correctness. Moreover, the company should also develop a unified standard and process for data processing, so that whenever a new data set is added to the analysis, it will be standardized in this process, maintaining the integrity and consistency of the data, so that the data can be more accurately applied to data processing.

3.2 Encrypting the Data

To solve the data leakage problem, enterprises should strengthen both external protection and internal prevention. First of all, strict access policies and management rights should be formulated so that only core personnel or authorized personnel can access. When accessing and processing important data, you can let multiple top-ranking data analytics scientists have some of the permissions, and only when the majority of scientists are authorized can they legally access it. In this way, even if there is a problem with an insider, it will not affect the entire data leakage. Second, the data can be strictly protected using encryption so that even if the company's data is hacked, it cannot be accessed illegally. This is equivalent to provide a double insurance for the protection of data, for example, this paper can use irreversible encryption algorithm for data processing, which can effectively prevent reverse decryption. Just like the new encryption method of data shown in the study by Eltayieb et al., (2024) the encryption method of CLPRE-CRF is proposed to prevent data leakage and enhance data security. Then this paper also needs to strengthen the risk assessment of data reading to prevent violent decryption to obtain data. Finally, an inspection team can be established, specifically responsible for data protection, to conduct regular data security checks, identify and repair potential security gaps, to ensure that data security measures are always effective. Once a data leakage problem occurs, it will be immediately investigated and solved to ensure that the data leakage problem is solved at the first time.

3.3 Elevation Tools, Talents and Models

To improve the accuracy of data analysis, it is necessary to improve both tools and talents. For the selection of multiple analysis tools, enterprises should choose appropriate analysis tools and technologies according to the project and market clear

needs and goals. In addition, a variety of analysis tools and techniques can be selected to get different results. Multiple small-scale pilot projects can be established based on the different results of the analysis before applying them to the whole project, so as to test out the most accurate and compliant analysis tools. For professionals, companies should increase the education and training of analysts to improve the data analysis ability of employees. In addition, it can also be improved by external factors, such as the introduction of excellent data analysis talents with high technical literacy, and the cooperation with professional data analysis organizations to complete the project. By further improving the tools and talents, the accuracy of data analysis will be greatly guaranteed. In addition, the accuracy can be ensured by adjusting the model itself. Enterprises can also continuously optimize the analysis model and assess the accuracy and applicability of the model through the verification of historical data (Vanwinckelen, 2012). For example, previous historical data can be used as a validation set to observe the fitting effect of the analytical model, so as to adjust and analyze the accuracy of the model. The accuracy of the prediction can also be improved by using several different models for prediction and analyzing the results comparatively. In conclusion. The prediction of the accuracy of data analysis must not be too high, when the enterprise is making decisions, it is necessary to consider the uncertainty of the prediction results and reduce the risk of decision-making by developing a variety of response options and risk management strategies.

3.4 Enhancing Collaboration Across Sectors

Firstly, this paper must focus on strengthening the synergy between all parts of the organization. This means strengthening the communication and collaboration between the data analysis team and the business units to ensure that the analysis results are highly aligned with the actual business needs. In order to achieve this goal, this paper should establish a " double-check " working concept, i.e., in every project, the business department and the data analysis department should communicate and report to each other twice. This minimizes errors and ensures that both parties have the same understanding and expectations of the project. Through frequent communication and feedback, this paper is able to identify and resolve potential problems in a timely manner, thereby improving the quality and efficiency of the project. Second, to avoid errors due to poor

presentation and understanding, this paper should present data and results in a more visual way. This can be done through the use of charts, documents and other forms of presenting complex data processing processes and results so that other departments can understand and apply the information in a more intuitive way. The impact of visualization in big data analysis is very important for business. It is a technique of displaying information in a graphical format that makes it easy for users to understand the information (Mahajan, & Gokhale, 2017). For example, this paper can create data dashboards, reports, and visual models that present key metrics and trends to business units in a clear and understandable way. As a result, task assignment and execution will become more efficient and precise because everyone will be able to make decisions based on the same understanding and data. Finally, before each project acceptance, this paper should convene a meeting of all project department heads to provide joint feedback on the results and their application in decision-making and action, and to understand the specific progress so far. During the meeting, departments can share their experiences and findings, make suggestions for improvement, and adjust data analysis and optimization strategies based on feedback from other departments. This cross-departmental cooperation and communication will help ensure smooth progress and provide valuable lessons learned for future projects. By working together, this paper can continue to improve the quality and value of data analysis and provide strong support for business development. In conclusion, through enhanced collaboration, visual data presentation and cross-departmental feedback mechanisms, this paper can improve the accuracy and usefulness of data analysis to better support business decisions and actions. This will help enterprises gain advantages in the competitive market and realize sustainable development.

4 CONCLUSIONS

In the current business environment, data analytics has undoubtedly become a core tool for enterprises to gain competitive advantage. However, data analytics faces many challenges in its practical application. Through in-depth analysis of four major problems: data quality, data leakage, inaccurate data analysis, and data analysis application, this study proposes targeted solutions, including the establishment of a sound data processing mechanism, the enhancement of data encryption technology, the improvement of

the match between analytical tools and talents, and the promotion of cross-departmental collaboration. These countermeasures can help enterprises deal with various problems in data analysis more effectively and ensure the accuracy and security of data, so that they can develop better with the support of data analysis.

Future research directions can focus on the innovation and practical application of data analysis. Firstly, with the continuous development of artificial intelligence and machine learning technology, the technology of data analysis is also improving, how to better apply these technologies in data analysis to improve the accuracy of prediction and rationality of decision-making will be one of the important future research directions. Secondly, better application of data analytics to company development and decision making, more accurate as well as rational application in business is also a very important research direction in the future.

REFERENCES

Cheng, L., Liu, F., & Yao, D. (2017). Enterprise data breach: causes, challenges, prevention, and future directions. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7(5), e1211.

Choughri, R., Kamaleddine, K., Soubjaki, M., & Baytieh, M. (2018). The Challenge of Data Accuracy in Business Analytics that Affect Managers Decision Making–Case Study of Saudi Arabia & Lebanon. IOSR Journal of Business and Management (IOSR-JBM).

Chu, X., Ilyas, I. F., Krishnan, S., & Wang, J. (2016, June). Data cleaning: Overview and emerging challenges. In Proceedings of the 2016 international conference on management of data (pp. 2201-2206).

Divyashree, S, Christy Jackson Joshua, Abdul Quadir Md, et al. (2024). Enabling business sustainability for stock market data using machine learning and deep learning approaches. Annals of Operations Research (prepublish), 1-36.

Eltayieb, N., Elhabob, R., Abdelgader, A. M., Liao, Y., Li, F., & Zhou, S. (2024). Certificateless Proxy Re-encryption with Cryptographic Reverse Firewalls for Secure Cloud Data Sharing. Future Generation Computer Systems.

Hu, Y. (2018). Marketing and business analysis in the era of big data. American Journal of Industrial and Business Management, 8(07), 1747.

Huang, C., Jiang, W., Wu, J., & Wang, G. (2020). Personalized review recommendation based on users' aspect sentiment. ACM Transactions on Internet Technology (TOIT), 20(4), 1-26.

Mahajan, K., & Gokhale, L. A. (2017). Significance of digital data visualization tools in big data analysis for business decisions. International Journal of Computer Applications, 975, 8887.

Vanwinckelen, G., Blockeel, H., De Baets, B., Manderick, B., Rademaker, M., & Waegeman, W. (2012). On estimating model accuracy with repeated cross-validation. In BeneLearn 2012: Proceedings of the 21st Belgian-Dutch Conference on Machine Learning (pp. 39–44).

Wang, Yitong. (2024). An Introduction to the Application of Big Data Analytics Technology in Business Intelligence. Engineering and Management Science (3).