applied sciences

MDPI

Article

Value Assessment of UGC Short Videos through Element Mining and Data Analysis

Jinyu Fang, Yuan Ni * and Jian Zhang

School of Economics and Management, Beijing Information Science and Technology University, Beijing 100192, China; 2021020711@bistu.edu.cn (J.F.); zhangjian@bistu.edu.cn (J.Z.)
* Correspondence: niyuan@bistu.edu.cn

Abstract: UGC short videos play a crucial role in sharing information and disseminating content in the era of new information technology. Accurately assessing the value of UGC short videos is highly significant for the sustainable development of self-media platforms and the secure governance of cyberspace. This study proposes a method for assessing the value of UGC short videos from the perspective of element mining and data analysis. The method involves three steps. Firstly, the text clustering algorithm and topic mapping visualization technology are utilized to identify elements for assessing the value of UGC short videos and construct an assessment index system. Secondly, structured data indexes are quantified using platform data statistics, while unstructured data indexes are quantified using the LSTM fine-grained sentiment analysis model. Lastly, the VIKOR model, incorporating an improved gray correlation coefficient, is employed to effectively evaluate the value of UGC short videos. The empirical results indicate that the value of current domestic UGC short videos is primarily associated with three dimensions: the creators, the platforms, and the users. It encompasses 11 value elements, including fan popularity, economic returns of creation, and frequency of interaction. Additionally, we assess the value of short videos within the mainstream partitions of the Bilibili platform and generate a value radar chart. Our findings reveal that short videos in game partitions generate higher revenue for creators and platforms but may neglect users' needs for knowledge, culture, and other content. Conversely, short videos in the knowledge, food, and music partitions demonstrate specific distinctions in fulfilling users' requirements. Ultimately, we offer personalized recommendations for the future development of high-value UGC short videos within the mainstream partitions.

Keywords: UGC short video; LSTM; value assessment; social media; data analysis

Citation: Fang, J.; Ni, Y.; Zhang, J. Value Assessment of UGC Short Videos through Element Mining and Data Analysis. *Appl. Sci.* **2023**, *13*, 9418. https://doi.org/10.3390/app13169418

Academic Editors: Krzysztof Koszela, Dickson K.W. Chiu and Kevin K.W. Ho

Received: 26 June 2023
Revised: 31 July 2023
Accepted: 16 August 2023
Published: 19 August 2023

1. Introduction

During the era of Web 2.0, the field of digital media has experienced rapid growth, with various forms of digital media constantly evolving. One such form is UGC short videos, which serve as a prominent means for creating digital resources and engaging in information interaction [1]. UGC short videos refer to short video content created and shared by users on digital platforms. Users typically create these videos to capture life moments, express creativity, share opinions, or showcase their talents [2]. In recent years, UGC short videos have gained significant worldwide success and influence due to their ease of distribution and diverse creativity. This has resulted in the emergence of UGC short video sharing platforms such as YouTube, TikTok, Bilibili, etc. The 2023 China Network Audiovisual Development Research Report, released on 29 March 2023, reveals that as of December 2022, the number of Chinese short video users reached 1.012 billion. This reflects a year-on-year increase of 77 million users, corresponding to a growth rate of 8.3% [3].

However, the increasing number of UGC short videos poses significant challenges for the content management of video sharing platforms. On one hand, the platform witnesses the widespread dissemination of UGC short videos with negative effects, including defaming others' images, spreading false information, and violating privacy. Users may doubt the platform's content filtering and auditing mechanisms and question the authenticity of the video content, thus harming the platform's reputation and credibility [4]. On the other hand, the platform has overly focused on the traffic generated by UGC short videos, disregarding their inherent value and hindering the development of other videos that offer more creative and valuable content, thereby restricting the diversity of UGC short videos on the platform. Consequently, this limitation hampers the diversity of UGC short videos on the platform, which undermines effective content strategy optimization and hinders improvements in the user experience [5]. Therefore, it is crucial to assess the value of UGC short videos to enhance the operational capacity of the video sharing platform and foster its sustainable development in the future. For example, a user on the TikTok platform uploaded a short video demonstrating how to prepare a specialty dish at home in a vivid and engaging manner. The platform can accurately understand which types of content users are more interested in by evaluating various dimensions. For instance, they might be more interested in cooking skills conveyed through videos or in local Chinese cultural specialties that are popularized by science. This is crucial for platforms to optimize their content recommendation strategies. By adjusting their recommendation algorithms according to users' feedback on different types of food videos, platforms can provide content that better matches users' preferences, ultimately leading to increased user activity and retention.

This paper aims to address two questions and provide theories and methods for evaluating the value of user-generated content (UGC) short videos on video sharing platforms in mainland China. (1) What elements contribute to the value of UGC short videos in mainland China? (2) How can we scientifically establish a value assessment model for UGC short videos?

This paper proposes a method for assessing the value of UGC short videos. First, the value elements for assessing the value of UGC short videos in the professional field are identified, and an index system for UGC short video value assessment is constructed. Second, the LSTM fine-grained sentiment analysis model is employed to calculate the sentiment value for the unstructured data indexes in the assessment system, achieving index quantification. The quantification of structured data indicators is based on statistics from multi-source platforms, and all indicators are assigned values using a combination of methods. Finally, the value of UGC short videos is evaluated using an improved VIKOR method with gray correlation coefficients. A radar chart of value dimensions is then created to provide decision support for the platform's evaluation of UGC short videos.

2. Literature Review

2.1. Study on the Value Assessment of UGC Short Videos Using Technical Methods

Two main methods are commonly used to evaluate the value of UGC short videos: traditional technical methods and deep learning [6]. The traditional technical approach involves utilizing computer vision and signal processing techniques to process and analyze UGC short videos, aiming to determine their value level. For instance, Sun et al. [7] developed a VQA model that effectively assesses the value of UGC videos. They achieved this by training an end-to-end spatial feature extraction network, enabling the model to directly learn value-aware spatial features from the original pixels of UGC videos. Tu et al. [8] enhanced the performance of the reference-free VQA video value assessment model (BVQA) by incorporating a feature selection strategy, which demonstrated favorable results. Nanne et al. [9] compared the performance of three mainstream computer vision technology models (YOLOV2, Google Cloud Vision, and Clarifai) for evaluating UGC value. Their findings revealed that the Google Cloud Vision model outperformed the others, especially in assessing brand promotional UGCs. Deep learning methods, being the prevailing AI technology, enable automatic learning and extraction of high-dimensional and nonlinear data features through the construction of deep neural network models. This facilitates the efficient evaluation of UGC's value. For instance, Wang et al. [10] introduced a framework based on deep neural networks (DNNs) that considers features like technical value and video compression level. This framework aims to perceive the significance of UGC video value and ultimately assign a value score for assessment. Su et al. [11] developed a hyper-network-based NR-IQA model that can dynamically adjust video value prediction parameters. Their model exhibited superior performance in differentiating UGC values.

2.2. Study on Value Assessment of UGC Short Videos through the Construction of an Index System

The value assessment method based on constructing an index system refers to the process of developing a UGC value assessment index system. This method involves enhancing the existing index system through qualitative or quantitative methods and subsequently utilizing the index system for assessing the value of UGC short videos. It serves as a crucial foundation for evaluating the value of UGC short videos. For instance, Meng et al. [12] developed a comprehensive six-dimensional UGC short video value assessment index system that encompasses attributes of publishers, viewer behavior, video content, and more. They conducted a questionnaire survey and identified video quality, title, and viewer interest as the primary factors influencing the value of UGC short videos. Chen et al. [13] applied rooting theory to identify the fundamental value elements of rural landscape short videos. Additionally, Yu et al. [14] employed data analysis and gathered short video value factors using a randomly administered online questionnaire. Their findings revealed that enhancing user stickiness and increasing user scale can substantially boost the value of short UGC videos. Mohammad et al. [15] developed an online brand-based UGC value assessment system and employed partial least squares-structural equation modeling (PLS-SEM) to analyze empirical data. They revealed that content and technology are the primary elements that influence the value of UGC short videos. Zhang et al. [16] analyzed the value elements of short video communication using the all-information emotion theory and the super IP theory. They employed the fuzzy set theory-DEMATEL model to identify the key influencing factors on the secondary indicators and identified 11 value elements, which include public expression and popularity demand, among others.

2.3. Study on the Value Assessment of UGC Short Videos Using Data Modeling

The value assessment method based on data modeling involves automating the assessment of UGC short videos by constructing data models and algorithms using user-generated content data. For example, Cai [17] proposed a value assessment model for UGC short videos based on random forest and BP neural networks. Through experiments, it was found that the constructed model significantly improved the accuracy of UGC short video value assessment compared to a single model. Manikandan et al. [18] compared multiple models, including linear regression, random forest, and XGBoost, for assessing the value and performance of UGC short videos. They discovered that the random forest model outperformed other machine learning models. Leszczuk et al. [19] used the random forest model to train and test using professional content data from UGC short videos, along with video technology data. They achieved better performance in evaluating the value of IP-based UGC videos. Rui et al. [20] evaluated the value of UGC short videos using online gradient descent and least squares methods, respectively. They compared the methods and found that the least squares method yielded more accurate value outputs. Gupta et al. [21] compiled a comprehensive list of machine learning and artificial intelligence techniques used in assessing the value of UGC short videos. They discovered that these techniques aid creators in assessing the authenticity of their content and enhance the production of high-quality short video resources. Mekouar et al. [22] employed a simple machine learning model to assess the distribution value of UGC short videos based on short video parameters and a proposed popularity function. They conducted an empirical study using YouTube short videos as an example and found that their proposed model exhibited good assessment performance.

2.4. Research Gap

In summary, recent years have seen some progress in research assessing the value of UGC short videos, but further improvements can still be made.

The research methods primarily focus on evaluating the value of UGC short videos in the technical fields of computer vision and signal processing. These methods excessively emphasize the technical elements of UGC short videos while neglecting the content attributes and value attributes of UGC short videos as a form of user-generated content information product. Additionally, value assessment methods based on constructing index systems have started to investigate the content and emotional characteristics of UGC short videos. However, the data modeling-based approach allows for the quantification of various attributes of UGC short videos. Moreover, the machine learning model demonstrates excellent performance in evaluation. This aspect has practical application. Nevertheless, the following problems persist:

First, the evaluation index system is typically constructed using questionnaires and a literature review. However, it lacks multidimensional value element exploration, correlation analysis, and model representation of UGC short videos through the integration of multiple authoritative data sources. Second, the existing indexes quantify the factors in a singular manner and fail to effectively capture the impact of unstructured data indicators on the value of UGC short videos. This limitation makes it challenging to fully capture the emotional characteristics of users within the indexes. Third, the current evaluation models mostly rely on machine learning techniques, which do not adequately consider the correlation and weightage between different indicators. Consequently, it becomes difficult to comprehensively assess the variations in value among the research subjects. Finally, most of these models are black-box models, and interpreting the evaluation results is often challenging. Thus, additional statistical analysis and visualization methods are required to understand the underlying patterns.

Considering these issues, this paper aims to address the strong subjectivity of the index system, the insufficient quantification of key features, and the limited interpretation of evaluation results in the evaluation process by mining the value elements of UGC short videos. The research focuses on the following aspects: (1) Mining the value elements for assessing UGC short videos by utilizing text clustering algorithms, visualization technology for theme mapping, and incorporating professional literature and expert commentary texts. This will facilitate the construction of a comprehensive UGC short video value assessment index system; (2) employing a diverse processing approach for multi-dimensional index data: structured data indicators are statistically quantified from UGC short video operation platforms and third-party platforms, while unstructured data indicators, such as user comments, are quantified through LSTM fine-grained sentiment analysis; and (3) utilizing the AHP method and CRITIC method to combine normalized index data for weighting, which is then input into an improved VIKOR model with gray correlation coefficients. This enables the realization of value assessments for UGC short videos.

3. Methodology

This paper aims to assess the value of UGC short videos. It utilizes authoritative texts, including literature collections in the field and expert commentary articles, as the basis. Text clustering algorithms and visualization techniques are employed for topic mapping. Through this process, the UGC short video value elements are identified, and a comprehensive UGC short video value assessment index system is constructed. Subsequently, a UGC value assessment model is established to achieve effective value assessment of UGC short videos, as illustrated in Figure 1.

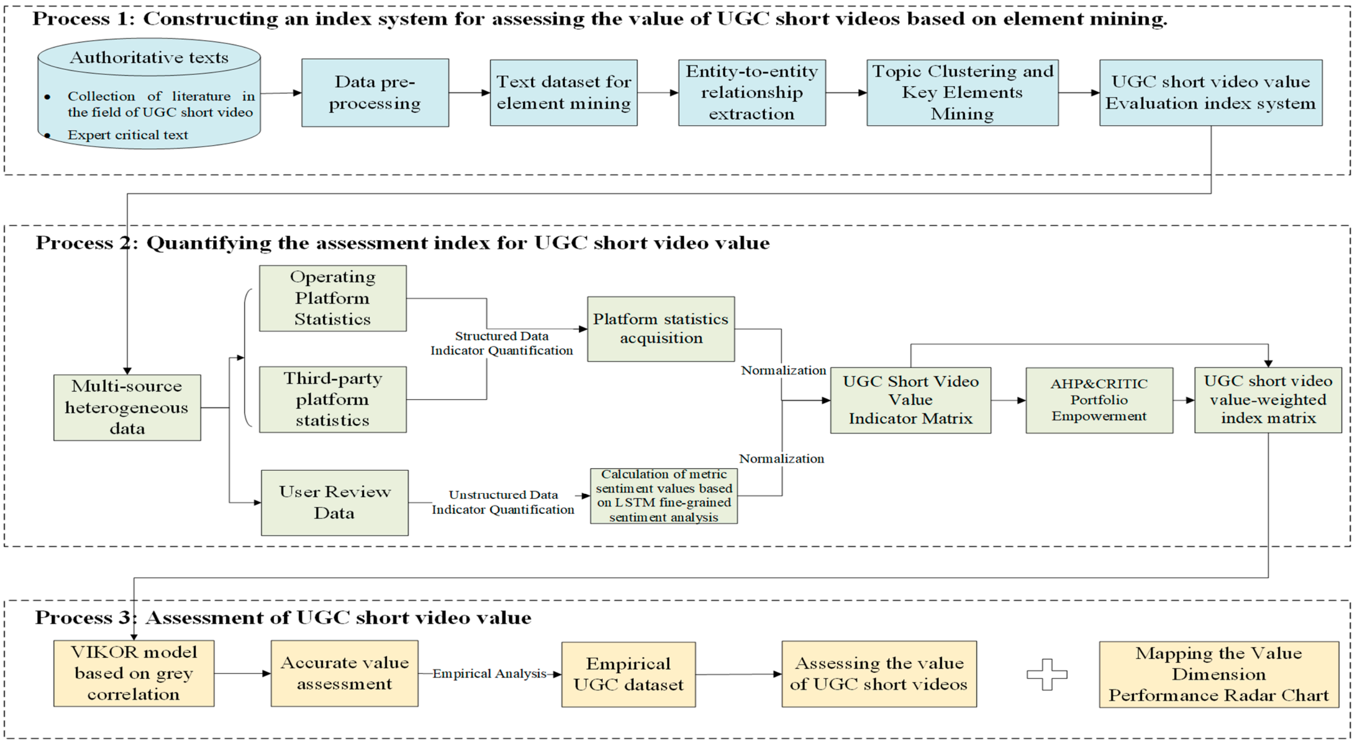

Figure 1. UGC short video value assessment model.

3.1. *Process 1: Constructing an Index System for Assessing the Value of UGC Short Videos Based on Element Mining*

This study initially retrieves relevant literature on UGC short video value assessment from databases such as CNKI, Wanfang, and WOS. The collected literature includes fields such as title, abstract, keywords, and publication time. Additionally, expert review texts related to UGC short videos are collected as Supplementary Data. The collected data undergoes pre-processing, which involves text data normalization, merging, and the removal of duplicate and invalid samples. As a result of this process, a dataset for UGC short video value mining is obtained.

Text clustering algorithms and visualization techniques, such as topic mapping, are employed to mine UGC short video value elements and construct the UGC short video value evaluation index system. The specific process is illustrated in Figure 2.

After pre-processing the data, the text dataset undergoes entity and inter-entity relationship extraction. Specifically, this process includes: (a) The extraction of entities. This paper performs data cleaning on the text and employs the jieba word splitting tool to accurately segment the text data based on specific word splitting patterns, thereby eliminating redundant data. Additionally, the TF-IDF strategy in natural language processing is utilized to remove punctuation and stop words, minimizing interference with word segmentation and obtaining the subject words for each analyzed content; (b) The calculation of inter-entity relationships. The K-means method is utilized to cluster the subject words, and then the strength of entity association, inter-entity spacing, and inter-entity position are calculated. The inter-entity association relationship is visually presented by controlling the entity color; and (c) The construction of topic mapping. The topic mapping is constructed using VOSviewer to visualize the association relationship between elements.

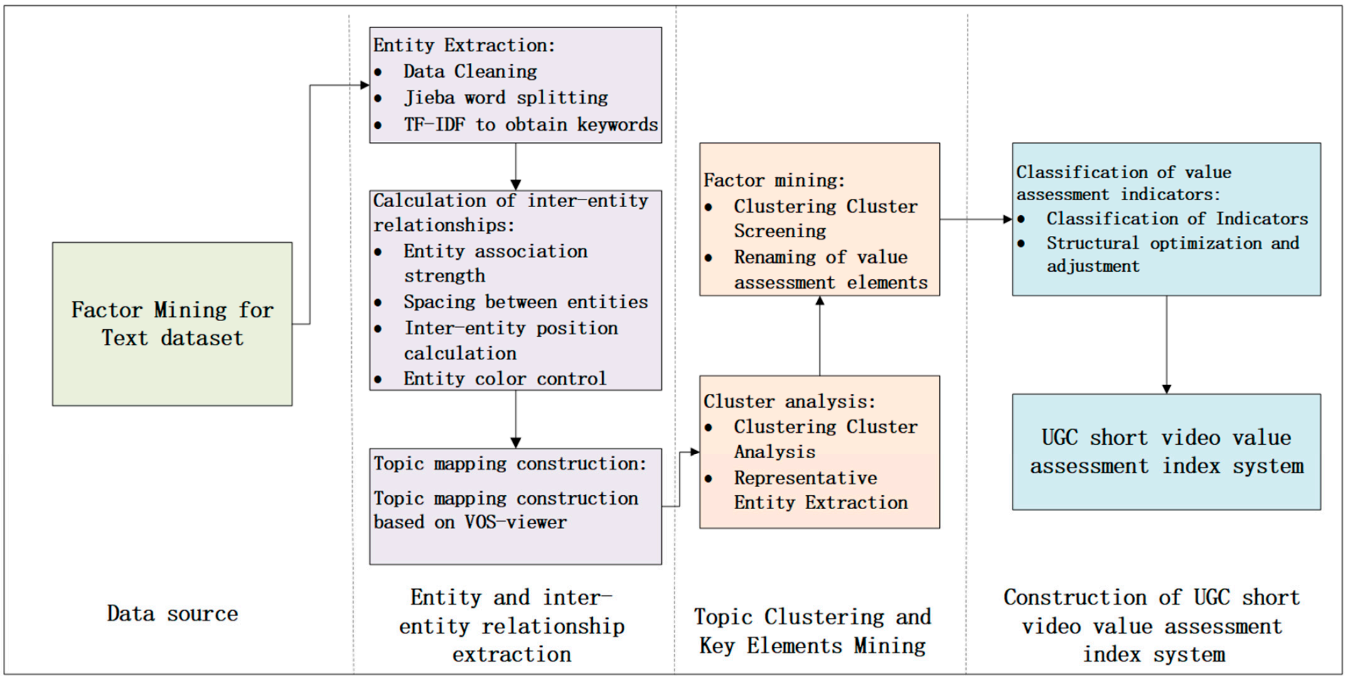

Figure 2. The process of constructing the UGC short video value assessment index system.

Subsequently, theme clustering and key element mining are conducted. This process includes: (a) The analysis of clusters. Utilizing the constructed topic mapping, cluster analysis and representative entity word extraction are performed to analyze the characteristics of entity words within each cluster category; and (b) The mining of key elements. The clusters that fulfill the value assessment criteria for UGC short videos are selected, and these selected clusters are renamed by incorporating the value characteristics of UGC short videos.

Lastly, utilizing classical theory or a priori knowledge, the mined value assessment elements are classified and organized hierarchically. The structure is then optimized and adjusted based on the similarity and difference of indicators, resulting in the construction of a UGC short video value assessment index system.

3.2. Process 2: Quantifying the Assessment Index for UGC Short Video Value

This paper focuses on quantifying indicators using heterogeneous data from multiple sources. The main sources include structured data from UGC short video operation platforms and third-party platform statistics, as well as unstructured data from user comments. The data processing steps are outlined as follows:

Step 1: Processing of structured data indicators. To mitigate the impact of indicator data dimensions and enable comparability among different indicators, the indicator data obtained from various platforms should undergo individual maximum–minimum normalization;

Step 2: Processing of unstructured data indexes. In the social media environment, users express their emotions, opinions, and attitudes through the creation and publication of UGC short videos. These videos encapsulate the intense emotional catharsis and clash of ideas experienced by users [23]. Thus, user attributes related to emotional tendencies become fundamental attributes of UGC short video value [24]. The index data for these user attributes cannot be directly quantified. Typically, quantification is accomplished through text data analysis techniques such as word frequency statistics [25] and hierarchical analysis [26] based on feature words. However, these methods do not originate from the user's inherent attributes, thus making it less likely to explore the user's genuine emotional expressions and emotional tendencies. Consequently, accurate quantification of UGC short video value indexes that encompass user attributes cannot be achieved through these methods.

The investigation of user review text sentiment intensity has gained significant attention in the academic, business, and political spheres in recent years. Numerous studies have indicated the crucial role of calculating user review text sentiment intensity in optimizing domain-specific sentiment lexicons [27] and devising brand development strategies. Hence, this paper begins with user review texts and achieves precise quantification of metrics related to user attributes by computing sentiment values extracted from user reviews.

It has been observed that computer information technology performs effectively in processing textual information. However, traditional neural network methods can suffer from issues like gradient explosion and gradient disappearance with an increasing number of learning layers. On the other hand, LSTM models exhibit robustness and generalization capabilities [28,29], which can effectively address these challenges in textual information processing [30–32]. Therefore, this paper integrates the domain sentiment lexicon with the LSTM deep learning model to calculate the sentiment value of indicators and achieve the quantification of user comment data. The specific steps consist of four parts: data processing, feature word extraction, feature sentence extraction, and training and testing based on the LSTM model. Refer to Figure 3 for detailed information.

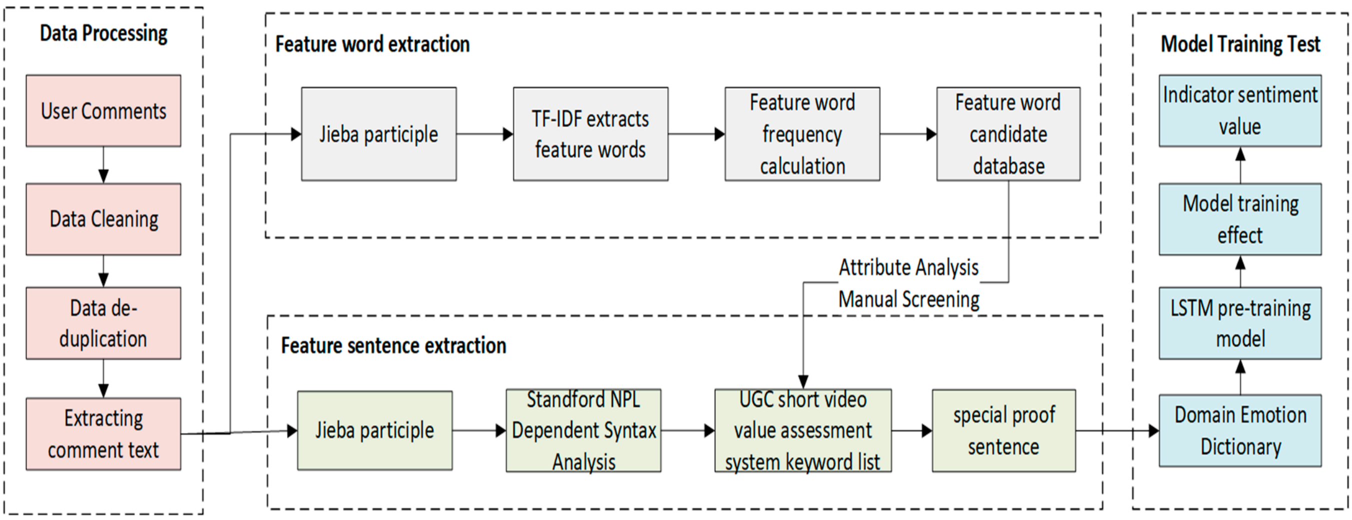

Figure 3. Emotional intensity calculation model based on the LSTM model.

(1) Data processing. Prior to extracting feature words and special evidence sentences from the comment text, it is necessary to perform data cleaning and de-weighting on the comment data. This step significantly improves the quality and quantity of the comment text while preventing interference or reduction of accuracy in calculating the sentiment value of indicators due to irrelevant or duplicate comments;

(2) Feature word extraction. Term Frequency (TF) represents the frequency of words, which is the number of times a word appears in the text. The processed result is the word frequency, TF. The importance of a word in the text increases with its frequency. The feature word frequency is obtained by utilizing the counter library to create a candidate library of feature words. The feature words are classified into the UGC short video value assessment index system keyword table through manual screening, with reference to the constructed UGC short video value assessment index system;

(3) Feature sentence extraction. After performing word-by-word traversal using jieba splitting, the Stanford NPL is utilized for dependency sentence analysis. The result is then compared with the keyword table of the UGC short video value assessment index system. The matching feature words are considered the feature attributes of the comment text, thus extracting the comment as a special evidence sentence;

(4) LSTM-based model training and testing. To reduce the manual sentiment labeling time, this study labels the extracted idiosyncratic sentences with sentiments (-1 for negative, 0 for neutral, and 1 for positive) using a sentiment dictionary. The Boson NLP sentiment lexicon, constructed based on microblogging, news, forums, and other data sources, serves as the base lexicon, supplemented by the KnowNet HowNet negative word and degree adverb lexicon, ultimately forming the domain sentiment lexicon.

For model training and testing, the Long Short-Term Memory (LSTM) neural network is used. First, the data is divided into a training set and a test set with a ratio of 4:1. Second, the tanh function is chosen as the activation function for the LSTM model, commonly used for transforming input continuous real values into output values between -1 and 1. The word vector dimension is set to 100, representing the semantic information of each word as a 100-dimensional vector. The data batch size is 32, ensuring computational efficiency and providing sufficient data for the model to learn. To prevent overfitting, the dropout technique is introduced, randomly disabling a portion of neurons during training. After cross-validation, a dropout value of 0.5 is selected as it generates the most diverse network structures;

Step 3, the index data obtained from the normalization process in Steps 1 and 2 are combined together to form the UGC short video value index matrix;

Step 4, the combination of indicators is determined using the AHP method and the CRITIC method. This approach integrates the subjective experience of decision-makers with the objective attributes of data information.

Hierarchical analysis (AHP) is a multi-criteria decision-making method introduced by Saaty, an American operations researcher [33]. This method simplifies multi-objective decision-making problems and enables fuzzy quantification of qualitative indicators, making it simple and practical. However, it heavily relies on subjective judgment and does not offer novel solutions for decision-makers, thus being categorized as a subjective empowerment method [34]. The CRITIC method, proposed by Diakoulaki et al. in 1995 [35], is an objective empowerment method. It provides a comprehensive assessment of the objective weights of indicators by considering the comparative strengths of evaluation indicators and the conflicts between them. However, the objective weighting method demands a significant amount of data, involves complex calculations, and does not consider the rater's importance, leading to a notable discrepancy between the assessed weights and the actual importance [36].

In this study, the AHP method and the CRITIC method are integrated with subjective and objective assignments to conduct a comprehensive assessment of the weights of evaluation indices. This approach addresses the limitations of single assignment methods [37,38] and provides comprehensive weights for each index. It effectively incorporates both the subjective intentions of decision-makers and the objective attributes of data information.

The subjective weight of the jth indicator calculated using the AHP method is denoted as W_{Pj}, and the objective weight of the jth indicator calculated using the CRITIC method is denoted as W_{Vj}, assuming a total of i indicators. Since there exists a certain interdependence among UGC short video value indicators, this paper utilizes the multiplicative synthesis method to calculate the combined weights of the indicators [39].

$$W_j = \frac{W_{Pj} \cdot W_{Vj}}{\sum_{j=1}^{j=i} W_{Pj} \cdot W_{Vj}}, j = 1, 2, 3, \cdots i; \tag{1}$$

Step 5, the UGC short video value weighting index matrix is determined based on the UGC short video value index matrix and the combination weights of the UGC short video value assessment indices. This matrix serves as the input data for the subsequent value assessment model.

3.3. Process 3: Assessment of UGC Short Video Value

This study effectively evaluates the value of UGC short videos using the improved VIKOR model incorporating gray correlation coefficients.

The VIKOR (multicriteria compromise optimal solution) method, proposed by Opricovic and Tzeng [40], is a compromise-based multicriteria decision-making method. It obtains the feasible compromise solution that is closest to the multi-attribute positive ideal solution by making concessions between attributes, thereby determining the relative superiority or inferiority of the evaluated object based on the evaluation index [41,42]. This approach enables the value assessment of UGC short videos.

Recognizing that utilizing a standardized method for calculating the gray correlation coefficients of each assessment index can significantly enhance the model's accuracy and convergence speed [43,44], this study integrates the gray correlation coefficients into the VIKOR model to enhance the validity of decision results. The calculation steps are outlined below:

(1) Unquantified processing of indicators. The unquantified processing of gray correlation involves determining the reference sequence $X_0(X_0 = x_{01}, x_{02}, \cdots, x_{0n})$, the comparison sequence $X_i(X_i = x_{i1}, x_{i2}, \cdots, x_{in})$, and the normalized index f_{ij} (correlation coefficients). The unquantified formula is as follows:

$$f_{ij} = \frac{\underset{m}{min}\underset{n}{min}|x_{0j} - x_{ij}| + \rho \underset{m}{max}\underset{n}{max}|x_{0j} - x_{ij}|}{|x_{0j} - x_{ij}| + \rho \underset{m}{max}\underset{n}{max}|x_{0j} - x_{ij}|}, \tag{2}$$

the quantization adopts a resolution coefficient $\rho \in (0, 1]$, set to 0.5. The variables m and n indicate the starting and ending positions of the subsequence, respectively, where both m and n are integers satisfying $1 \leq m \leq n$. The resolution coefficient ρ is employed in gray correlation analysis to adjust the similarity between sequences. The value range of this parameter is (0, 1]. When ρ approaches 1, the correlation coefficient of all sequences becomes close to 1, indicating reduced dissimilarity between the sequences. Conversely, as ρ approaches 0, the dissimilarity between sequences increases. In this paper, ρ is set to 0.5. This choice ensures a suitable balance between difference and similarity, allowing the correlation coefficients between sequences to reflect their similarity while also showing their difference to a certain extent. Additionally, setting ρ to 0.5 helps avoid potential computational fluctuations;

(2) Calculate the positive ideal solution f_j^+ and the negative ideal solution f_j^- for each indicator as follows:

$$f_j^+ = [(\underset{i}{max} f_{ij} \mid j \in I_1), (\underset{i}{max} f_{ij} \mid j \in I_2)], \tag{3}$$

$$f_j^- = [(\underset{i}{min} f_{ij} \mid j \in I_1), (\underset{i}{min} f_{ij} \mid j \in I_2)], \tag{4}$$

where I_1 represents a benefit-based indicator and I_2 represents a cost-based indicator;

(3) Calculate the group benefit value S_i and the individual regret value R_i for each assessment object as follows:

$$S_i = \sum_{j=1}^{n} \omega_j \frac{\left(f_j^+ - f_{ij}\right)}{\left(f_j^+ - f_j^-\right)}, \tag{5}$$

$$R_i = max \omega_j \frac{(f_j^+ - f_{ij})}{(f_j^+ - f_j^-)}, \tag{6}$$

where S_i represents the group effect of the assessed object and a smaller S_i value indicates a lower group effect value. R_i represents the individual regret value and a smaller R_i value indicates a lower individual regret. ω_j is used to indicate the weight of the jth evaluation indicator;

(4) Calculate the interest ratio of the appraisal object, which serves as a crucial basis for assessing the value of UGC short videos:

$$Q_i = v\frac{(S_i - S^+)}{(S^- - S^+)} + (1-v)\frac{(R_i - R^+)}{(R^- - R^+)}, \tag{7}$$

where S^+= minS_i, S^-= maxS_i, R^+= minR_i, and R^-= maxR_i. v is the decision mechanism coefficient, which is taken as 0.5. The rationale behind this choice is that v aims to strike a reasonable balance between S_i and R_i, determining whether the evaluation strategy favors the majority-criterion optimal solution (minimum S_i) or the maximum group consensus solution (minimum R_i). Typically, v has values between 0 and 1. A value of v = 0 considers only the group benefit value, while v = 1 considers only the individual regret value. When v is set to 0.5, both the group benefit and the individual regret carry equal weight, representing a typical equal decision-making mechanism. This approach ensures that the evaluation and comparison of the short video's value account for both the overall benefit and the impact of individual elements. The choice of 0.5 is common because it achieves a well-balanced consideration of group and individual impacts.

Following the aforementioned process, the value of UGC short videos as empirical objects is assessed. Additionally, a radar chart depicting the UGC short video value dimensions is generated, aiding the platform in discerning the value characteristics of various UGC short video categories.

4. Results

4.1. Construction of the UGC Short Video Value Assessment Index System

4.1.1. Data Acquisition and Processing

In mainland China, the main databases selected for literature searches were CNKI and Wanfang, supplemented by expert review texts. The keywords were chosen through literature research, resulting in the following search formula: Topic/Keyword/Abstract = (short video and user-generated content) AND (short video and UGC) AND (short video and content generation) AND (TikTok and content generation) AND (Bilibili and content generation). For foreign databases, Web of Science was selected as the primary data source, and the search formula used was: TS = ("short video" AND "UGC") OR TS = ("short video" AND "user generated content") OR TS = ("short video" AND "content production") OR TS = ("Tiktok" AND "user generated content") OR TS = ("short video" AND "consumer generated media"). The search period spanned from 1 January 2010 to 1 July 2022, focusing on papers and journals related to the research topic of "UGC short video value". Irrelevant literature was excluded, resulting in a total of 212 valid related articles. Additionally, 41 articles were collected from reputable platforms such as People's Daily Online and Guangming.com that were relevant to the research topic. After the data pre-processing step referred to as "Process 1", a total of 253 text data pieces were obtained.

4.1.2. Element Mining Based on Text Clustering Algorithm and Topic Mapping

The commentary text data is processed using the relationship extraction technique between entities in Process 1. The standardized association strength selection method is employed with default values. The random parameter starts with a value of 1, and the maximum number of iterations is set to 1000. The initial step size is 1, and the step size convergence coefficient is 0.001. The clustering resolution and minimum number of clustering items are both set to 1. The process is repeated 10 times to obtain the topic mapping of feature words shown in Figure 4. The mapping represents the location and associations of feature words using various visual cues such as color, size, ring thickness, and perspective.

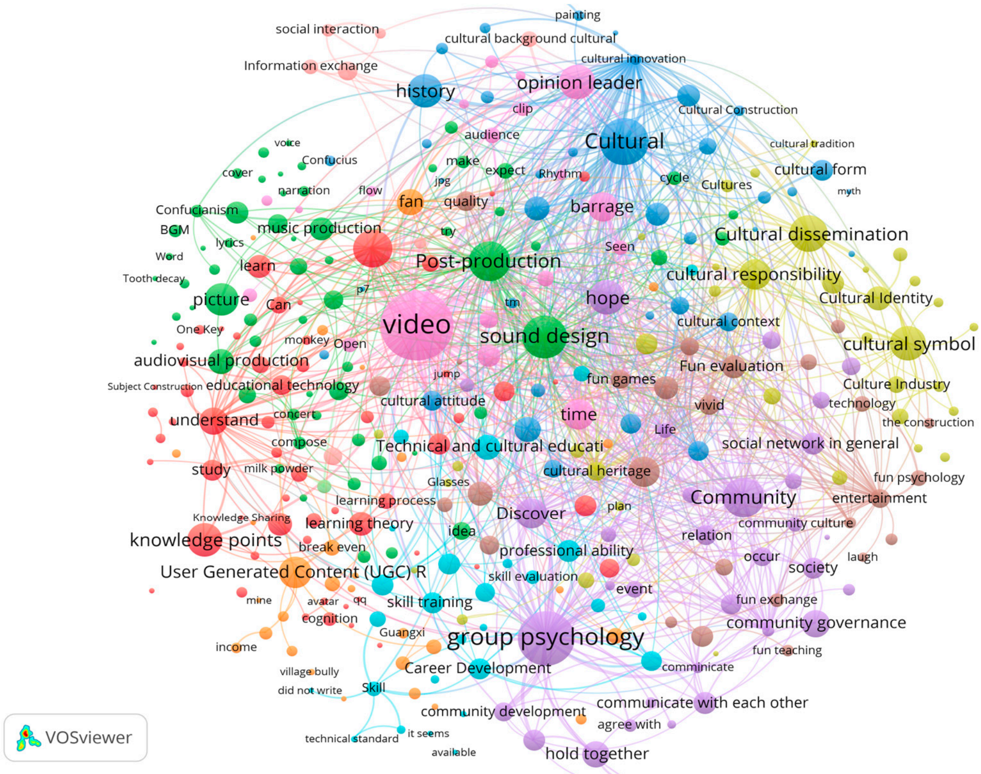

Figure 4. Clustering mapping of UGC short video value elements.

The mapping resulted in the formation of 24 clusters consisting of 358 entity words. These clusters include thematic categories such as UGC short video revenue, the number of fans, and platform interaction. By analyzing the clusters within the topic mapping and identifying representative entity words, a detailed analysis of the characteristics of entities within each topic category is conducted. For this study, 11 clusters were selected based on quantifiability, ease of understanding, and applicability of indicators, aligning with the value assessment of UGC short videos. The selected clusters were renamed to reflect the inherent value characteristics of UGC short videos. Table 1 below presents the renamed clusters:

Table 1. Theme clustering statistics.

Theme Clusters	Theme Clusters	Clustered Entities	Clustering Renaming	Accuracy	Recall Rate	F1
Cluster 1: UGC short video revenue category	15	Work revenue, profit model, economic benefits, income, realization strategy, etc.	Creating financial returns	0.6236	0.4831	0.5341
Cluster 2: Number of fans category	21	Number of fans, growth, number of opinion leaders, popularity, user activity, etc.	Fan popularity	0.6738	0.4496	0.5869
Cluster 3: Platform interaction category	37	Social interaction, social networking, information interaction, user experience, user engagement, etc.	Frequency of interaction	0.7124	0.5489	0.6134
Cluster 4: Platform dissemination category	24	Web communication, social media marketing, media effect, brand communication, brand promotion, etc.	Information reprint volume	0.6807	0.5274	0.5826

Table 1. *Cont.*

Theme Clusters	Theme Clusters	Clustered Entities	Clustering Renaming	Accuracy	Recall Rate	F1
Cluster 5: Cultural content category	35	Cultural communication, cultural responsibility, cultural values, cultural identity, cultural innovation, cultural creativity, etc.	Cultural carrying capacity	0.6471	0.5109	0.5316
Cluster 6: Knowledge category	63	Learning, knowledge points, cognition, knowledge transformation, knowledge, knowledge sharing, etc.	Knowledgeable	0.7201	0.4562	0.5401
Cluster 7: Skill category	28	Skill development, practicality, vocational competence, practical technology, skill demand, etc.	Skillfulness	0.6329	0.5368	0.6339
Cluster 8: Culture category	42	Cultural history, cultural heritage, cultural symbols, cultural awareness, cultural, cultural innovation, cultural expression, cultural construction, etc.	Cultural identity	0.6893	0.4697	0.5243
Cluster 9: Community category	36	Community management, community building, community culture, social network, community culture, etc.	Community identity	0.6472	0.5197	0.5695
Cluster 10: Audiovisual category	61	Audio-visual production, music composition, picture composition, editing techniques, video shooting, audio post-production, etc.	Audiovisuality	0.7452	0.5518	0.5879
Cluster 11: Fun category	27	Fun, entertaining, aesthetic interest, fun communication, fun experience, etc.	Fun	0.7125	0.5156	0.6451

Based on the statistical data, it is evident that the accuracy of the 11 topic clusters related to the value assessment of UGC short videos is above 0.5. This indicates the effectiveness of the text clustering model in producing high-quality clusters and sets the stage for constructing the subsequent UGC short video value assessment index system.

4.1.3. Construction of the UGC Short Video Value Assessment Index System

The information ecological chain theory, introduced in the 1980s, focuses on the interconnectedness of information flow among various information subjects in information systems [45,46]. Analysis shows that the application scenarios based on this theory involve information flow activities among three types of information subjects: information generators, information disseminators, and information consumers. These subjects exhibit a certain degree of interdependence, forming a chain-like relationship. In the context of this study, we posit that UGC short videos also involve interactions among three key subjects during the dissemination process: creators, platforms, and users. Creators primarily undertake the creation and production of UGC short videos, while platforms serve as intermediaries between creators and users, handling the operation and management of UGC short videos. Users, in turn, receive and engage with UGC's short video content.

Thus, this study integrates the 11 identified UGC short video value features and establishes a hierarchical relationship among indicators based on three dimensions: creators, platforms, and users. The study analyzes the similarities and differences between the indicators to optimize and adjust them, ultimately constructing the UGC short video value assessment index system. Refer to Table 2 for more information.

Table 2. UGC short video value assessment index system.

Tier 1 Indicators	Secondary Indicators	Tertiary Indicators	Variables	Explanation of Indicators	Measurement Method
Creators Dimension	Domain visibility	Fan popularity	X1	Reflects the increased visibility that the release of the video brings to the creator	Number of new followers of the creator after the content is published
	Content revenue	Creating financial returns	X2	Reflects the financial return that creators receive from content distribution platforms for newly created content	Total financial return to creators from content distribution platforms for newly created content
Platform Dimension	Internal interaction power	Frequency of interaction	X3	Reflecting users' interactive behavior on original videos with the help of the platform, thus further releasing the vitality of the platform	Total number of interactive behaviors such as user retweeting and collecting UGC short videos
	External Influence	Information reprint volume	X4	Reflecting UGC videos are reported by third-party platforms with their own quality value, bringing influence to the platform	Total number of UGC short videos retweeted and reported by other platforms
	Ideology Leadership	Cultural carrying capacity	X5	Reflecting UGC short videos to spread cultural content and help platform awareness leadership	Frequency of "culture"-related feature words in UGC short video content
User Dimension	Perceived usefulness	Knowledgeable	X6	Reflecting UGC short video content can meet users' knowledge needs	Sentiment value of feature sentences related to "knowledge" in user comments
		Skillfulness	X7	Reflecting that UGC short videos can bring users practical skills to improve	Sentiment value of "skillfulness" in user comments
	Perceptual identity	Cultural identity	X8	Reflecting UGC short videos to improve users' cultural literacy	Sentiment value of "culture"-related special testimonials in user comments
		Community identity	X9	Reflecting the inner sense of belonging that UGC short videos bring to users in the platform community	Sentiment value of "community" related sentences in user comments
	Perceptual entertainment	Audiovisuality	X10	Reflecting that UGC short videos can meet users' entertainment needs in terms of audio and visual	Sentiment value of the special evidence sentences related to "audiovisuality" in user comments
		Fun	X11	Reflecting UGC short video content and its presentation has the quality of attracting viewers' interest and human touch, which can meet users' interesting needs	The sentiment value of the special testimonials related to "interesting" in user comments

4.2. Quantification of the UGC Short Video Value Assessment Index

4.2.1. Data Collection

Bilibili (referred to as B station) is a prominent UGC short video creation and sharing platform in mainland China, offering a wide range of video resources and data. For this study, we utilized a Python crawler tool to randomly select four popular video categories on B station in 2022: "Game Zone", "Knowledge Zone", "Culinary Zone", and "Music Zone". Specifically, we randomly chose videos with traffic exceeding 1 million views as empirical objects, collected between 1 July 2022, and 31 December 2022. The collected data, presented in Table 3, includes the UGC video title, user comments, and the creator's number of fans, among other information. In total, 46,234 comments were obtained.

Table 3. Name of the empirical video.

Video Zone	Number	Video Name
Game Zone	1	[Genshin impact] 3.0 Sumeru treasure chest full collection (achievement number 572)
	2	Werewolf Fool
	3	Bad guys 2
	4	Sheep Village (1)
	5	Werewolf Silly 2
Knowledge Zone	6	What crime was involved in the outrageous Tangshan beating case?
	7	[Liang Ji biological identification] network hot biological identification 38
	8	How to skin care for men? I Two steps to solve 90% of the skin problems
	9	What colony wants a suzerain state to beg for independence? [Odd little country 32]
	10	[Half Buddha] To live is to simmer; to live is everything
Culinary Zone	11	Spend 7 days making a piece of meat! Come in and feel what indulgence means!
	12	Wanzhou Roasted Fish Expo Cook's Visit ¥217
	13	One of the top 10 buffets in the world! What is the experience of eating 7 days and 7 nights on a luxury cruise
	14	After delivering takeout to this soccer team, I broke down.
	15	This may be the world's best food prison! UP for food went to prison
Music Zone	16	"Please bury me in, in that geography".
	17	"Myopia, every day is a gamble"
	18	"It's easy to hide from the open gun, but it's hard to defend against secret love".
	19	A good day ends with a brush in this video
	20	Reenactment of the classic handheld game "Temple Run" sound effects! [MayTreeMayTree]

4.2.2. Indicator Processing Process

Utilizing the LSTM fine-grained sentiment analysis model, we calculated the sentiment value for the indicators related to user comments in the UGC short video value assessment index system. The calculation results are presented in Table 4.

Table 4. Quantification of the sentiment value of user review data indicators.

Video Number	Knowledgeable	Skillfulness	Cultural Identity	Community Identity	Audiovisuality	Fun
1	0.625	0.512	0.41	0.856	0.766	0.899
2	0.213	0.32	0.125	0.812	0.782	0.785
3	0.12	0.23	0.216	0.8	0.654	0.725
4	0.36	0.16	0.149	0.824	0.684	0.763
5	0.1	0.16	0.127	0.743	0.755	0.749
6	0.864	0.886	0.452	0.846	0.672	0.659
7	0.756	0.824	0.654	0.754	0.549	0.632
8	0.8	0.795	0.439	0.632	0.523	0.771
9	0.694	0.721	0.359	0.755	0.421	0.645
10	0.765	0.751	0.548	0.721	0.439	0.663
11	0.213	0.123	0.659	0.522	0.895	0.645
12	0.156	0.278	0.644	0.439	0.862	0.752
13	0.42	0.321	0.721	0.325	0.821	0.712
14	0.126	0.11	0.664	0.267	0.793	0.6
15	0.249	0.15	0.545	0.545	0.766	0.645
16	0.521	0.645	0.894	0.751	0.64	0.756
17	0.623	0.546	0.887	0.669	0.554	0.754
18	0.324	0.325	0.756	0.743	0.61	0.69
19	0.632	0.469	0.845	0.62	0.557	0.61
20	0.559	0.477	0.76	0.6	0.68	0.68

The structured data indicators are quantified as follows: the "FeiGua Data" platform provides statistics on the number of new fans gained by creators after releasing a video, enabling quantification of "fan popularity" of the creator; the popular revenue conversion formula of B station, "1w plays can earn 30 RMB", is used to quantify the "economic return on creation" for each video by the creator; the interaction index of each UGC short video on the "FeiGua Data" platform is calculated by considering users' interactive behaviors such as likes, favorites, and coins, which quantifies the "frequency of interaction" on the platform; the "CMHI" data detection and analysis platform calculates the interaction index of third-party platforms for each UGC short video released through the B station platform. This is performed by counting the total number of reposts and reports by third-party platforms, which quantifies the "information repost volume" of the platform. To quantify the platform's "cultural carrying capacity", video content is converted into textual information, and the TF-IDF method is applied to identify culture-related keywords for quantification. The total count of these keywords in the text data determines the quantification.

To facilitate the comparison and assignment of indicators, all quantified indicators were separately normalized before the combination assignment, resulting in the UGC short video value indicator matrix. By considering subjective and objective assignments and calculating the weights for indicator combinations, the results in Table 5 are obtained.

Table 5. Weighting weights for indicator combinations.

Tertiary Indicators	CRITIC Objective Weights	AHP Subjective Weights	Portfolio Weights
Fan popularity	0.0604	0.0870	0.0564
Creating financial returns	0.0603	0.0663	0.0429
Frequency of interaction	0.0847	0.0902	0.0820
Information reprint volume	0.0676	0.0657	0.0477
Cultural carrying capacity	0.1049	0.0730	0.0822
Knowledgeable	0.1020	0.1322	0.1447
Skillfulness	0.0995	0.1469	0.1568
Cultural identity	0.1368	0.1018	0.1495
Community identity	0.0750	0.0535	0.0431
Audiovisuality	0.1358	0.0764	0.1114
Fun	0.0724	0.1064	0.0827

Based on the quantified indexes and weights of the UGC short video value assessment indexes, the final weighting matrix for UGC short video value is derived. The gray correlation coefficient of each index is calculated and inputted into the VIKOR model to obtain the group effect value S, individual regret value R, and benefit ratio Q for the 20 UGC short videos.

4.3. UGC Short Video Value Assessment

4.3.1. Assessment Principles

The VIKOR method calculates the benefit ratio value (Q) by measuring the distances between the empirical UGC videos and the positive and negative ideal solutions. This ratio serves as the value of the UGC videos. The group effect value (S) represents the sum of the distance ratios to the positive ideal solution. A smaller value indicates that the research object is closer to the "optimal solution". The individual regret rate (R) represents the maximum distance ratio between the evaluation object and the positive ideal solution. A smaller value suggests that the research object is closer to its own "optimal solution". A lower value signifies a higher value of the research object itself, serving as an important criterion for judging the value of UGC short videos.

4.3.2. Analysis of Assessment Results

In this paper, we utilize element mining technology to extract the components of UGC short video value indexes and build the UGC short video value assessment index system. We obtained the UGC short video value assessment results and the weighted ranking results of UGC short video indexes, which are presented in Figures 5 and 6.

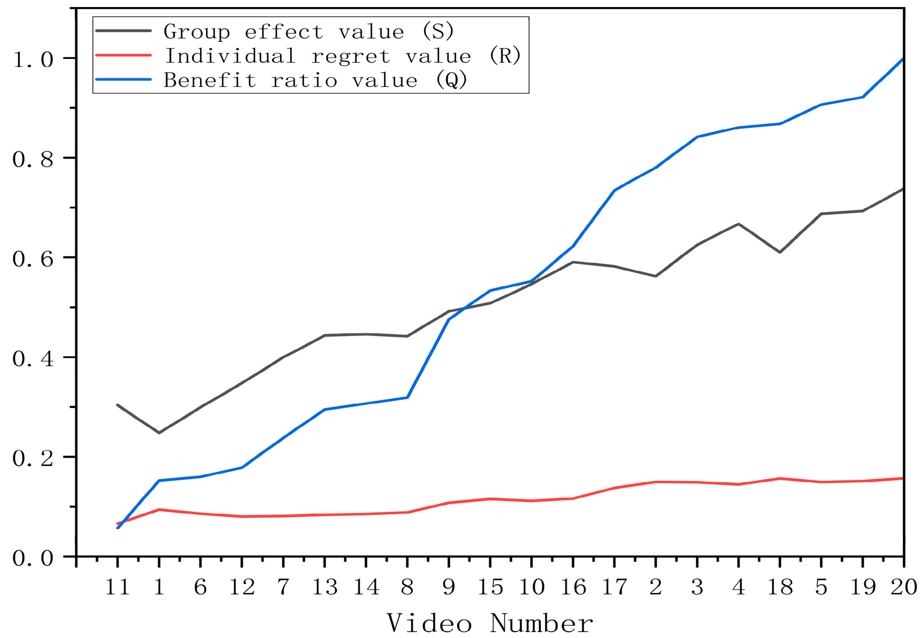

Figure 5. UGC short video value assessment result chart.

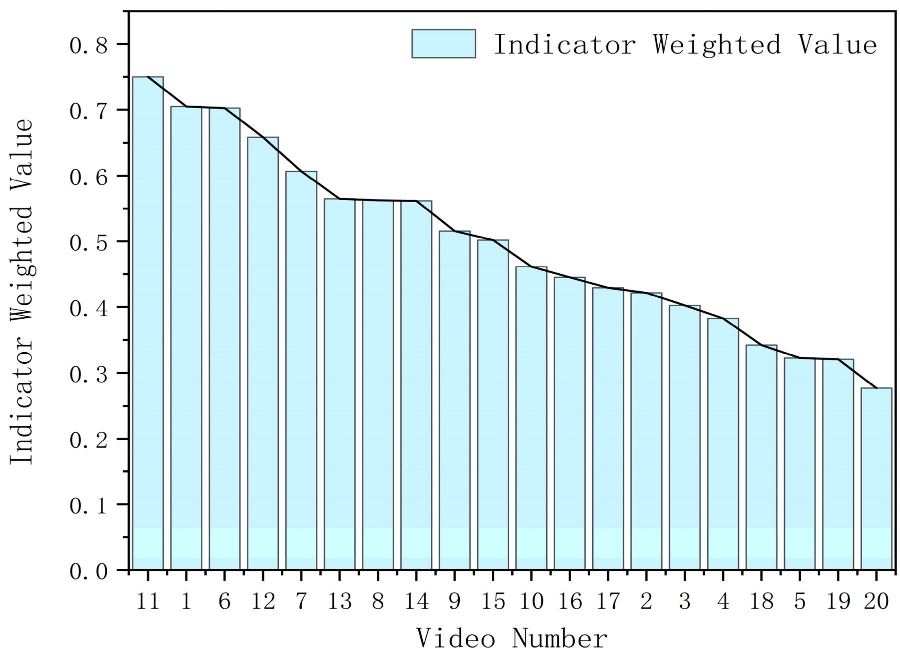

Figure 6. Bar chart of weighted results of UGC short video indicators.

Based on the principle of UGC short video value assessment, the benefit ratio value (Q) reflects the magnitude of the video's corresponding value, with smaller values indicating higher video value and higher ranking. As shown in Figure 5, the value ranking (using video numbers instead of video names) results for the 20 selected UGC short videos in this paper are as follows:

11 > 1 > 6 > 12 > 7 > 13 > 14 > 8 > 9 > 15 > 10 > 16 > 17 > 2 > 3 > 4 > 18 > 5 > 19 > 20.

As shown in Figure 6, the final results of the indicator weighting values represent the relative value of UGC short videos, with larger values indicating a higher video value. The ranking of the indicator weighting values for the 20 selected UGC short videos in this paper is as follows:

$$11 > 1 > 6 > 12 > 7 > 13 > 8 > 14 > 9 > 15 > 10 > 16 > 17 > 2 > 3 > 4 > 18 > 5 > 19 > 20.$$

The results indicate that the weighted ranking of UGC short video indicators and the ranking of UGC short video value show a similar trend, with slight differences in the ranking order of individual videos. This suggests that the UGC short video value assessment model proposed in this paper focuses on the mutual influence among the indicators in the UGC short video value assessment system while ensuring evaluation accuracy. The comparison and analysis of the best and worst result items in the assessment object demonstrate the superiority of the VIKOR model in handling multi-attribute decision problems.

This paper selected 20 UGC short videos, which were categorized into four partitions: game, knowledge, food, and music. To analyze the value performance of UGC short videos in each partition, the weighted quantification results of the videos were divided based on the creator dimension, platform dimension, and user dimension. Radar charts were created using the average values of the three dimensions for each thematic partition, as presented in Figure 7. These radar charts effectively illustrate the value and performance of short videos in different partitions.

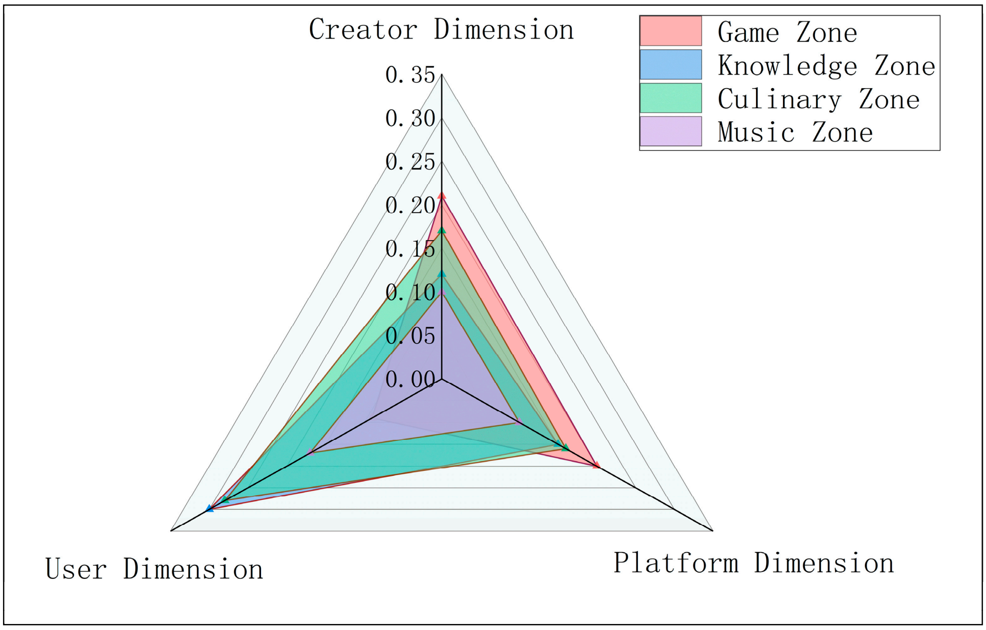

Figure 7. Performance of UGC short video values in different partitions.

5. Discussion

In this study, we develop a comprehensive set of models for the scientific evaluation of the value of UGC short videos. This involves three processes: (1) constructing the index system for evaluating the value of UGC short videos based on element mining; (2) quantifying the indexes used to evaluate the value of UGC short videos; and (3) conducting an evaluation of the value of UGC short videos.

Firstly, we initiate the process by analyzing the professional text of UGC short videos. We employ text clustering algorithms and visualization technology for theme mapping to thoroughly explore the value characteristics of UGC short videos, thus constructing the UGC short video value assessment index system. Secondly, we quantitatively evaluate the indexes using heterogeneous data from multiple sources. For structured data indicators, we statistically quantify them using data from various platforms, while the LSTM fine-grained sentiment analysis model calculates the sentiment value of unstructured data indicators to enable precise quantification. Finally, by assigning weights to the subjective-objective combinations of all indicators, we utilize the gray correlation coefficient of the improved VIKOR model to accurately assess the value of UGC short videos. The empirical analysis yields the following findings:

Finding 1: Deconstructing the Value Characteristics of UGC Short Videos.

In contrast to single-dimensional value evaluation, this study employed value element mining technology to discover ecological value elements in UGC short videos and establish a multi-dimensional value evaluation index system. We identified that the current value elements of UGC short videos in China mainly encompass three dimensions: the creators, the platforms, and the users. These dimensions consist of the following value elements: The creator dimension includes fan popularity and creative economic return; the platform dimension includes interaction frequency, information reprint volume, and cultural carrying capacity; and the user dimension includes knowledge, skills, cultural identity, community identity, audio-visual appeal, and entertainment value. The value of UGC short videos should not be assessed solely based on their content and technical attributes. Instead, it is an organic unity resulting from the interactions among UGC short video creators, communication platforms, and users [47]. This system aligns with the interactive nature of UGC short videos as a form of Internet communication [48] and holds valuable implications for the future value evaluation of UGC short videos;

Finding 2: Differentiated Value Performance of UGC Short Videos in Mainstream Categories.

1. Game Zone: Among the mainstream video divisions, the game partition exhibits the highest indicator weighting results in the creator and platform dimensions, suggesting that game videos in this category can generate more tangible benefits and value for creators and platforms [49]. However, compared to other partitions, the weighted results of indicators in the user dimension are average, indicating a relatively lower value contribution of game short videos to users. This may be attributed to the emphasis placed on short-term revenue and video traffic growth, which overlooks the genuine needs and experiences of users. Consequently, the overall short video value score of the division is relatively low [50]. In the future, practitioners should create game-short videos with a user-centric approach, focusing on delivering new audiovisual experiences while addressing the diverse knowledge and cultural needs of users to enhance their overall experience. It is noteworthy that a domestic game-themed short video, "Genshin impact", ranks second in value. Compared to other game-short videos, it stands out in terms of "cultural carrying capacity" and "cultural identity" index values. This highlights the importance of integrating cultural elements into the creation of high-value game-short videos, as it enhances the artistic and aesthetic value of the videos, attracting greater user attention and engagement [51]. Moreover, cultural elements serve as a valuable source of inspiration for game-short video creation [52], fostering more innovative ideas and expanding the imaginative space, thereby further enhancing the value and distinctiveness of video content;

2. Knowledge Zone: The high user value scores of short videos in the knowledge partition indicate that these videos effectively fulfill viewers' knowledge and skill-related needs. This exceptional user value performance also reflects the growing demand for quality knowledge content among users today. With the widespread popularity of mobile internet, people increasingly turn to innovative media such as short videos to access information and acquire knowledge [53], further bolstering the prominence of short videos in the knowledge domain. Additionally, knowledge-based short videos provide an excellent platform for knowledge creators to share their expertise and insights, serving as a conduit for the dissemination of wisdom. Popularizing knowledge through short videos not only facilitates rapid learning and mastery for viewers but also enhances the reputation and influence of knowledge creators [2]. For instance, "Luo Xiang speaks criminal law", a knowledge-focused up-owner examined in this study, has garnered the distinction of being the fastest video creator to reach 10 million fans in the history of B-station. However, in terms of the creator dimension and the platform dimension, the performance of knowledge-based short videos remains average. This indicates the need for creators to align their

video content with current social trends and produce more targeted videos [54]. Moreover, creators should delve deeper into professional knowledge to create videos with both depth and breadth, thereby delivering a more comprehensive and diverse video content experience. Additionally, platforms should enhance the promotion of high-quality content and provide robust resource support for outstanding creators, ensuring that viewers find delight in their knowledge acquisition journey. In today's society, knowledge-based short videos have become a preferred leisure choice for individuals, contributing to the accumulation of knowledge and fostering innovation for societal development. Consequently, creators and platforms should continuously optimize and improve content value while enhancing viewers' overall viewing experience;

3. Gourmet Zone: Short videos in the food category receive high user scores, indicating that the creators of these videos effectively showcase their cooking skills and professionalism from a user experience perspective while teaching viewers cooking techniques [55]. Additionally, the data demonstrates that these food-related short videos outperform other indicators in terms of "cultural carrying capacity", suggesting that they provide viewers with rich cultural content in the context of China's culinary culture [56], further strengthening their cultural identity. The study's findings also highlight the exceptional performance of short videos in the food category in terms of "interestingness". Creators are gradually departing from traditional cooking processes, employing humorous and witty copywriting as well as exquisite editing skills to create highly engaging audio-visual experiences [57]. The content covers various topics such as "Production", "Taste Exploration", and "Judging", allowing viewers to appreciate the beauty of cooking and adding diversity and richness to the video topics. Looking ahead, creators should continue to explore new topics and formats in food-related short videos, showcasing unique ingredients and production methods. Additionally, incorporating interspersed storytelling elements can bring food to life, offering viewers a fresh and unprecedented experience [58];

4. Music Zone: Short videos in the music category receive relatively high scores in the user dimension. Compared to other mainstream video divisions, these videos demonstrate higher quantitative results in the "audio-visual" index. This indicates the importance of enhancing the musical presentation in music videos [59], showcasing the performers' virtuosity, deep emotions, and excellent artistic performances. Creators can utilize camera work, editing techniques, and post-production to further emphasize the inherent power of the music [60]. The study's results also suggest that the performance of music videos is moderate in the creator dimension and platform dimension, indirectly highlighting the increasing significance of music subject matter and values in future music video production [61]. Creators should ensure that the music aligns with the video's theme, fully expresses their own emotions, resonates with the audience's values, and reflects contemporary social trends [62]. Additionally, creators should continuously expand their video production skills by incorporating storytelling elements [63], enriching the content and viewing value of music videos. This approach will make short music videos more captivating and resonant with viewers, thereby contributing to the overall enhancement of video value.

This study holds both theoretical and practical significance. Theoretically, it establishes a new paradigm for assessing the value of UGC short videos. By mining key value elements from authoritative texts, employing differentiated indicator quantification methods for multi-source data, and conducting value assessment, the study effectively addresses the issues of strong subjectivity, insufficient quantification of key features, and weak interpretability of assessment results in the current UGC short video value assessment indicators. Consequently, it provides a novel assessment theory and direction for future evaluations of the value of short videos and other related Internet content communication media. In practical terms, this study introduces a new and efficient video value assessment tool for short video operation platforms. By applying our study, platforms can effectively identify high-value short videos and curtail the dissemination of videos influenced by misguided opinions, thereby enhancing their reputation and credibility.

Furthermore, the research method presented in this paper can enhance the diversity of video content on the platform, further improving the user experience and enhancing the platform's operational capabilities.

6. Conclusions

However, this paper still has some shortcomings and room for improvement. On the one hand, while we employ text clustering algorithms and topic mapping visualization technology to mine elements for UGC short video value assessment, effectively addressing the issues of strong subjectivity in indicators and insufficient theoretical basis, we solely rely on literature and expert review articles. In the future, we should expand our data sources to fully explore additional elemental information relevant to UGC short video value assessment. On the other hand, the empirical research on the proposed method focuses only on Bilibili, a mainstream domestic short video platform. To verify the method's scientificity and reliability, future analyses should be carried out on other mainstream short video platforms.

Author Contributions: Conceptualization, J.F. and Y.N.; methodology, J.F.; software, J.Z.; validation, J.F., Y.N. and J.Z.; formal analysis, J.F.; investigation, Y.N.; resources, J.Z.; data curation, Y.N.; writing—original draft preparation, J.Z.; writing—review and editing, J.F. and Y.N.; visualization, J.F.; supervision, J.Z.; project administration, Y.N.; funding acquisition, Y.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by the National Key R&D Program of China (2021YFF0900200).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, Y.; Chiu, D.K.W.; Ho, K.K.W. Short-form videos for public library marketing: Performance analytics of douyin in China. *Appl. Sci.* **2023**, *13*, 3386. [CrossRef]
2. Li, L.; Dai, D.; Liu, H.; Yuan, Y.; Ding, L.; Xu, Y. Research on Short Video Hotspot Classification Based on LDA Feature Fusion and Improved BiLSTM. *Appl. Sci.* **2022**, *12*, 11902. [CrossRef]
3. "2023 China Network Audiovisual Development Research Report" Released: Over 1 Billion Users in My Country Are Watching Short Videos. Available online: https://new.qq.com/rain/a/20230329A09P8200 (accessed on 2 March 2023).
4. He, M.; Ma, C.; Wang, R. A Data-Driven Approach for University Public Opinion Analysis and Its Applications. *Appl. Sci.* **2022**, *12*, 9136. [CrossRef]
5. Psallidas, T.; Koromilas, P.; Giannakopoulos, T.; Spyrou, E. Multimodal summarization of user-generated videos. *Appl. Sci.* **2021**, *11*, 5260. [CrossRef]
6. Han, W. Research on Short Video Marketing Model in the New Media Era. In Proceedings of the 2022 International Conference on Comprehensive Art and Cultural Communication (CACC 2022), Online, 14 May 2022; Atlantis Press: Amsterdam, The Netherlands, 2022; pp. 195–198.
7. Sun, W.; Min, X.; Lu, W.; Zhai, G. A deep learning based no-reference quality assessment model for UGC videos. In Proceedings of the 30th ACM International Conference on Multimedia, Lisboa, Portugal, 10–14 October 2022; pp. 856–865.
8. Tu, Z.; Wang, Y.; Birkbeck, N.; Adsumilli, B.; Bovik, A.C. UGC-VQA: Benchmarking blind video quality assessment for user generated content. *IEEE Trans. Image Process.* **2021**, *30*, 4449–4464. [CrossRef]
9. Nanne, A.J.; Antheunis, M.L.; Van Der Lee, C.G.; Postma, E.O.; Wubben, S.; Van Noort, G. The use of computer vision to analyze brand-related user generated image content. *J. Interact. Mark.* **2020**, *50*, 156–167. [CrossRef]
10. Wang, Y.; Ke, J.; Talebi, H.; Yim, J.G.; Birkbeck, N.; Adsumilli, B.; Milanfar, P.; Yang, F. Rich features for perceptual quality assessment of UGC videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13435–13444.
11. Su, S.; Yan, Q.; Zhu, Y.; Zhang, C.; Ge, X.; Sun, J.; Zhang, Y. Blindly Assess Image Quality in the Wild Guided by a Self-Adaptive Hyper Network. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 3664–3673.

12. Meng, J.; Zhang, X.; Tan, X. Evaluation system of short video influence based on TOPSIS-Kmeans method. In Proceedings of the 2022 China Automation Congress (CAC), Xiamen, China, 25–27 November 2022; IEEE: Seattle, WA, USA, 2022; pp. 1989–1995.
13. Chen, H.; Wang, M.; Zhang, Z. Research on rural landscape preference based on TikTok short video content and user comments. *Int. J. Environ. Res. Public Health* **2022**, *19*, 10115. [CrossRef]
14. Yu, K.; Zhang, L.; Yang, S. Research on user satisfaction of mobile short video based on rough set. In Proceedings of the 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 14–16 April 2022; IEEE: Seattle, WA, USA, 2022; pp. 211–215.
15. Mohammad, J.; Quoquab, F.; Thurasamy, R.; Alolayyan, M.N. The effect of user-generated content quality on brand engagement: The mediating role of functional and emotional values. *J. Electron. Commer. Res.* **2020**, *21*, 39–55.
16. Zhang, H.; Zhang, X.; Zhou, H.; Wang, X. Research on the evaluation of short video communication effect based on user preference and content characteristics. *Libr. Inf. Serv.* **2020**, *64*, 81.
17. Cai, Q.; Tan, W. Box Office Forecast Model Based on Random Forest and BP Neural Network. In Proceedings of the 5th International Conference on Big Data Technologies, Qingdao, China, 23–25 September 2022; pp. 69–75.
18. Manikandan, P.; Manimuthu, A.; Rajam, S. Prediction of YouTube View Count using Supervised and Ensemble Machine Learning Techniques. In Proceedings of the 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 13–15 December 2022; IEEE: Seattle, WA, USA, 2022; pp. 1038–1042.
19. Leszczuk, M.; Kobosko, M.; Nawała, J.; Korus, F.; Grega, M. "In the Wild" Video Content as a Special Case of User Generated Content and a System for Its Recognition. *Sensors* **2023**, *23*, 1769. [CrossRef]
20. Rui, L.T.; Afif, Z.A.; Saedudin, R.D.R.; Mustapha, A.; Razali, N. A regression approach for prediction of Youtube views. *Bull. Electr. Eng. Inform.* **2019**, *8*, 1502–1506. [CrossRef]
21. Gupta, V.; Diwan, A.; Chadha, C.; Khanna, A.; Gupta, D. Machine Learning enabled models for YouTube Ranking Mechanism and Views Prediction. *arXiv* **2022**, arXiv:2211.11528.
22. Mekouar, S.; Zrira, N.; Bouyakhf, E.H. Popularity prediction of videos in youtube as case study: A regression analysis study. In Proceedings of the 2nd International Conference on Big Data, Cloud and Applications, Tetouan, Morocco, 29–30 March 2017; pp. 1–6.
23. Wu, X.; Tian, M.; Zhai, L. ICANet: A Method of Short Video Emotion Recognition Driven by Multimodal Data. *arXiv* **2022**, arXiv:2208.11346.
24. Ren, C.; Zhang, B. The factors of user emotion and behaviour solidification caused by information dissemination in network media. In Proceedings of the Second International Conference on Green Communication, Network, and Internet of Things (CNIoT 2022), Xiangtan, China, 8 March 2023; SPIE: Bellingham, WA, USA, 2023; Volume 12586, pp. 238–244.
25. Yu, S.; He, S.; Cai, Z.; Lee, I.; Naseriparsa, M.; Xia, F. Exploring public sentiment during COVID-19: A Cross Country Analysis. *IEEE Trans. Comput. Soc. Syst.* **2022**, *10*, 1083–1094. [CrossRef]
26. Wang, W.; Wei, T.; Yu, S.; Chen, J.; Yang, X. A Telecommuting Experience Service Design Decision Model Based on BP Neural Network. *Psychol. Res. Behav. Manag.* **2022**, *15*, 3147–3166. [CrossRef]
27. Zhang, Z.; Luo, M.; Hu, Z.; Niu, H. Textual emotional tone and financial crisis identification in Chinese companies: A multi-source data analysis based on machine learning. *Appl. Sci.* **2022**, *12*, 6662. [CrossRef]
28. Jia, F.; Chen, C.C. Emotional characteristics and time series analysis of Internet public opinion participants based on emotional feature words. *Int. J. Adv. Robot. Syst.* **2020**, *17*, 1729881420904213. [CrossRef]
29. Li, H.; Chen, Q.; Zhong, Z.; Gong, R.; Han, G. E-word of mouth sentiment analysis for user behavior studies. *Inf. Process. Manag.* **2022**, *59*, 102784. [CrossRef]
30. Jang, B.; Kim, M.; Harerimana, G.; Kang, S.U.; Kim, J.W. Bi-LSTM model to increase accuracy in text classification: Combining Word2vec CNN and attention mechanism. *Appl. Sci.* **2020**, *10*, 5841. [CrossRef]
31. Khan, L.; Amjad, A.; Afaq, K.M.; Chang, H.-T. Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media. *Appl. Sci.* **2022**, *12*, 2694. [CrossRef]
32. Shaikh, S.; Daudpota, S.M.; Imran, A.S.; Kastrati, Z. Towards improved classification accuracy on highly imbalanced text dataset using deep neural language models. *Appl. Sci.* **2021**, *11*, 869. [CrossRef]
33. Nazim, M.; Mohammad, C.W.; Sadiq, M. A comparison between fuzzy AHP and fuzzy TOPSIS methods to software requirements selection. *Alex. Eng. J.* **2022**, *61*, 10851–10870. [CrossRef]
34. Gao, J.; Ren, H.; Cai, W. Inundation risk assessment based on G-DEMATEL-AHP and its application to Zhengzhou flooding disaster. *Sustain. Cities Soc.* **2022**, *86*, 104138.
35. Wang, S.; Wei, G.; Lu, J.; Wu, J.; Wei, C.; Chen, X. GRP and CRITIC method for probabilistic uncertain linguistic MAGDM and its application to site selection of hospital constructions. *Soft Comput.* **2022**, *26*, 237–251. [CrossRef]
36. Kumaran, S. Financial performance index of IPO firms using VIKOR-CRITIC techniques. *Financ. Res. Lett.* **2022**, *47*, 102542. [CrossRef]
37. Bhadra, D.; Dhar, N.R.; Salam, M.A. Sensitivity analysis of the integrated AHP-TOPSIS and CRITIC-TOPSIS method for selection of the natural fiber. *Mater. Today Proc.* **2022**, *56*, 2618–2629. [CrossRef]
38. Li, Y.; Shen, J.; Huang, M.; Peng, Z. Debris Flow Classification and Risk Assessment Based on Combination Weighting Method and Cluster Analysis: A Case Study of Debris Flow Clusters in Longmenshan Town, Pengzhou, China. *Appl. Sci.* **2023**, *13*, 7551. [CrossRef]

39. Zhong, S.; Chen, Y.; Miao, Y. Using improved CRITIC method to evaluate thermal coal suppliers. *Sci. Rep.* **2023**, *13*, 195. [CrossRef] [PubMed]
40. Bera, B.; Shit, P.K.; Sengupta, N.; Saha, S.; Bhattacharjee, S. Susceptibility of deforestation hotspots in Terai-Dooars belt of Himalayan Foothills: A comparative analysis of VIKOR and TOPSIS models. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 8794–8806. [CrossRef]
41. Zhong, X.; Cheng, Y.; Yang, J.; Tian, L. Evaluation and Optimization of In-Vehicle HUD Design by Applying an Entropy Weight-VIKOR Hybrid Method. *Appl. Sci.* **2023**, *13*, 3789. [CrossRef]
42. Wang, J.; Ye, N.; Ge, L. Steady-state power quality synthetic evaluation based on the triangular fuzzy BW method and interval VIKOR method. *Appl. Sci.* **2020**, *10*, 2839. [CrossRef]
43. Jin, B.; Liu, Z.; Qin, J. Two-stage real-time track correlation algorithm based on gray correlation. *Acta Armamentarii* **2020**, *41*, 1330.
44. Cheng, Y.-T.; Liang, R.-Q.; Ho, M.-C.; Wu, C.-H. Sensory Evaluation of Low-Carbon City Tourism by Gray Relational Analysis. In Proceedings of the E3S Web of Conferences, Online, 13 January 2021; EDP Sciences: Les Ulis, France, 2021; Volume 228, p. 01003.
45. Deng, C.; Yu, Q.; Luo, G.; Zhao, Z.; Li, Y. Big data-driven intelligent governance of college students' physical health: System and strategy. *Front. Public Health* **2022**, *10*, 924025. [CrossRef]
46. Qu, J. Research on Mobile learning in a teaching information service system based on a big data driven environment. *Educ. Inf. Technol.* **2021**, *26*, 6183–6201. [CrossRef]
47. Xie, X.-Z.; Tsai, N.-C.; Xu, S.-Q.; Zhang, B.-Y. Does customer co-creation value lead to electronic word-of-mouth? An empirical study on the short-video platform industry. *Soc. Sci. J.* **2019**, *56*, 401–416. [CrossRef]
48. Zhou, Q. Understanding User Behaviors of Creative Practice on Short Video Sharing Platforms—A Case Study of TikTok and Bilibili. Master's Thesis, University of Cincinnati, Cincinnati, OH, USA, 2019.
49. Kleinlein, R.; Luna-Jiménez, C.; Arias-Cuadrado, D.; Ferreiros, J.; Fernández-Martínez, F. Topic-oriented text features can match visual deep models of video memorability. *Appl. Sci.* **2021**, *11*, 7406. [CrossRef]
50. Mustač, K.; Bačić, K.; Skorin-Kapov, L.; Sužnjević, M. Predicting player churn of a Free-to-Play mobile video game using supervised machine learning. *Appl. Sci.* **2022**, *12*, 2795. [CrossRef]
51. Kaye, D.B.V.; Zeng, J.; Wikstrom, P. *TikTok: Creativity and Culture in Short Video*; John Wiley & Sons: Hoboken, NJ, USA, 2022.
52. Bach, P.D. Bondy Valdovinos Kaye, Jing Zeng, and Patrik Wikström, TikTok: Creativity and Culture in Short Video. *Int. J. Commun.* **2023**, *17*, 3.
53. Liu, X.; Liu, S.; Ma, Z. A Framework for Short Video Recognition Based on Motion Estimation and Feature Curves on SPD Manifolds. *Appl. Sci.* **2022**, *12*, 4669. [CrossRef]
54. Zheng, C. Research on the flow experience and social influences of users of short online videos. A case study of DouYin. *Sci. Rep.* **2023**, *13*, 3312. [CrossRef]
55. Deng, D.S.; Seo, S.; Li, Z.; Austin, E.W. What people TikTok (Douyin) about influencer-endorsed short videos on wine? An exploration of gender and generational differences. *J. Hosp. Tour. Technol.* **2022**, *13*, 683–698. [CrossRef]
56. Cao, S.; Foth, M.; Powell, W.; McQueenie, J. What are the effects of short video storytelling in delivering Blockchain-credentialed Australian beef products to China? *Foods* **2021**, *10*, 2403. [CrossRef] [PubMed]
57. de Araújo, T.P.; de Moraes, M.; Afonso, C.; Rodrigues, S. Food and nutrition literacy: An education booklet and video on ultra-processed food. *Eur. J. Public Health* **2021**, *31* (Suppl. S3), ckab164.221.
58. Li, J.; Li, C.; Liu, J.; Zhang, J.; Zhuo, L.; Wang, M. Personalized mobile video recommendation based on user preference modeling by deep features and social tags. *Appl. Sci.* **2019**, *9*, 3858. [CrossRef]
59. Becerra Martinez, H.; Hines, A.; Farias, M.C.Q. Perceptual quality of audio-visual content with common video and audio degradations. *Appl. Sci.* **2021**, *11*, 5813. [CrossRef]
60. Surís, D.; Vondrick, C.; Russell, B.; Salamon, J. It's Time for Artistic Correspondence in Music and Video. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10564–10574.
61. Vernallis, C. The aesthetics of music video: An analysis of Madonna's 'Cherish'. *Pop. Music* **1998**, *17*, 153–185. [CrossRef]
62. Wang, W.; Wu, J. Research Perspectives on TikTok & Its Legacy Apps | Short Video Platforms and Local Community Building in China. *Int. J. Commun.* **2021**, *15*, 23.
63. Mystakidis, S.; Filippousis, G.; Tolis, D.; Tseregkouni, E. Playful metaphors for narrative-driven e-learning. *Appl. Sci.* **2021**, *11*, 11682. [CrossRef]