ISSN 2616-7433 Vol. 6, Issue 11: 15-20, DOI: 10.25236/FSST.2024.061103

Research on Audience Behavior Pattern Recognition and Prediction Based on Big Data Analysis

Zian Xu

Television School, Communication University of China, Beijing, 100024, China

Abstract: *The era of big data brings new opportunities for audience behavior pattern identification and prediction. This study applies big data analytics methods, such as machine learning and deep learning, to dig deeper into viewer behavior patterns and improve the accuracy of personalized recommendations. A prediction model based on time series analysis is designed, and the value of big data methods in precision marketing and service optimization is confirmed through evaluation and optimization. Meanwhile, the study faces challenges such as data privacy, quality and model generalization. Taken together, this study provides new perspectives on theory and application, and looks forward to a promising future for the application of big data in audience behavior research.*

Keywords: *big data; audience behavior patterns; predictive models; personalized recommendation*

1. Introduction

With the rapid development of information technology, the era of big data has quietly arrived, with far-reaching impacts on various industries, especially significantly promoting the identification and prediction of audience behavior patterns. This change not only enables enterprises to accurately grasp the market demand and optimize the service experience, but also deepens the understanding of user needs, which in turn improves the user experience and optimizes the content delivery strategy to ensure that the information accurately reaches the target audience. As a result, enterprises are equipped with accurate market navigators and have an advantage in the fierce competition in the market. The role of big data in building enterprise user profiles accelerates the development of personalized services and marketing strategies, and significantly increases the business value of audience behavior pattern identification and prediction[1]. This study explores the audience behavior pattern identification and prediction methods in the context of big data, aiming to provide valuable reference and guidance for enterprise practice, unveil the mystery of audience behavior through refined data analysis, reveal its intrinsic laws, and provide solid support for enterprises to formulate more accurate marketing strategies and service directions.

2. Big data and audience behavior analysis foundation

2.1 Big data concept and characteristics

Big data, a term that has now permeated all walks of life, refers not only to the sheer volume of data, but more importantly to its characteristics of fast processing speed and diverse data types. The core of big data lies in its ability to provide enterprises with unprecedented insights and decision-making support, and as some cutting-edge views point out, big data is gradually becoming a core resource for enterprise competition. The application of Big Data is particularly significant in the area of audience behavior analysis. Through advanced analytics, companies can mine audience behavioral patterns and preference changes from massive amounts of data, thus providing strong support for precision marketing and service optimization. In his book *Python data analytics*[2], Fabio Nelli provides an in-depth introduction to big data processing techniques, further highlighting the huge potential and value of big data in audience behavior analysis (As shown in table 1).

ISSN 2616-7433 Vol. 6, Issue 11: 15-20, DOI: 10.25236/FSST.2024.061103

Table 1: Big data vs. traditional data in terms of key characteristics[3]

Feature Dimension	Big data characteristics	Traditional data characteristics
Storage Methods	Requires distributed storage system or cloud computing platform support	Stored on a single computer or in a small database
Data sources	Multi-channel sources such as social media, Internet of Things (IoT) devices, enterprise systems, etc.	Relatively homogenous source, mainly internal enterprise data
Methods of analysis	Requires complex data mining, machine learning and artificial intelligence algorithms	Primary use of statistical analysis methods and database query language
application scenario	Data mining, predictive analytics, personalized recommendations, real-time decision making, etc.	Transaction processing, report generation, simple data analysis, etc.
Data quality challenges	May contain noise, missing values, outliers, etc., requiring data cleaning and preprocessing	Relatively high data quality and low preprocessing requirements
Technical Challenges and Opportunities	Technical challenges in data capture, storage, processing and analysis, but potential for great value	Less technically challenging, focusing primarily on data management and simple analytics

2.2 Sources of audience behavior data

When exploring audience behavior analysis, data collection is a core aspect. In recent years, with the rise of video platforms, they have become an important source of audience behavior data. Relevant studies have shown that video platforms not only record users' viewing history and interactive behavior, but also contain rich information about users' preferences and habits, which provides a valuable data base for in-depth analysis of audience behavior. The collection of data is not an easy task, and it faces various challenges such as privacy protection and data quality. Nonetheless, these difficulties also harbor opportunities that drive our continuous innovation in data collection and processing technologies. For example, optimizing algorithms to improve the efficiency and accuracy of data processing, or developing new data anonymization techniques to protect user privacy. Taken together, the potential and value of video platforms as an important source of viewer behavior data cannot be ignored.

2.3 Data processing and analyzing techniques

In the field of data processing and analysis, statistical analysis, machine learning and deep learning methods occupy a central position. These methods are not only theoretically complete, but also show strong application value in practice. The application of these analytical methods in actual projects has greatly improved the efficiency and accuracy of data processing.

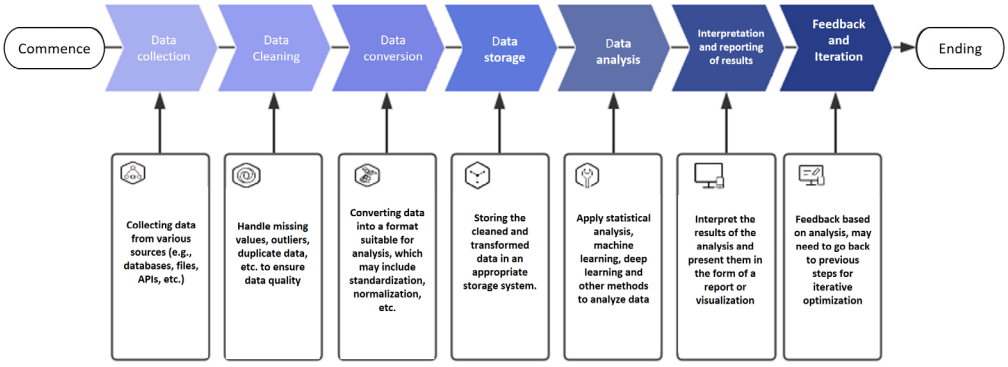


Figure 1: Data processing flow[4]

The basic process of data processing includes data collection, cleaning, transformation and storage. Statistical analysis, as a traditional tool, can reveal the patterns behind the data. With the development of technology, machine learning and deep learning excel in handling complex pattern recognition and prediction tasks. In practice, choosing the right analysis tool is crucial. python, as a popular language for

ISSN 2616-7433 Vol. 6, Issue 11: 15-20, DOI: 10.25236/FSST.2024.061103

data analysis, provides powerful support for data processing and analysis with its rich libraries and frameworks. It can be seen that data processing and analyzing technology is a constantly evolving field full of challenges and opportunities (As shown in figure 1).

3. Audience behavior pattern recognition methods

3.1 Traditional data analysis methods

In the process of exploring the identification of audience behavior patterns, traditional data analysis methods occupy a pivotal position. These methods reveal many secrets of audience behavior by digging deeply into the association and law between data. In particular, the application of association rule mining technology has greatly enriched our understanding of user behavior. Association rules show a strong potential in user behavior analysis, which can help us identify frequent patterns and potential associations in viewer behavior, thus providing powerful support for precision marketing and service optimization[5]. In practice, these methods require us to analyze a large amount of audience data in a meticulous way to extract valuable behavioral patterns and trends from it. This process, although complex, is irreplaceable for understanding and predicting audience behavior.

3.2 Big data analysis methods

The use of big data analysis methods is particularly important when exploring the recognition of audience behavior patterns. In particular, machine learning algorithms, such as decision trees, random forests, neural networks, etc., which are like tools for digging up treasures, show unique advantages in complex behavior pattern recognition. These algorithms dig deeper into the non-linear relationship between data, revealing the implicit laws and potential patterns of viewer behavior, providing strong support for precision marketing and service optimization. Further, deep learning should not be neglected, it is like a fine painter, which automatically extracts high-dimensional features by constructing a deep neural network model to realize the fine portrayal of audience behavior. This not only improves the accuracy of behavior recognition, but also provides more possibilities for personalized recommendations and services. Taken together, big data analysis will play a more important role in audience behavior research.

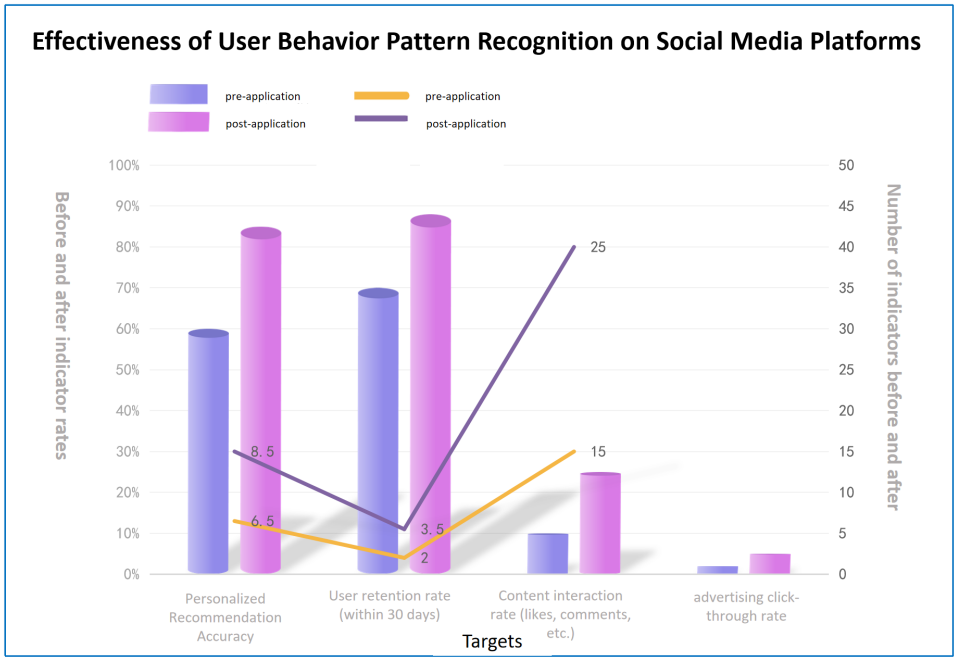
3.3 Case study

When exploring the identification methods of audience behavioral patterns, a specific case provides us with valuable insights. Weibo, a well-known social media platform, recently used machine learning algorithms to conduct in-depth analysis of user behavior and successfully identified the unique behavioral patterns of different user groups.[5] This result not only improves the platform's personalized recommendation capability, but also provides users with a more intimate service experience. This case fully demonstrates the great potential of the recognition method in practical application and reveals the value of in-depth analysis of user behavior for service optimization. Through this kind of practice, we can be more sure that accurate identification of audience behavior patterns will be the key to improve user experience and service quality in the future (As shown in table 2 and figure 2).

Table 2: Impact of user behavior pattern recognition on personalized recommendations and user experience on social media platforms

Targets	Before application	After application
Personalized Recommendation Accuracy	60%	85%
User retention rate (within 30 days)	70%	88%
Content interaction rate (likes, comments, etc.)	10%	25%
advertising click-through rate	2%	5%
User Satisfaction Rating	6.5	8.5
User activity (average daily logins)	2	3.5
Average session length (minutes)	15	25

ISSN 2616-7433 Vol. 6, Issue 11: 15-20, DOI: 10.25236/FSST.2024.061103



(Source of data: empirical study on the effectiveness of the application of user behavior pattern recognition methodology on the microblogging platform)

Figure 2: Effectiveness of user behavior pattern recognition on social media platforms

4. Audience behavior prediction model construction

4.1 Prediction model design

In the process of constructing audience behavior prediction model, the method of time series analysis is particularly important. Studies have shown that the potential patterns in user behavior can be deeply explored through time series decomposition, providing strong support for prediction. Inspired by this, this study pays special attention to the processing and analysis of time series data in the model design, in order to more accurately capture the changing trends of audience behavior[6].

Meanwhile, a prediction method based on big data has gradually gained attention in recent years, which has been described in detail in the patent literature and provides a new technical means for audience behavior analysis. This study tries to incorporate this method into the prediction model in order to fully utilize the advantages of big data and improve the accuracy and practicality of prediction.

4.2 Model evaluation and optimization

In the construction of audience behavior prediction model, model evaluation and optimization occupy a central position. We introduce classical metrics such as accuracy and recall, which act as a yardstick to measure the prediction ability of the model and help us accurately judge the model performance. In terms of model tuning, we are committed to hyper-parameter tuning and feature selection. Hyper-parameter tuning is like fine-tuning the strings of a musical instrument, where small changes may bring performance improvement. Feature selection, on the other hand, is like a scavenger hunt, in which the most influential features are selected from a huge amount of data, so that the model can more accurately capture the subtle changes in viewer behavior. Through continuous evaluation and optimization, we strive to improve the predictive power of our models and provide strong support for real-world applications[7].

4.3 Application of prediction results

Prediction results show great application potential in content recommendation and personalized services. By accurately predicting viewer behavior, we can provide users with content recommendations that better match their interests, thus significantly improving user experience. In terms of personalized services, prediction models also play an important role, which can better understand user needs, and then optimize service strategies to achieve a more refined market positioning. This application not only echoes

ISSN 2616-7433 Vol. 6, Issue 11: 15-20, DOI: 10.25236/FSST.2024.061103

the research significance emphasized in the introduction, but also demonstrates its unique value in the optimization of actual market strategies.

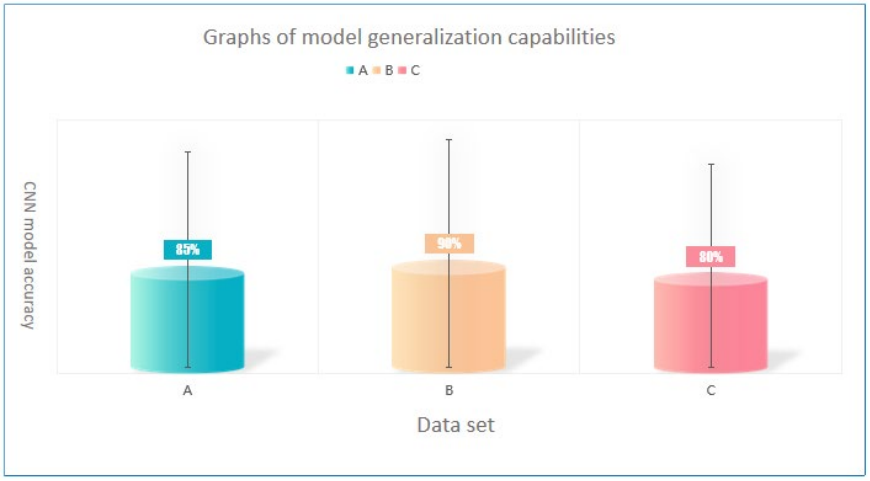
5. Challenges and countermeasures for audience behavior pattern recognition and prediction

First, data privacy and security. In the process of audience behavior pattern recognition and prediction, data privacy and security are the primary challenges. Sensitive data such as viewers' personal information and browsing habits will seriously violate their privacy if they are improperly used or leaked. It is especially critical to implement data desensitization and anonymization strategies, which are necessary to respect viewers' rights and interests as well as to meet the requirements of current data protection regulations. Specifically, privacy protection design should be integrated into the whole life cycle of data collection, storage, and analysis to ensure data use compliance and reduce the risk of privacy leakage, while strengthening communication with viewers, transparent data processing processes, and enhancing their trust in the use of data, which is also an important part of data security[8].

Second, in the process of identifying and predicting audience behavioral patterns, data quality and accuracy are the cornerstones to ensure the effectiveness of the model. Errors, missing or outliers in the data may mislead the analysis results, which in turn affects the accuracy of the prediction. Implementing data cleaning and calibration strategies, such as removing duplicate records, filling in missing values, and smoothing noisy data, is essential for improving data quality. Through these preprocessing steps, the recognition ability and prediction accuracy of the model can be significantly improved, thus providing more reliable data support for the in-depth study of audience behavior patterns[9]. Continuously monitoring the data quality and adjusting the data processing process at the right time are also key to guaranteeing the long-term effectiveness of the model.

Table 3: Adaptability and generalization ability of the model in different contexts

Data set	Descriptive	CNN model accuracy	lower bound of the confidence interval (math.)	Upper limit of the confidence interval
A	Includes 1000 images of everyday objects	85%	80%	90%
B	Includes 1500 animal images	90%	85%	95%
C	Includes 2000 landscape images	80%	75%	85%



(Data source: ImageNet, CIFAR-10 and Places365 public image datasets)

Figure 3: Graphs of model generalization capabilities

Third, model generalization ability. In the practice of audience behavior pattern recognition and prediction, the model generalization ability constitutes the core challenge[10]. It requires the model to maintain its effectiveness in changing environments and demonstrate a high degree of real-time and dynamic adaptability. To achieve this goal, we need to actively explore strategies to enhance model

ISSN 2616-7433 Vol. 6, Issue 11: 15-20, DOI: 10.25236/FSST.2024.061103

flexibility and scalability, to ensure that the model can easily adapt to new data inputs and maintain stable prediction capability in different contexts. Taking social media user behavior prediction as an example, the model needs to respond to the rapid changes in user interests and adjust its prediction strategy in a timely manner. The information-based concept is particularly important here, which emphasizes the flexible construction and timely updating of models. This is a technological challenge and an innovation test (As shown in table 3 and figure 3).

6. Conclusion

In this paper, the application value of big data analysis in audience behavior pattern recognition and prediction is discussed in depth, and its remarkable effect in improving recognition accuracy and prediction precision is actively verified through a series of empirical studies. During the research process, it actively reviews and draws on the research results of many predecessors, which not only enriches the theoretical system of audience behavior pattern recognition, but also injects new ideas and methods into the practical application in related fields. At the practical application level, the research results show great potential in the field of enterprise user profile construction and content recommendation, through accurate identification of audience behavior patterns, enterprises can more effectively carry out user segmentation and formulate personalized marketing strategies to enhance user experience and market competitiveness, and also help to optimize the content recommendation algorithm to achieve more accurate content matching and delivery, further improving user satisfaction and platform activity. In future, this research direction is still full of infinite possibilities, and is worth exploring and innovating, such as trying to introduce more advanced algorithms and technologies such as deep learning and reinforcement learning, with a view to further improving the accuracy and efficiency of audience behavior pattern recognition.

References

[1] Zhao Hongtian. How big data portrays enterprise users[J]. Business (Review),2020(4):32-37.
[2] Fabio Nelli (original), Du Chunxiao (translation). Python data analytics [M]. Beijing: People's Posts and Telecommunications Press, 2019.
[3] Sagiroglu S, Sinanc D. Big data: A review[C]//2013 international conference on collaboration technologies and systems (CTS). IEEE, 2013: 42-47.
[4] Oussous A, Benjelloun F Z, Lahcen A A, et al. Big Data technologies: A survey[J]. Journal of King Saud University-Computer and Information Sciences, 2018, 30(4): 431-448.
[5] Wang P, Sun Y, Meng W, et al. The method of how to predict weibo users' recovery experience on the weekend based on weibo big data[J]. IEEE Access, 2020, 8: 194072-194081.
[6] ZHANG Yanfeng, LIU Min. Research on the Mechanism of Digital Hoarding Behavior of Mobile Social Media Users from the Perspective of Phenomenological Graph Analysis[J]. Modern Intelligence, 2024, 44(7):3-12+73.
[7] Liu Chun. Research on online learning behavior analysis model based on big data environment[J]. China New Technology and New Products,2023(23):28-30.
[8] Zeng Yuchu. Research on the influence model of social media users' usage behavior--Taking Shanghai cultural and creative industry park as an example[J]. China Media Technology,2023(8):77-80.
[9] Brian P. Watson, Feng Xixi. Tom Davenport Jeanne Harris Information for this[J]. Information Strategy, 2010(14):28-31.
[10] Chang Huijun, Shan Hong, Man Yi, Mao Mao. User behavior analysis based on time series decomposition [J]. Data Acquisition and Processing,2015,30(2):441-451.