



## **The Dilemma and Analysis of Ideological and Political Education in Short Online Videos under Big Data Mining Algorithm**

Jian Qin, Jiajing Li  
Hebei Jiaotong Vocational and Technical College  
Shijiazhuang, 050035, China  
[lijiajing19911113@163.com](mailto:lijiajing19911113@163.com)

---

### **ABSTRACT**

*With the development of technology, short online videos have rapidly grown and become an effective means of promoting socialist core values. However, due to their unique advantages, some issues limit promoting socialist core values. Therefore, we explore this issue by using big data mining technology. Through systematic data mining and analysis, we find various problems with the current ideological and political content. Hence, we propose measures to address these issues, such as strict content review and supervision, improvement of recommendation mechanisms, and enhancing users' civility awareness. These measures will serve as important guidance for ideological and political propaganda in short online videos.*

**Received: 27 January 2024**

**Revised: 5 April 2024**

**Accepted: 14 April 2024**

**Copyright: with Author**

---

**Keywords:** Big Data Mining, Short Online Videos, Ideological and Political Work, New Media

### **1. Introduction**

In recent decades, with the rapid development of the Internet, short online videos have become a popular means of information dissemination. They are favored by a wide audience due to their concise text, rich presentation techniques, and entertaining programs [1]. Moreover, they can broaden our thinking and help the younger generation better accept and understand concepts and values of ideological and political education. Although short online videos have succeeded in recent years, they have also brought some challenges. Hence, with the help of big data mining algorithms, we can explore and effectively respond to the current social situation from multiple perspectives, thus improving the development of society[2]. With the popularity of short videos, some content creators sacrifice the depth and seriousness of ideological and political content to gain more clicks and traffic [3]. They produce low-quality entertainment videos, which pose significant challenges to disseminating ideological and political education on short online video platforms, making it challenging for severe ideological and political content to gain user acceptance. Information overload and fragmentation have become increasingly prominent issues with the development of short online video platforms. Due to the short

length and fast information dissemination of short videos, a large amount of content attracts users' attention, but it lacks systematic and orderly ideological and political education, seriously affecting social development. When users browse short videos, they often only skim through surface entertainment content without truly understanding and grasping ideological and political education's core ideas and values [4]. The algorithmic recommendation mechanisms of short online video platforms also have some issues. To improve user experience and click-through rates, these platforms often use algorithmic recommendation systems to recommend related videos to users. Although some algorithms can identify low-quality content and pursue click-through rates, this approach restricts the recommendation of ideological and political education content. As a result, even high-quality ideological and political education videos struggle to gain sufficient exposure and recommendations on short online video platforms, affecting their dissemination effectiveness. This article will use big data mining algorithms to explore the current ideological and political education situation in short online videos, thus discovering the problems and challenges therein [5]. To address this, we will take a series of measures, including strengthening content review and supervision, optimizing recommendation algorithms, and enhancing users' media literacy, to resolve these issues. After in-depth research, we identified the challenges of ideological and political education in short online videos and proposed effective solutions. These measures will promote the innovation and development of ideological and political education and enhance the ideological and political qualities of the younger generation [6].

## **2. Related Work**

In recent years, due to the various adverse effects of short online videos, they have faced numerous difficulties and received widespread attention from scholars, experts, companies, and other parties [7]. One prominent challenge is the large number of low-quality ideological and political education resources on these platforms. To effectively promote the development of ideological and moral education, we should strengthen the review of ideological and moral education content and establish a sound supervision system to ensure that all ideological and moral education content can be effectively disseminated and have a positive social impact [8]. Additionally, we should adopt a series of powerful incentive measures, such as implementing a reward mechanism, to promote the development of ideological and moral education. With the development of internet technology, there are increasing numbers of short videos on the web. However, their rapid dissemination and compact length have led to information dispersion and excessive repetition becoming commonplace. According to relevant surveys, this situation severely hinders people from better understanding and comprehending ideological and moral concepts. To address the fragmentation problem in today's society and the excessive accumulation of information [9], we should conduct in-depth research on user needs, thereby discovering more ideological and political courses and more comprehensive knowledge points to provide more accurate information to the public. At the same time, we should comprehensively summarize these fragmented video resources to construct a comprehensive ideological and political curriculum system. We have taken several measures to improve the quality of short videos to improve efficiency. Firstly, we divide all short videos according to themes and give them clear tags. Secondly, we use the latest technology to improve recommendation algorithms to meet consumers' needs. The current recommendation algorithms have certain limitations as they only focus on high-quality content and ignore user feedback, making it difficult for high-quality ideological and political education content to be widely accepted and recommended. To address this, we have developed a recommendation algorithm based on content quality and user feedback, which can provide personalized recommendation services more accurately to meet users' needs [10]. In conclusion, short online videos for ideological and political education face many challenges but also many effective solutions. For example, strengthening the collection of high-quality content, eliminating the influence of fragmented information, optimizing recommendation algorithms, and improving information overload conditions. Although short online videos for ideological and political education have achieved some results, there are still many challenges, such as how to strike a balance between ideological and political education and entertainment content and how to encourage users to participate in ideological and political education actively. Therefore, we should conduct more in-depth research in this field and seek better solutions.

## **3. Big Data Mining Algorithm Design**

With the development of social media, the ideological and political aspects of short online videos have been increasingly influenced. However, due to the advancement of big data technology, effective

classification of information in these short videos and improving quality to meet society's needs have become significant challenges today. To better identify short videos on the internet, we must first subject their information to careful preprocessing, such as eliminating noise, removing unnecessary information, and checking for abnormal data. Additionally, to evaluate the performance of algorithms effectively, extensive datasets with features should be searched and organized. We can extract essential information from videos, such as titles, descriptions, keywords, and tags using text mining techniques. This content will be used in the preprocessing of short video log data, where one of the tasks is to remove user-specific features and eliminate irrelevant information to obtain more accurate results. Furthermore, we will utilize data mining algorithms to analyze the information content in the videos and provide corresponding part-of-speech tags based on different features to achieve effective noise reduction. By removing redundant image information, filling in blank areas, ensuring image integrity, and reducing noise, we can use the following method for calculations:

$$P_{dens} = \frac{1}{P_{dist}(0, S)} \quad (1)$$

In this formula,  $P_{dens}$  represents a numerical indicator of denoising,  $P_{dist}$  is a denoising concept, and  $S$  is a measure describing the density of numerical information. Additionally, we need to eliminate information that occupies more of users' videos through digital cleaning, following the specific steps outlined below:

$$\sigma(0, S) = \sqrt{\frac{\sum_{s \in S} d(0, S)^2}{|S|}} \quad (2)$$

In the formula,  $\sigma(0, S)$  refers to the approximate range of the preprocessed data, where  $s$  represents the limit of the processed data, and  $d$  denotes the deviation of the processed data. Furthermore, this formula can help us better understand users' behaviors when using short videos and their preferences. When the noise generated by users while using short videos accounts for less than 5% of the entire dataset, the data can be considered as having no value, and appropriate measures should be taken to handle it. By utilizing machine learning algorithms, we can effectively identify the content of short videos and extract relevant information such as video duration, view count, likes, comments, etc. On the other hand, text mining techniques can be used to extract information, such as keywords, titles, descriptions, etc., from the content of the videos. To analyze the ideology of short videos on the internet, we should employ some traditional analysis methods such as decision trees, naive Bayes, support vector machines, and random forests, which can provide accurate and efficient results and can be applied to various scenarios, such as text, images, and music. To improve efficiency, we recommend using deep learning algorithms, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), which can effectively capture and analyze complex information, providing users with accurate and highly personalized short video recommendations. To ensure the reliability of this technique, we must collect a sufficient number of users to evaluate its reliability, and we can also compare it with traditional short video recommendation techniques (e.g., content and danmaku analysis) to ensure its reliability.

$$P_r = \frac{TP}{(TP + FP)} \quad (3)$$

In this formula,  $P_r$  represents accuracy. Through effective data processing techniques, the performance of the classification algorithm can be improved. Therefore, we can divide the dataset into two parts, the training set and the test set, and update the model parameters in real-time through continuous monitoring to achieve better classification results. Next, we evaluate the model using the test set and calculate metrics such as accuracy, recall, F1 score, etc., to assess the model's performance. In practical applications, we need to optimize and improve the performance of the classification algorithm continuously. This can be achieved by increasing the training data size to enhance the model's generalization ability, adjusting the weights of features and selecting appropriate feature subsets, trying different classification algorithms and model structures, and introducing domain knowledge and rules to improve

classification accuracy. Through in-depth research on big data mining algorithms, we have found that they can effectively help us accurately identify information in short videos on the internet. We have used data preprocessing, feature sampling, algorithm selection, training, evaluation, optimization, and improvement methods to develop an efficient identification system that better assists us in accurately identifying this information.

#### 4. Experimental Design and Analysis

Through big data mining algorithms, Sub-I can effectively explore and research how to address issues related to short online videos, thus achieving effective improvement. Therefore, this study aims to explore an effective experimental design and analysis method based on big data mining algorithms and conduct targeted research. Although the ideological aspect of short online videos faces numerous challenges, we should still strive to address them, such as fine content segmentation, meeting the audience's enthusiasm, and better meeting their needs. To achieve this, we should combine the current situation and develop one or more effective practical plans to achieve this goal. In this study, we will collect many online short videos. These video segments will be collected using web crawler technology from various channels. Furthermore, the study will focus on categorizing the research topics into different domains, such as ideological and political courses. In the collected raw information, meticulous data preprocessing is necessary to improve efficiency. Effective methods include noise elimination, removing unnecessary information, and detecting abnormal data. In the study of short online videos, specific selection, filtering, and optimization of information should be done to improve the quality and reliability of the information. By using deep neural network technology, we can construct a personalized new media short video recommendation system. This process involves converting video information into two independent numerical representations to better understand user needs. Then, we use these two numerical representations to build a neural network to better predict user feedback. By deeply analyzing the features of videos and users, we can convert their original numbers into two feature vectors with high semantic meaning. By testing the differences between the two features, we can infer their correlation, thereby improving the prediction of short videos on new media.

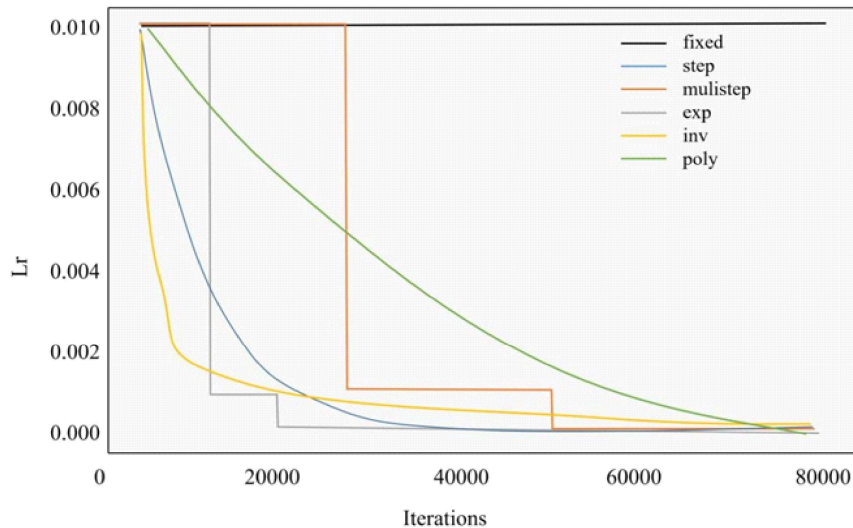
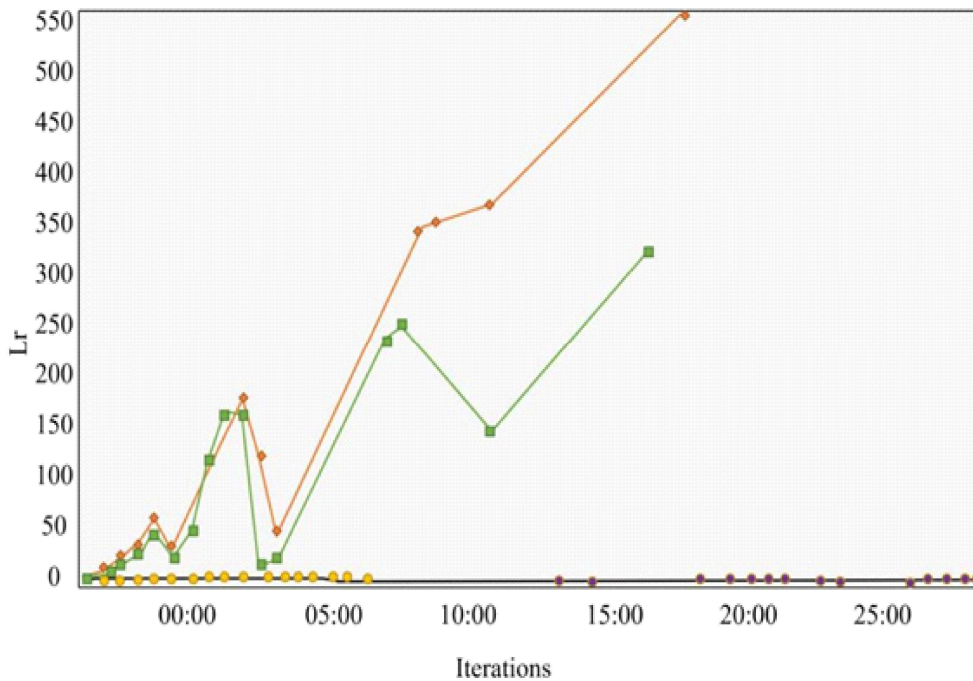


Figure 1. The Impact of Learning Rate on the Performance of the Recommendation System

Based on Figure 1, it can be observed that the learning rate can be used to evaluate the model's performance. If the learning rate is too low, the model's performance cannot reach the expected level, thus affecting the accuracy and reliability of the model. On the other hand, if the learning rate is too high, the model's performance cannot reach the expected level, thus impacting the model's performance. By utilizing the *Simuworks* platform, we can compare the synchronous performance of the traditional computer-assisted course with the performance of our proposed new device under the same experimental conditions. We can also present these differences by plotting corresponding graphs.



**Figure 2. Results of User Concurrency Testing**

According to the performance shown in Figure 2, the system provided in this paper achieves significantly higher user concurrency compared to the traditional approach, reaching 73 times. This indicates that the system is highly efficient. Therefore, when conducting experiments, planning the experimental process to achieve optimal results is essential. By using various methods, such as cross-validation, we can effectively improve the model's accuracy. These methods include multiple pieces of training to ensure that the model accurately captures information from the network and carefully selected features to reflect genuine information accurately. To better study the videos, we need to comprehensively examine their contents, including text content, view counts, number of likes, and comments. For this research purpose, we need to select features relevant to ideological and political content and then extract and process these features effectively. Finally, we need to use suitable big data mining techniques to train the model. Some traditional approaches, such as decision trees, naive Bayes, support vector machines, and random forests, are effective simulation methods. These approaches can be implemented through open platforms and simple programs. Increasing efficiency offers an effective way to understand and better predict more complex data.

Through systematic experiments, we will obtain a series of data, including but not limited to classification accuracy, recall rate, F1 score, pattern stability, prediction accuracy, and prediction error. By deeply analyzing the experimental results, we can better understand and improve the model. To meet different application scenarios, we can also improve the model's accuracy by changing its basic parameters or larger parameters. By analyzing the biases and errors in the experimental results, we can identify systematic issues and take effective measures to improve them. We adopt three different storage methods to handle uploaded data more efficiently: high-frequency, medium-frequency, and low-frequency. This allows us to retain copies of mining data, write storage codes to reduce data reading steps and establish an intelligent channel to utilize mining data in the medium-frequency storage channel effectively. According to Figure 3, the stored data exhibits noticeable fluctuation patterns.

Through systematic experimental design and analysis, we can effectively address the challenges of ideological and political content in network short videos. These steps include defining experimental objectives, data collection, data preprocessing, experimental design, feature selection, model selection and training, experimental analysis, and result interpretation and optimization. These steps help us



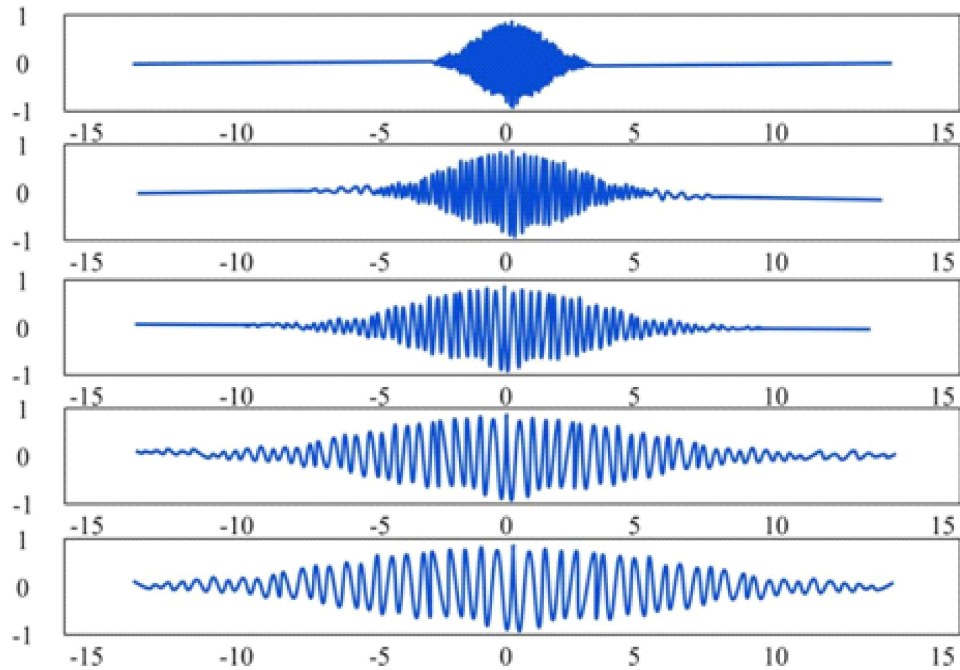


Figure 3. Waveform Chart of Stored Data

to classify more accurately and improve our work efficiency. However, in practical applications, continuous adjustments and optimizations are still necessary to achieve more outstanding experimental results.

### 5. Conclusions

Ideological and political content in network short videos is essential today. However, due to challenges such as the accuracy of ideological content classification and user interest in ideological videos, it is necessary to analyze and solve these issues through big data mining algorithms. In this paper, we propose a feasible solution to address the challenges of ideological and political content in network short videos through experimental design and analysis. In the experimental design phase, we clarify the objective of the experiment, which is to improve the accuracy of ideological content classification. Through data collection and preprocessing, we obtain a certain amount of network short video dataset and filter and clean it to ensure data quality and usability. In the feature selection stage, we select features relevant to ideological content classification, such as text information, view counts, number of likes, and comments for videos. Then, we choose appropriate big data mining algorithms for model training, such as decision trees, naive Bayes, and support vector machines. Finally, through experimental analysis, we evaluate the model's performance, calculate metrics such as classification accuracy, recall rate, F1 score, and conduct statistical analysis and result interpretation. According to the experimental results, we draw the following conclusions: Firstly, big data mining algorithms can effectively improve the accuracy of ideological and political content classification in network short videos. Our model achieves high accuracy and recall rates through training and testing the dataset, demonstrating the algorithm's effectiveness. This strongly supports accurate ideological and political content classification in network short videos. Through in-depth research, we find that carefully selected features are the key factors in determining whether network short videos can be correctly classified into the field of ideological and political content. To this end, we have selected features from multiple dimensions that can reflect this field, including but not limited to text content, user ratings, and other potential reference factors.

### Acknowledgements

The study was supported by the 2023 Hebei Province Social Science development research project results, the project name: New Era of Higher Vocational Colleges Ideological and Political Teaching Innovation.

## References

- [1] Ji, J., Zhao, N., Liu, G. (2022). Research on the application of Krathwohl classification in ideological and political education of professional courses. *Creative Education*, 13 (5), 1554-1563.
- [2] Wang, Y. (2018). Research on the application of Internet information technology in higher mathematics classroom teaching. In: Proceedings of the 2018 5th International Conference on Education, Management, Arts, Economics and Social Science (ICEMAESS 2018) (pp. 845-850). Atlantis Press.
- [3] Xia, Y. (2021). Big data based research on the management system framework of ideological and political education in colleges and universities. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*, 40 (2), 102-113.
- [4] Yan, J., Zhang, Y., Chen, L., et al. (2021). The impact of video information and publisher's characteristics in Tik Tok platform on the spreading effect of poverty alleviation by e-commerce. In: Proceedings of the Fourteenth International Conference on Management Science and Engineering Management: Volume 2 (pp. 763-777). Springer International Publishing.
- [5] Ahmad, S. Z., Ahmad, N., & Bakar, A. R. A. (2018). Reflections of entrepreneurs of small and medium-sized enterprises concerning the adoption of social media and its impact on performance outcomes: Evidence from the UAE. *Telematics and Informatics*, 35 (1), 6-17.
- [6] Shen, Z. (2022). Evaluation index system of ideological and political teaching in colleges and universities based on data mining under the background of big data. *Computational Intelligence and Neuroscience*, 2022, Article ID 2022.
- [7] Yang, P. (2022). Tik Tok and microcelebrities: An analysis of the impact of short video apps on Chinese culture and communication. *China Media Research*, 18 (1).
- [8] Khatri, P., Singh, S. R., Belani, N. K., et al. (2020). YouTube as a source of information on 2019 novel coronavirus outbreak: A cross-sectional study of English and Mandarin content. *Travel Medicine and Infectious Disease*, 35, 101636.
- [9] Lu, J. (2021). Research on the ideological and political teaching mode of dual system curriculum in colleges and universities based on MOOC. In: Proceedings of the 7th EAI International Conference, eLEOT 2021, Xinxiang, China, June 20-21, 2021, Part I (pp. 231-242). Springer International Publishing.
- [10] Chen, T. (2022). Research on the dilemma and breakthrough path of ideological and political education in colleges and universities in the era of big data. *Journal of Higher Education Research*, 3 (2), 203-206.

## **Book Review**

### **Change and Stability in Thesis and Dissertation Writing**

#### ***The Evolution of an Academic Genre***

**Brian Paltridge and Sue Starfield**

BLOOMSBURY ACADEMIC  
Bloomsbury Publishing Plc  
London, New York, NY, Dublin  
First published in Great Britain 2024  
Copyright © Brian Paltridge and Sue Starfield, 2024

**ISBN: HB: 978-1-3501-4657-0**

**ePDF: 978-1-3501-4658-7**

**eBook: 978-1-3501-4659-4**

It has ten chapters, an appendix, references and an index.

The authors of this book have read many PhD theses and dissertations.

This book has an excellent introduction.

“Doctoral theses and dissertations, have been a relatively stable genre for many years. This has begun to change, however, especially as doctoral degrees are offered in an increasing range of academic disciplines where alternate scholarly forms are now being accepted for the award of the degree .....”

In **Chapter 2**, the authors traced the evolution and development of these over time. In this chapter, they conclude that the technologies impact the changes in the structure of theses.

The **3 chapter** extensively reviews the research methods used in past theses and dissertations. It discusses textography, which includes observations, document analysis, and other data sources. “A textography, thus, aims to provide an understanding of the context in which a text is produced to gain an understanding of why texts are written as they are as well as extend students’ understandings of the purposes for which they are writing” (Evans 2022).

**Chapter 4** reported on research carried out using a larger and more international data set to explore changes in thesis and dissertation types over time. Chapter 5 reviewed and extended it with a more recent set of doctoral theses to chart more recent developments.

**Chapter 6** deals with the Doctor of Education degree, while **chapters 7 and 8** discuss doctoral degrees in the visual arts and music theses. Using ethnographic interviews helps students and supervisors understand the challenges posed by these theses.

**Chapter 9** explores the ‘Doctorate by publication’, a collection of journal articles that have already been published or are about to be published. This is a relatively new form of doctorate in the humanities & social sciences. The genre shares similarities with the ‘prudential research article’ and the ‘more conventional’ doctoral thesis. Drawing on textual analysis of recently submitted doctoral theses by publications and interviewing students and supervisors, this chapter examines the challenges students and supervisors face and their decisions as authors and supervisors working in a relatively unknown territory. Chapter 10 explores change and stability in the field of doctorates and what the future holds