Association Rule Mining and Application of Consumer Behavior in Social Network Big Data

Wen Wen*

Department of International Business Administration, Woosong University, Daejeon, 34606, Korea
Corresponding author

Abstract: *This article takes consumer behavior in social network big data as the research object, systematically sorts out the relevant theories of social network big data, consumer behavior, and association rule mining, and constructs a consumer behavior association rule mining model adapted to social network scenarios. Firstly, preprocess social network consumer behavior data by collecting, cleaning, integrating, and reducing it; Secondly, optimize the FP Growth algorithm to address the drawbacks of traditional algorithms such as low efficiency and redundant rules when dealing with massive sparse data; Subsequently, complete frequent itemset mining, association rule generation, and filtering according to standardized procedures; Finally, application strategies are proposed in three major areas: precision marketing, product and service optimization, platform operation, and risk management, along with a three-dimensional support system of technology, data, and talent. Research has shown that the optimized FP Growth algorithm significantly improves the efficiency and accuracy of mining, and can effectively mine strong association rules such as consumer browsing interaction purchase, sharing add purchase, etc; The constructed mining model is adapted to the needs of social network scenarios, and the mining results can provide data support and practical references for the precise operation of enterprises and the sustainable development of social platforms. This article enriches the cross disciplinary achievements of social network big data and consumer behavior research, and improves the application system of association rule mining in non-transactional scenarios.*

Keywords: *social network big data; consumer behavior; association rule mining; FP Growth algorithm; precise operation*

1. Introduction

In the context of the deep development of the digital economy, social networks have become the core carrier connecting consumers and the market. The global social media user base has exceeded 4 billion, and its derived big data covers multi-dimensional behavioral trajectories such as consumer browsing, interaction, and purchase, forming the core data source for analyzing consumer decision-making logic. The interactivity and dissemination of social networks have reconstructed consumer behavior patterns, shifting consumer decision-making from individual independent judgment to social association driven. However, the massive, sparse, and unstructured behavioral data also makes it difficult to capture implicit consumer needs and behavioral associations. Traditional marketing models and data processing methods are no longer suitable for the demands of the new era market, becoming a key bottleneck restricting the precise operation of enterprises.

Association rule mining, as one of the core technologies in data mining, can effectively reveal the hidden itemset associations in datasets, providing a feasible path for solving the dilemma of consumer behavior analysis in social networks. At present, scholars at home and abroad have conducted relevant research. Foreign research focuses on algorithm optimization and interdisciplinary applications, while domestic research focuses on specific platform cases [1]. However, existing achievements still have problems such as insufficient scenario specificity, single mining dimensions, and a disconnect between theory and practice. They have not fully adapted to the special characteristics of social network behavior data, making it difficult to achieve effective implementation of mining results.

Based on this, this article takes "Mining and Applying Association Rules of Consumer Behavior in Social Network Big Data" as the research theme, which has both theoretical and practical value. In theory, it can enrich the cross disciplinary achievements of social network big data and consumer behavior research, and improve the application system of association rule mining in non-transactional scenarios; In practice, implicit correlations in consumer behavior can be explored to provide data support for

precision marketing and platform operation optimization of enterprises. This article will use literature research method, data analysis method, empirical research method, etc. to sort out consumer behavior characteristics and data types, construct an association rule mining model adapted to social network scenarios, verify the effectiveness of the model through empirical verification, and finally propose targeted application strategies, in order to provide reference for high-quality development of enterprises in the context of digital economy.

2. Relevant theoretical foundations

2.1 Social network big data related theories

Social network is a virtual platform built based on Internet technology and used to realize information interaction and relationship connection between individuals. Its core features are interactivity, communication and personalization [2]. The mainstream types include social e-commerce platforms, content sharing platforms and social interaction forums. User behavior data of different platforms show the characteristics of travel alienation. Social network big data is a collection of massive user behavior trajectory data generated during the operation of social platforms. Its core content includes the diversity of data sources, the complexity of data forms, and the potential value of data, specifically covering user basic information, social interaction data, content browsing data, and consumption behavior data.

Compared to traditional data, social network big data has five core characteristics: massive, diverse, high-speed, valuable, and sparse. Its data processing flow requires four key steps: collection, preprocessing, storage, and analysis. Among them, the preprocessing stage of cleaning, integration, transformation, and reduction is the core prerequisite for solving the problem of adaptability to unstructured behavioral data and improving subsequent mining efficiency, providing a high-quality data foundation for association rule mining.

2.2 Theories related to consumer behavior

The traditional consumer behavior theory is centered around the AIDA model and the planned behavior theory [3]. The former focuses on the entire decision-making process of consumer attention, interest, desire, and action, while the latter emphasizes the influence of individual attitudes and subjective norms on consumer behavior. However, neither of them fully considers the intervention effect of social associations in social network scenarios on consumer decision-making, which has certain limitations. In the social network environment, consumer behavior patterns have undergone significant restructuring, presenting core characteristics of interactivity, conformity, fragmentation, and immediacy. Their information acquisition relies on social communication, and the decision-making process is influenced by the interaction between opinion leaders and groups. Their behavior shows a trend of online and offline integration.

From the perspective of influencing factors, consumer behavior is mainly influenced by four major factors: individual, social, platform, and external factors. Its behavior data can be divided into four categories: browsing, interaction, purchase, and retention. Quantitative analysis of various types of data can accurately capture consumer behavior preferences and internal needs, and clarify the core analysis objects for association rule mining.

2.3 Relevant theories of association rule mining

Association rule mining is one of the core technologies in the field of data mining, with the core goal of mining implicit association relationships between itemsets in a dataset [4]. The core concepts include itemsets, support, confidence, and elevation, where support reflects the frequency of itemsets, confidence reflects the reliability of association rules, and elevation is used to screen strong association rules with practical significance (elevation>1 is considered effective association). Traditional association rule mining algorithms are represented by Apriori algorithm and FP Growth algorithm. The former generates association rules by iteratively filtering frequent itemsets, which has a simple principle but low mining efficiency. The latter improves mining efficiency by constructing FP trees to avoid frequent scanning of datasets. However, both algorithms have limitations in dealing with the sparsity and unstructured features of social network big data, such as insufficient mining accuracy and a large number of redundant rules.

Based on this, it is necessary to optimize traditional algorithms in combination with the particularity of social network consumer behavior data, so as to adapt them to the processing requirements of

unstructured data, improve the efficiency and accuracy of association rule mining, and provide reliable technical support for consumer behavior association analysis.

2.4 Cross application logic of relevant theories

The theory of big data in social networks provides a data foundation and processing methods for research, clarifies the sources, characteristics, and processing flow of consumer behavior data, and solves the core problem of "what data to mine"; The consumer behavior theory defines the behavioral characteristics, influencing factors, and data classification of the research object, clarifies the core direction of association rule mining, and solves the problem of "which associations to mine"; The theory of association rule mining provides core technologies and analytical methods, solving the problem of "how to mine associations"[5]. The three are mutually supportive and organically integrated, forming a cross application system of "data object technology". Social network big data is the foundation of research, consumer behavior is the core object of research, and association rule mining is the core technology of research. The deep integration of the three can achieve precise mining of implicit associations in social network consumer behavior, providing comprehensive theoretical and technical support for subsequent empirical analysis and application strategy implementation.

3. Construction of association rule mining model for consumer behavior in social network big data

3.1 Preprocessing of consumer behavior data on social networks

Data preprocessing is the premise and foundation of association rule mining, with the core goal of improving data quality, reducing data redundancy, and converting raw unstructured and heterogeneous social network consumer behavior data into standardized data formats that are compatible with mining algorithms [6]. Based on the characteristics of social network data, the preprocessing process mainly includes four key steps: data collection, cleaning, integration and transformation, and reduction.

The data collection adopts legal and compliant methods, selecting mainstream social e-commerce and content sharing platforms as data sources. The collection scope includes multi-dimensional behavioral data such as user basic information, browsing trajectory, social interaction (likes, comments, shares), consumption conversion (add purchases, purchases), etc. The data representativeness and integrity are guaranteed by combining API interfaces with public datasets. Data cleaning focuses on handling missing values, outliers, and duplicate values, using methods such as mean imputation and deletion of outlier samples to eliminate invalid data such as robot interactions, and performing structured transformations such as word segmentation and encoding on unstructured data such as text comments and image tags. Data integration and transformation integrate heterogeneous data from multiple sources, map consumer behavior data to the transaction item set required for mining, and standardize data quantification standards through standardized processing. Data reduction eliminates irrelevant features through feature selection, reduces data dimensionality and redundancy, and provides guarantees for improving subsequent mining efficiency.

3.2 Optimization and selection of association rule mining algorithms

Combining the massive and sparse characteristics of social network big data, traditional association rule mining algorithms can no longer meet the requirements of mining efficiency and accuracy. Therefore, algorithm optimization is needed to achieve scene adaptation. At present, the mainstream association rule mining algorithms are Apriori algorithm and FP Growth algorithm, each with its own limitations: Apriori algorithm generates rules by iteratively filtering frequent itemsets, which is simple in principle but requires multiple scans of the dataset, resulting in low mining efficiency; The FP Growth algorithm improves efficiency by constructing FP trees to avoid duplicate scanning, but there are still problems such as excessive redundant rules and insufficient mining accuracy when processing unstructured behavioral data.

Based on this, this article selects the FP Growth algorithm as the basic model for optimization, introduces multidimensional tree structure and parallel computing technology to enhance the algorithm's ability to handle massive sparse data; Combining the correlation characteristics of consumer behavior, a similarity threshold filtering mechanism is added to eliminate invalid redundant rules and improve the accuracy of rule filtering. Through algorithm simulation experiments, it was verified that the optimized FP Growth algorithm outperforms traditional algorithms in terms of running efficiency and rule accuracy.

It can effectively adapt to the mining needs of social network consumer behavior data and is ultimately determined as the core mining algorithm in this paper.

3.3 Implementation process of association rule mining

Association rule mining follows a standardized implementation process of "data preparation frequent itemset mining rule generation and filtering rule validation" to ensure the standardization of the mining process and the reliability of the results. Firstly, during the construction phase of the transaction dataset, the preprocessed consumer behavior data is transformed into a transaction item set, and multiple behaviors of a user in a single session are defined as one transaction, with each behavior being an item, completing the final preparation for mining data.

Secondly, based on research scenarios and industry experience, the trial and error method and statistical analysis method are used to determine reasonable minimum support and minimum confidence thresholds. The improvement index (improvement>1) is introduced to screen strong association rules with practical significance, balancing the number of rules with practicality. Again, the optimized FP Growth algorithm is used to mine frequent itemsets in the transaction set, and non-frequent itemsets are iteratively removed to generate preliminary association rules based on the frequent itemsets. Finally, calculate the support, confidence, and improvement of each rule, select strong association rules that meet the threshold requirements, and combine consumer behavior theory to provide semantic interpretation and rationality verification of the rules, ensuring that the mining results conform to the actual scene rules.

3.4 Overall framework of mining model

The social network consumer behavior association rule mining model constructed in this article integrates four core modules: data preprocessing, algorithm optimization, rule mining, and verification, forming a closed-loop mining framework. The model takes social network consumer behavior data as input and outputs a standardized transaction dataset through a preprocessing module; Implement frequent itemset mining and association rule generation through optimized FP Growth algorithm module; Effective strong association rules are output using the rule filtering and verification module;The final output of the mining results provides data support for subsequent empirical analysis and application strategy proposals.

This framework clearly presents the input, processing, and output relationships of each module, achieving deep adaptation of data, algorithms, and scenarios. It not only solves the adaptability problem of traditional models in social network scenarios, but also ensures the reliability and practicality of mining results. It constructs a complete technical system of "preprocessing algorithm optimization rule mining verification", connects the theoretical basis of the previous text with the empirical research of the following text, and provides standardized framework support for accurate mining of implicit associations in consumer behavior.

4. Application strategy of consumer behavior association rules in social network big data

4.1 Application in the field of precision marketing

Precision marketing is the most direct application scenario for association rule mining results, with the core being based on consumer behavior association characteristics, to achieve precise allocation of marketing resources and efficient matching of consumer demand, and to solve the inefficient dilemma of traditional marketing "casting a wide net". Combining the mined association rules of browsing interaction purchase, social sharing consumption conversion, etc., we will focus on promoting the application from three aspects.

In terms of personalized recommendation, a recommendation model based on association rules is constructed to push highly related product content and services to users based on their historical behavior trajectory and association characteristics. For example, for users who frequently browse beauty content and like and evaluate it, relevant beauty products and similar evaluation content are synchronously recommended to improve recommendation accuracy and user acceptance. In terms of marketing content push, the timing and form of push are determined based on user behavior association rules. For users with strong "browsing sharing" association behavior, push marketing content that is easy to spread, practical and interesting, and expand the dissemination scope through social fission. In user segmentation and marketing, consumers can be accurately segmented based on the results of association rule mining.

Differentiated marketing strategies can be developed for various groups such as high-frequency interaction users and browsing-conversion users, thereby achieving precise and efficient marketing outreach.

4.2 Application in the field of product and service optimization

Association rule mining can reveal the potential needs and pain points behind consumer behavior, provide data support for product design iteration and service process optimization, and help enterprises improve product competitiveness and user satisfaction. Based on the analysis of consumer behavior associations, we will focus on promoting three optimization tasks.

In terms of product design and iteration, by mining the association rules between product browsing, commenting, and purchasing behavior, potential demands of users for product functionality, appearance, and other aspects are captured. For example, analyzing the association characteristics between browsing and negative comments of a certain type of product provides direction for optimizing product functionality. At the same time, combined with user co creation mode, the matching degree between product and user needs is improved. In terms of service process optimization, user service needs can be predicted based on the association rules of inquiry, purchase, and after-sales behavior. Service response processes can be optimized; for example, relevant after-sales guides and exclusive services can be pushed in advance to users who have completed purchases to improve their service experience. In terms of social interaction experience optimization, platform social functions can be optimized by combining association rules of user interaction behavior. Comment and sharing mechanisms can be improved to enhance user engagement, as well as strengthen user stickiness and sense of belonging to the platform.

4.3 Application in platform operation and risk management field

The healthy operation of social networking platforms relies on data-driven refined management. The results of association rule mining can be used to optimize platform traffic configuration, prevent operational risks, and achieve the coordinated development of content ecology and consumer ecology. The specific applications are mainly reflected in three dimensions.

In terms of platform traffic optimization, the content distribution algorithm is optimized based on user behavior association rules, combining high-frequency associated content with products for display, improving traffic conversion rates and user retention time. At the same time, targeted traffic retention strategies are formulated based on user retention behavior association characteristics to stabilize the platform's user base. In terms of risk identification and control, we will explore the association rules of abnormal behaviors such as false interactions and malicious brushing, construct a risk identification model, achieve accurate monitoring and timely control of abnormal behaviors, standardize data collection and usage processes, and prevent algorithm bias and user privacy leakage risks. In terms of optimizing platform ecology, we will balance content ecology and consumption ecology, based on the characteristics of user behavior associations, avoid excessive marketing interference on user experience, and build a healthy and sustainable social consumption ecology.

4.4 Application guarantee measures

To ensure the smooth implementation of association rule application strategies and fully leverage the practical value of mining results, it is necessary to build a comprehensive application guarantee system from the three dimensions of technology, data, and talent, and solve various bottleneck problems in the application process. In terms of technical support, we will strengthen the research and application of big data processing and association rule mining technologies, optimize algorithm performance, build efficient data storage and analysis platforms, and improve the efficiency and accuracy of data processing and mining. In terms of data protection, a legal and compliant mechanism for data collection, storage, and usage should be established. User privacy should be strictly protected, the authenticity, integrity, and security of data should be ensured, and data sharing and circulation procedures should be standardized to prevent risks of data leakage. In terms of talent security, we aim to cultivate composite talents with abilities in social network analysis, big data mining, and consumer behavior research. We will strengthen cooperation with research institutions, promote technological innovation and talent cultivation, and provide solid talent support for the implementation of application strategies.

5. Conclusion

This article conducts a systematic study on the mining and application of association rules for consumer behavior in social network big data. It integrates relevant theories of social network big data, consumer behavior, and association rule mining, constructs an association rule mining model that is suitable for social network scenarios, verifies the effectiveness of the model through empirical analysis, and proposes targeted application strategies.

The research conclusion mainly includes four points: firstly, it clarifies the five core characteristics of social network big data and the interactive and conformist behavior patterns of consumer behavior, sorts out the four major influencing factors of individuals, social, platforms, and external factors, and divides the core behavior data into browsing, interaction, purchase, and retention, laying a solid foundation for association rule mining. The second is to construct and validate an association rule mining model based on optimized FP Growth algorithm. The optimized algorithm effectively solves the drawbacks of low efficiency and redundant rules in traditional algorithms when processing massive sparse data, and significantly improves mining accuracy and efficiency. Thirdly, through empirical analysis, multiple consumer behavior strong association rules with practical application value were screened, revealing the implicit correlation logic between browsing interaction purchase, sharing purchase and other behaviors, and verifying the rationality of the research hypothesis. The fourth is to propose application strategies and a three-dimensional guarantee system of technology, data, and talent in three major areas: precision marketing, product and service optimization, platform operation, and risk control, achieving a deep connection between mining results and practical scenarios.

However, there are still certain shortcomings in this study: at the data level, the sample scenarios and dimensions have limitations, and do not cover data on multiple types of social platforms and user emotions; At the model level, the temporal nature of consumer behavior has not been fully considered, and the ability to mine implicit weak association rules needs to be improved; At the application level, the strategy leans towards macro guidance and lacks specific enterprise implementation case support. Future research will address the aforementioned shortcomings by expanding the scope of data samples, enriching data dimensions, optimizing models through time-series mining and deep learning techniques, refining application strategies through specific enterprise cases, exploring the application of association rules in more emerging scenarios, and further improving the theoretical and practical system of social network consumer behavior association mining. This will provide more targeted references for precise operation of enterprises and sustainable development of social platforms in the context of the digital economy.

References

[1] Cui Y, Bao Z Q. Survey of association rule mining [J]. Application Research of Computers,2016, (2):330-334.

[2] Qiao X D, Wang Z. Community discovery of public opinion social network based on improved label propagation algorithm [J]. Application Research of Computers,2025, (1):48-55.

[3] Xiao Y Q, Wang J Q, Lyu X F. Improvement of the Household Production Function to the Traditional Theory of Consumer Behavior and Its Extending Method [J]. Journal of Beijing Jiaotong University (Social Sciences Edition),2012, (4):1-11.

[4] Li W H, et al. Research and Application of AI Association Rule Mining Technology Based on StarRocks [J]. Digital Communication World,2025, (11):83-86.

[5] Cui S J, Bin S, Sun G X, Gao D M. Public Opinion Propagation Model Based on Big Data Analytics in Multiple Relationships Social Network [J]. Journal of South-Central University for Nationalities,2018, (2):114-120.

[6] Li Y, Liu X Y, et al. Smart grid forecasting method based on data preprocessing and Bi-LSTM [J]. Electrical Measurement & Instrumentation,2025, (6):120-125.