Data Science and Big Data Technology, 2025, 1(1), 10-13
https://www.hillpublisher.com/Journals/dsbdt/
ISSN Online:

Discussion on Data Mining Technology of Software Engineering

Tianrui Xia*, Yue Liu, Li Wang

Wuhan Donghu College, Wuhan 430212, Hubei, China.

How to cite this paper: Tianrui Xia, Yue Liu, Li Wang. (2025). Discussion on Data Mining Technology of Software Engineering. *Data Science and Big Data Technology*, 1(1), 10-13.
DOI: 10.26855/dsbdt.2025.12.003

Received: September 22, 2025
Accepted: October 16, 2025
Published: November 14, 2025

***Corresponding author:** Tianrui Xia, Wuhan Donghu College, Wuhan 430212, Hubei, China.

Abstract

Based on the current software engineering data mining technology problems, this paper puts forward targeted measures to strengthen the data mining processing of software maintenance, so as to maximize the application effect of software engineering data mining technology. In the concept and practical application of software engineering, its own application status is relatively common, which can be regarded as a rich and scientific major measure. Therefore, in the new era, data mining technology in today has a very great significance in the current era.

Keywords

Software engineering; Data mining; Mining technology; Data source

With the advent of the information and technological age, the application of various high technologies can not only adapt to the development needs of various industries in the new era, but also effectively collect and utilize existing data. A comprehensive analysis of the current situation reveals that the collection and application of various data is gradually gaining importance across the entire industry. To achieve this goal, it is necessary to extract target data from existing data to support it, and a complete data extraction process from this data can be called data mining. Currently, its application is quite widespread in the concept and practice of software engineering, and can be regarded as a significant and scientific measure. Therefore, in this new era, data mining technology has extremely important significance (Chen & Zhu, 2022).

1. Application of Data Mining Techniques in Software Engineering

In today's society, with the rapid development of science and technology, various new technologies have been widely used. In particular, the Internet and information technologies have become an indispensable part of development, bringing more convenience to users (Liu, 2021).

1.1 Data Acquisition

Data mining extracts useful data from massive data resources; this mining and sifting is a crucial step. In a sense, data mining is a method for verifying driving data; its goal is to effectively analyze various types of data from their essence, thus changing current data-driven approaches. When validating and analyzing driving data, users must objectively analyze hypothetical information and continuously collect data in a targeted manner. Applying this method scientifically and rationally helps verify the validity of hypotheses. Currently, in relation to reality, data and information storage exhibit complex characteristics and are massive in scale. Faced with this situation, it's difficult to efficiently identify and mine the entire database using only a confirmatory analysis method. Therefore, when driving data is proposed and applied, it can perform deep mining and filtering, automatically identifying certain implicit and important information. In data mining, the main goal of collecting and using various types of data is to continuously

improve and innovate its products. Therefore, in information collection, various software measurement methods should be used scientifically and rationally to maximize their practical application and value in actual work (Liu, 2021).

1.2 Software Engineering Metrics

In the development and practical application of software engineering, it plays a crucial role in the development and application of various projects. Through a comprehensive study of the current development status of software developers, it can be found that in terms of overall and specific software monitoring, the quantification of indicators is gradually being realized, thereby promoting users' detailed analysis and understanding of the entire development process. Currently, quantitative data is widely used, and its overall usage level is relatively high; however, its analysis techniques and the operation of the entire system still face many problems. Generally speaking, to ensure the smooth progress of software design and data acquisition, it is necessary to organically link it with data processing. At this time, the characteristics of these symbols are diverse, and most of the time, they are used statically, and they also change over time. For example, most products can be reasonably evaluated in practice regarding product quality and production processes, but rarely can this be used as a scientific and reasonable standard and basis for judgment; therefore, the overall evaluation of the evaluation work has little substantial value (Lu, 2022).

2. Current Problems Encountered by Data Mining Techniques in Software Engineering

2.1 The Complexity of Existing Software Engineering Data

Research on data mining techniques in current software engineering reveals that data processing primarily involves two parts: structured data and partially unstructured data. In practice, structured data refers to version information, defect reports, etc., while unstructured data refers to various documents, data codes, etc. In reality, these two types of data cannot be processed and parsed using the same methods. These two aspects are closely related. For example, while the vulnerability in this report has a significant impact, it is itself structured data and can be categorized as unstructured data. This demonstrates that, in the overall context of software engineering, erroneous reports play a crucial role in its practical application (Wang, 2021).

2.2 Serious Shortcomings of Unconventional Methods

Data mining is a crucial component of any work process. It fundamentally ensures data transfer and maximizes the effective sharing of data resources in practical applications, thereby improving efficiency. In the past, many data mining methods were available for implementation, but they were typically presented in textual and graphical formats. From a practical perspective, software developers have high demands for various types of data. Due to the specific nature of these methods, they haven't been effectively applied in practice. Therefore, to ensure the system's effective execution and maximize its functionality, it's necessary to comprehensively analyze all relevant information and data, organically linking them to the actual situation and system limitations. This approach not only enhances the availability and utilization of existing data but also fundamentally strengthens existing knowledge and information, thereby improving the authenticity, timeliness, and effectiveness of data mining (Zhang, 2022).

2.3 Failure to Achieve a Consistent Assessment of Software Engineering Mining Results

Traditional mining techniques are often constrained by traditional ideas and working methods in practice, which leads to many problems in practice. However, in practice, due to its long usage cycle, a relatively complete set of evaluation criteria and requirements has been established within a certain range. From a comprehensive perspective, developers need to make full use of the integrity of information during the information mining process. However, due to the complexity and abstractness of information itself, it is difficult to show diversified characteristics, making it difficult for the entire system to work efficiently and achieve true organic integration [5]. In this case, it cannot be guaranteed whether the final evaluation value can be reflected in reality. Because there is often a large amount of data in software engineering, and this data is very complex, the demand on programmers is greater. Therefore, it is necessary to objectively analyze the various types of data involved according to actual needs and specific requirements, and to properly handle related explanations, code, and other issues. Otherwise, the complexity of technical work will be difficult to control.

3. Specific Applications of Data Mining Techniques in Software Engineering

3.1 Improving the Application of Data Mining Techniques in the Software Development Process

In the design and development of software engineering, it is essential to fully understand the significance and value of data mining technology. To ensure the effective utilization of data, attention must be paid to programming and the use of program results. First, during program development, a targeted search and parsing process should be conducted based on the program structure and functional technology to obtain relevant information. Based on this, the most suitable model for real-world needs can be selected from relevant coding resources, thereby enhancing the functionality of the entire search system. This method has been widely used in practice, reducing human resources required for actual work, enabling efficient identification of various types of information based solely on keywords. Simultaneously, it is integrated with the current user's real-world situation to achieve intelligent filtering, with the basic goal of continuously reinforcing information in real life. Second, static basic principles should be established in practice and applied scientifically and rationally, paying attention to inheritance and overloading to ensure the correct recording of various data and guarantee its accuracy and effectiveness. On this basis, existing recursive checks should be appropriately applied to better divide the work and maximize the correlation between different data. This approach ensures efficient classification of existing data and provides valuable reference for software developers' decision-making. Furthermore, it's crucial to adhere to dynamic principles, correctly utilize various functions, and ensure the proper use of mining objectives, making static coding more scientific and rational. This allows for the storage and replication of various data in the database, identifying errors and improving the reliability and effectiveness of data mining. When adjusting and optimizing compilation results, erroneous code can be reasonably eliminated, following overall development trends. In practice, it's essential to proactively adopt targeted methods based on the actual situation to locate, test, and analyze existing problems. During testing, it's crucial to ensure the accuracy and efficiency of the results to promptly identify issues and ensure the effective use of data.

3.2 Appropriate Application of Data Mining Techniques in Software Maintenance

In the routine maintenance of various software systems, the application of data mining techniques has a significant impact and role. Based on this, using data mining techniques for software repair improves system efficiency. By objectively analyzing existing problems and providing corresponding solutions, the overall functionality and features of the system are realized, achieving ideal application. To achieve better maintenance results, it is necessary to start from real-world needs, objectively analyze problems, conduct comprehensive assessments, evaluate specific maintenance levels, and select the most suitable maintenance route for practice. When assigning errors, the current textual data is usually converted first, and then a set of error-specific processing is applied. However, in practice, this method's overall results are unsatisfactory, and accuracy is not improved. Therefore, in specific implementation processes, enhanced testing techniques are often used to improve and optimize existing fault reporting techniques. Its basic goal is to ensure that natural language and additional implementation information have a high degree of similarity, thereby accurately, objectively, and reasonably describing the defects of existing software and promoting overall software repair. In the process of optimizing and improving the software system, it is necessary to clearly define the software design and programs, and appropriately utilize languages with different non-programming characteristics to achieve quantitative results. During the understanding of the system, objective analysis of the document design and entries can be conducted to maximize the protection of the overall system performance. Simultaneously, it is also essential to ensure the normal operation of the system and prevent massive data congestion.

4. Conclusion

Although the development and application of data mining techniques in software engineering have gradually progressed, many issues still exist that require improvement in practice. Therefore, we need to utilize data mining techniques as much as possible in software development and maintenance to ensure its long-term and stable development.

References

Chen, Q., & Zhu, X. (2022). Exploration of the Application Practice of Data Mining Technology in Software Engineering. *Information Recording Materials*, *23*(02), 148-150.

Liu, Y. (2021). Discussion on data mining technology in software engineering. *Computer Knowledge and Technology*, *17*(35), 70-71.

Liu, Z. (2021). Analysis on the application of data mining technology in software engineering. *Computer Knowledge and Technology*, *17*(24), 27-28.

Lu, S. (2022). Application Analysis of Data Mining Technology in Software Engineering. *Yangtze Information and Communication*, *35*(04), 113-118.

Wang, B. (2021). Research on the Application of Data Mining Technology in Software Engineering in the Era of "Internet+". *Journal of Beijing Institute of Printing*, *29*(04), 148-151.

Zhang, X. (2022). Research on data mining technology in software engineering. *Network Security Technology and Application*, (04), 43-44.