

Research Article

Information Fusion Algorithm for Big Data in Digital Publishing Industry Chain

Haixiang He 

School of Network Communication, Zhejiang Yuexiu University, Shaoxing 312000, Zhejiang, China

Correspondence should be addressed to Haixiang He; 20031015@zyufl.edu.cn

Received 13 March 2021; Revised 2 April 2021; Accepted 7 April 2021; Published 17 April 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Haixiang He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies the information of big data in the digital publishing industry chain and adopts advanced algorithms for its fusion calculation. The basic theory of digital publishing ecological chain is dissected, the construction requirements, construction methods, and construction paths of digital publishing ecological chain are analysed, and feasible construction measures are proposed. It also defines the connotation of the fusion of knowledge services between publishing institutions and libraries in the digital era; then analyses the characteristics and principles of the fusion of knowledge services between publishing institutions and libraries in the digital era; and finally sorts out the theoretical foundations such as synergy theory, information integration service theory, and game theory. Meanwhile, this paper also studies the flow of digital publishing resources and empirically analyses the eco-efficiency of digital publishing eco-chain by using metabolic network analysis, population analysis, and life cycle analysis, and it finds the eco-efficiency problems of the existing e-book eco-chain from its analysis. Finally, the imbalance in the digital publishing ecological chain and its hazards are analysed, and specific regulation and optimization measures are proposed. The research in this paper makes up for the deficiencies of related studies and can well solve the problems that exist in the previous and provide theoretical support for the healthy development of digital publishing enterprises.

1. Introduction

In the digital era, the application innovation and integration development of technologies such as big data and cloud computing, and the diversified and personalized development of users' knowledge needs have put forward new requirements for both publishers and libraries. To seek development, publishers and libraries must reach a consensus on the integration of knowledge services and realize digital transformation and upgrading as well as the innovation of knowledge services [1]. With the development of the mobile Internet and the progress of related digital technologies, the reading and consumption habits of audiences have changed greatly, and the demand for information and knowledge is more inclined to personalization [2]. The growing market demand of users for personalization, specialization, and high-precision docking has driven knowledge services forward.

These ecological problems seriously plague the development of digital publishing enterprises and restrict the transformation and upgrading of the whole publishing industry, which need to be solved urgently [3]. However, the existing digital publishing theories have many defects in dealing with such problems and lack effective means to solve them. Therefore, it is necessary to draw on new theories to conduct research on such problems and provide theoretical support for the healthy development of digital publishing enterprises [4]. This paper draws on the theories and methods of ecological research, starts from the ecological problems faced by digital publishing, and conducts research on the digital publishing ecological chain formed by digital publishing enterprises in digital publishing activities, takes the construction of the chain, resource flow, and ecological efficiency as the breakthrough of the research, combines them with case studies to provide a theoretical basis for analysing and solving such problems, and also provides

support for the sustainable development of the digital publishing industry.

Especially, at the current early stage of digital publishing development, this kind of practical guidance is promising. The second is to discover new research methods to serve the industrial analysis of digital publishing, eliminate hidden kinds of problems, and make digital publishing more vital [5]. The fundamental purpose of this paper is to promote the benign development of digital publishing. This paper is dedicated to creating a benign ecological cycle model of digital publishing and providing model support for the benign development of digital publishing.

At present, digital publishing enterprises, which are still in the primary stage of digital transformation, are relatively backward in their digital publishing platform information construction compared with other industries and have not yet formed a systematic customer management program, and the quality of customer behaviour data and efficient data analysis techniques cannot be guaranteed, so the management of customer data information of most publishing enterprises is in a bottleneck state. However, in the digital publishing industry, some digital publishing enterprises have already started to find a way to break through this bottleneck by referring to the implementation of customer relationship management strategies in other industries.

In order to solve the optimal solution of the big data information fusion algorithm in the digital publishing industry chain, the existing algorithms cannot meet the requirements of high precision and high efficiency in reality, so we compare several mainstream algorithms and fusion, the best algorithm is obtained by comparing the accuracy and efficiency of information fusion, and the best algorithm is used in the actual process.

2. Current Status of Research

For example, Davidson defined the concept of data warehouse (DW) and proposed a solution to the predata preparation problem of data mining [6]. Song pointed out the four main techniques used in the process of data mining [7]. Yang and other scholars elaborated the relevant aspects of data mining, etc. [8]. The authors of this paper have also discussed the content of data mining. In terms of software application development of data mining systems, the following are relatively influential ones in the world: Clementine from SPSS, Enterprise Miner from SA, and Intelligent Miner from IBM. In the actual software application process, there are many classic cases related to data mining [9]. The data mining model built based on application software not only effectively improves the proportion of data error detection but also reduces the time spent on data mining analysis with the improvement of data quality control automation, which reduces the time cost of enterprises [10]. The research on data mining mainly revolves around the introduction and correction of the elaboration of data mining algorithms and the implementation of data mining in specific industry sectors. Also, Chinese research scholars have conducted detailed research on data warehousing, such as Yudhistrya's research on data warehouse design methods

for Internet-based information systems, and proposed the development and implementation of information systems for applications such as analytical processing and decision support systems based on this system [11]. Ji and other scholars proposed a data warehouse-based customer analysis system [12]. Van combines the macro policy background and China's digital publishing industry situation and believes that China has made new attempts in content resources, product forms, and marketing methods among the convergent publishing based on knowledge services, but in the future, development should fully consider the characteristics of mobile Internet and improve the use of technical tools, content organization standards, product production process, and user interaction [13].

This capability requires the digital publishing ecosystem to provide sustainable development for digital publishing, while also allowing the entire digital publishing ecosystem to obtain a circular development mechanism. According to Mohamed, the new medium is nothing but new distribution platforms and new distribution channels [14]. The increase of channels will stimulate a stronger demand for excellent content, while the competition of platforms will continuously enhance the value of content and achieve the elimination of the best and the worst, thus producing a virtuous cycle of the digital publishing ecosystem [15].

Zhao et al. argue that the business ecology of publishing in the online environment has three characteristics: dynamic, organized, and evolutionary, with emphasis on growth and dynamics as he focuses on growth and dynamics as the endogenous causes of ecology [16]. Some scholars have borrowed the traditional ecological chain characteristics and applied them to digital publishing as the characteristics of digital publishing ecological chain. Using business ecosystems in describing the cultural industry, Liu argues that there are adaptive, evolutionary, networked, and self-organizing characteristics of business ecosystems [17]. In her study of new media ecology, Zeng gang argues that the characteristics of new media ecology are digital, interactive, hypertextual, networked, and virtual reality [18].

Through the previous research, we can know that the existing information fusion algorithms are not very effective in real-world application of the game. All algorithms have their advantages and disadvantages. We need an algorithm that integrates the advantages of all algorithms. These characteristics, although somewhat related, are not characteristics of the digital publishing ecosystem. However, most enterprises only pay attention to the research of data mining technology and are not yet aware of the importance of data preparation for data mining.

3. Digital Industry Chain Big Data Information Fusion Algorithm Analysis

3.1. Big Data Information Fusion Algorithm Design. A probabilistic graphical model is a general term for a model that expresses probabilistic dependencies between variables using a graph structure. In this way, the joint probability distribution of all variables in the graph can be decomposed into the product of a set of factors, each of which depends on

only a connected subset of the random variables. Bayesian networks are topologically ordered, i.e., the probability distribution of a variable node depends only on the values taken by its immediate parent node, independent of other ancestor nodes:

$$x_i | x_{pre(i)+pa(i)} | x_{pa(i)}, \quad (1)$$

where $pa(i)$ denotes the set of parent nodes of x_i and $pre(i)$ denotes the set of ancestor nodes of x_i .

Neither traditional feedforward neural networks nor convolutional neural networks can handle temporal information in data well, and modelling long-distance semantic dependencies between words are crucial to understanding text content in text analysis tasks. Also, traditional neural networks are only applicable to data of fixed length and cannot handle data of variable length. The recurrent neural network (RNN) solves the above problems by introducing recurrent structure and is widely used in natural language processing and video processing. The structure of RNN is shown in Figure 1.

The typical feature of RNN structure is cyclic connectivity in adjacent moments, which enables the RNN to update the current state based on past states and current input data. Moreover, the weight matrix of RNN is shared at all moments, which is one of the outstanding advantages over feedforward networks. However, due to the gradient disappearance and gradient explosion problems, RNNs cannot capture the relevant information when the distance between input data is large. To deal with the long-term dependence of sequence data, an improved long short-term memory (LSTM) is proposed to deal with the long-term dependence of sequence data, and the mathematical expression is

$$\begin{aligned} f_t &= \sigma(W_{fh}h_{t-1} + W_{fh}h_t x_t - b_f), \\ i_f &= \sigma(W_{fh}h_{t-1} - W_{fh}h_t x_t - b_i), \\ c_t &= \tanh(W_{ch}h_{t-1} + W_{cx}h_t x_t + b_c), \\ c_i &= f_i c_{i-1} - i_i c_t, \\ o_t &= \sigma(W_{oh}h_{t-1} + W_{ox}h_t x_t - b_o), \\ h_f &= o_t \times \tanh(c_t), \end{aligned} \quad (2)$$

where f_t denotes the memory state of the LSTM, i_f denotes the hidden state of the LSTM output, and W_{fh} denotes the new information in the input. When updating the memory state, the input gate h_t determines how much new information can be stored in the memory state, the forgetting gate b_f determines how much of the previous moment's memory state is discarded, and the output gate o_t can determine the output based on the memory state. The LSTM solves or mitigates the gradient. The LSTM solves or mitigates the problem of gradient disappearance and gradient explosion, enables long-range dependent memory, and adaptive addition and deletion of information through a gating mechanism. The structure of the LSTM can be easily extended to handle other forms of data, such as trees, graphs, and multidimensional data, and can be combined with convolutional neural networks [19].

To better model the topic-word association, the model splices the topic vector ec with the word representation H_1 and uses a new RNN to process the spliced word vector to obtain a topic-sensitive high-level word representation $H_2 = RNNs(H_1, ec)$, and H_2 highlights the topic-related features of the words. With H_2 as input, the model uses the attention module AT_{asp} to obtain the topic representation of the sentence v_a^h :

$$v_a^h, \beta_a = AT_{asp}(H_2, e_c). \quad (3)$$

The expression for the calculation of AT_{asp} is

$$\begin{aligned} M_a &= \tanh(W_h H_2 - W_c e_c \otimes L), \\ \alpha_a &= \text{soft max}(w_a^T M_a), \\ v_a^h &= H_2 \alpha_a, \end{aligned} \quad (4)$$

where W_h , M_a , and wa are model parameters and e_c denotes repeating the column vector ce horizontally L times. The high-level topic representation vector v_a^h is a weighted summation of all positional word representations using the attentional importance score as a weight. Since H_2 may discard the semantic information of the words themselves in the process of fusing topics, the model uses to integrate the information of H_1 to obtain the low-level topic representation $v_a^h = H_2$. Similarly, the attention module T_{asp} aa is used to obtain the sentiment representation of the sentence M_a and W_c . The model also introduces a shared attention model T_{asp} to generate sentence representations vs common to both topic identification and sentiment analysis tasks, modelling the semantic association between topic and sentiment.

Most algorithms consider only one aspect of meta-information and ignore other factors, which can easily result in the loss of useful information. Second, the model does not develop different fusion strategies when incorporating author and citation information according to the different ways of their influence on the distribution of paper topics and cannot effectively use the knowledge of different metadata. In general, an author is involved in a relatively wide range of research areas, and a document will only contain a limited number of topics. Thus, the topic distribution of a paper is partially correlated with the interest preferences of a single author. In other words, two documents with a citation relationship may have very similar topics or only a small part of their content may be related. Therefore, the general algorithm's strategy of driving the subject distribution of documents linked together as close as possible is not reasonable.

First, there is a clear difference in nature between the distribution of paper topics and the distribution of authors' topic interests. According to common sense, an author may be involved in several academic fields at the same time and be interested in different topics under the same field, making the corresponding topic interest distribution more dispersed with many elements taking large values. In contrast, a paper tends to address only a few closely related topics, with a more concentrated topic distribution. Thus, the distribution characteristics of the topics of a paper and a single author are

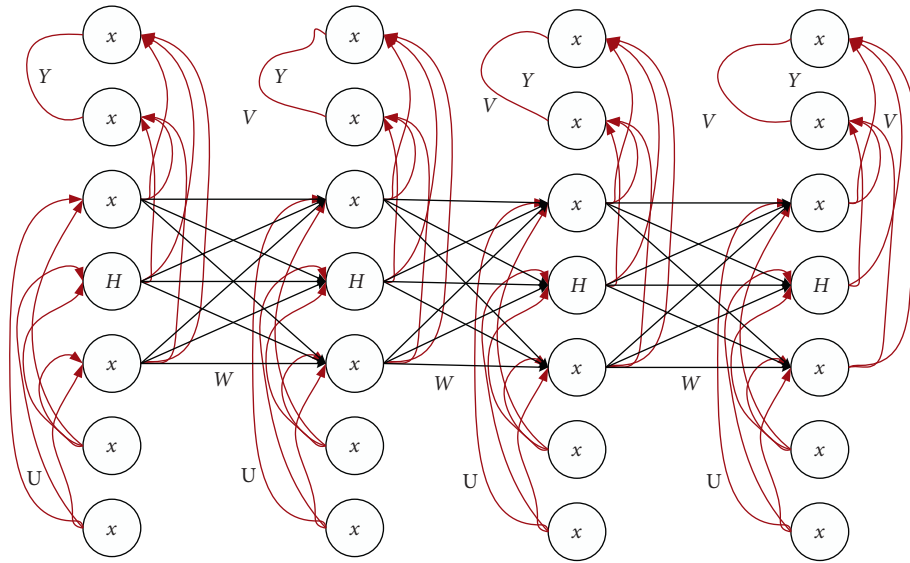


FIGURE 1: Schematic diagram of RNN structure.

inconsistent, which can adversely affect the results of the model if they are directly equated by summation without considering the essential differences between them, as shown in Table 1.

Second, in many cases, the distribution of topics in papers where citation relationships exist is not necessarily exactly similar. It is observed that a paper will cite a large number of references, but different references contribute differently to the content of the paper: some introduce background knowledge, some provide the necessary theoretical and technical foundations, some are historical literature on the same topic that is closely related to the paper, and some even cite just to ensure the completeness of the paper and have no significant relationship with the topic of the paper [20]. As an example, in the LDA paper, the authors propose a new probabilistic graph model and estimate the parameters of the model using a variational inference approach, and finally, they apply the model to text classification and collaborative filtering tasks. When other papers cite the LDA paper, they may be interested in the probabilistic graph structure of LDA, or they may need to refer to the variational inference method in the paper, or they may even just use it as a comparison algorithm for text classification experiments. Due to this complexity of the reasons for citation formation, although the topics of papers on the same edge in the citation network are related to some extent, the topic distributions of the two do not necessarily match. If the similarity of the topic distributions of connected papers is directly maximized without discrimination, it may mislead the model.

3.2. Experimental Design of Information Fusion in the Digital Publishing Industry Chain. The purpose of studying digital publishing ecological chain is to build a healthy and sustainable digital publishing ecological chain, and this requires not only understanding the basic theory of digital publishing ecological chain but also following certain principles and requirements to make the constructed digital publishing

ecological chain scientific, reasonable, and in line with the development of digital publishing enterprises.

This requires improving the digital publishing environment faced by digital publishing enterprises, adjusting the activities of digital publishing enterprises, promoting the formation and development of the digital publishing ecological chain, and ultimately realizing the construction of the digital publishing ecological chain. This section focuses on the principles and requirements for the construction of digital publishing ecological chain, the ways, and paths of the construction, and the key measures for the construction, to provide a reference for the construction of digital publishing ecological chain.

According to the characteristics of the template, the spatial filtering can be divided into two types: linear and nonlinear. Linear spatial filtering is often based on Fourier analysis, while nonlinear spatial filtering usually operates directly on the neighbourhood. According to the function of the spatial filter, the spatial filter can be divided into two types: a smoothing filter and a sharpening filter. The smoothing filter can be implemented with a low-pass filter, and the purpose is to blur the image (extract larger objects in the image to eliminate small objects or connect small discontinuities of objects) or eliminate image noise; the sharpening filter is implemented with high-pass filtering, and the purpose is to emphasize the details of the image being blurred.

The construction of digital publishing ecological chain is not a blind act but has its construction principles. In the construction of digital publishing ecological chain, the construction principles are not only the constraints that must be observed to improve the ecological efficiency of digital publishing and promote the virtuous cycle of digital publishing development but even include some other constraints that need to be observed such as improving the environment and relationships. The improvement of the general environment mainly relies on the government's macrocontrol and active foreign policy, as well as the various

TABLE 1: Variables and their meanings.

Variables	Meaning	Value
D	Number of papers	2
A	Number of authors	3
K	Number of topics	4
N	Number of words in paper d	5
C	Citation indicator variable	6

D, A, K, N, and C indicate that the parameter symbols are D, A, K, N, and C, respectively.

support roles provided by the government or industry authorities for the industry development. For example, providing digital publishing development bases, creating digital publishing industrial parks, encouraging digital publishing technology innovation, reducing administrative licenses for digital publishing, and improving comprehensive digital publishing services can improve the general environment of digital publishing. Of course, the most worthy of improvement is still the specific environment, such as the development environment of e-books and the development environment of knowledge bases, because the improvement of the specific digital publishing environment has a direct driving effect on the development of this type of digital publishing, so it is also the most effective, as shown in Figure 2.

Private Student Loans is information on students' loan behaviour through national policies. Teacher Loan Forgiveness Program is information on the loan behaviour of teachers through national policies. The data fusion described in this paper is the fusion of multiple sources of data, such as databases and knowledge bases from publishers and libraries, which can make full use of the complementary nature of multiple sources of data and the high-speed computing and intelligence of electronic computers to improve the quality of the resulting information [21, 22]. The key fusion object is the digital resource entity, which focuses on solving the problem of cross-system access to data and pays less attention to the content of data. We use a fusion algorithm that integrates multiple algorithms. We use 80% of the data as the training set for algorithm training and 20% as the test set for algorithm verification.

Information fusion is the process of obtaining relevant information from multiple information sources, such as sensors, databases, knowledge bases, and humans themselves, and filtering, correlating, and integrating them to form a representation architecture that is suitable for obtaining relevant decisions; it involves multivariate decision problems, i.e., the process of completing a given fusion decision task according to the decision task and the multiple information resources available to it, which can be accomplished through one. This process can be accomplished in one or more fusion processes. The information fusion described in this paper is the fusion of multiple sources of information, such as databases and knowledge bases from publishers and libraries, which can be used to derive more effective information and improve the effectiveness of the whole system by optimizing the combination of information. Specifically, the integration of information services

between publishers and libraries in the digital era refers to an integrated information service environment in which users can access digital resources provided by collaborating publishers and libraries through a single interface and from multiple perspectives without feeling the process of switching between resources and services. It relies on the concept of information organization, using information links, information portals, and other customary methods to describe and link digital resources of different nature, sources, and formats with a single standard, so that independent resource entities can form associations; its core integration object is the relationship between digital resource entities, focusing on revealing, linking, and linking the relationship between information, to implement users rely on a single portal to achieve their information service needs. The core object of integration is the relationship between digital resource entities, which focuses on revealing, linking, and linking information, to implement users' reliance on a single portal to achieve the satisfaction of their information service needs.

4. Analysis of Results

4.1. Analysis of the Performance Results of the Fusion Algorithm. In this article, all the data are collected by us, the total number of data is 80,000, of which 64,000 days are used for training the model in the training set, and 16,000 are used for the test mode.

We conducted experiments in two areas: a sentiment classification task for a given topic and joint topic detection and sentiment classification task, denoted as Task 1 and Task 2, respectively. Given a sentence, the sentiment classification task for a given topic predicts the sentiment expressed by the sentence on a given topic, while the joint topic detection and sentiment classification task requires the model to determine all occurrences of tuples talked about in the sentence without additional information, and the tuple is judged to be correctly classified only when both components of the tuple are correctly predicted. The classification accuracy and score were used as indicators to evaluate the classification quality of the model. Since sentiment classification is a triple classification problem, the experiments used macro-F1 as a measure of overall classification effectiveness, and the results are shown in Figure 3.

To combine specific topic information, the model first stitches the topic representation vectors to the word vectors of all words in the sentence and inputs them to the bidirectional LSTM to generate the hidden representations of the words. The model is based on a convolutional neural network and gating mechanism, using different convolutional filters to compute n-gram syntactic features with different granularity, and gating Tanh-ReLU unit to control the flow of sentiment information to the pooling layer, thus removing the sentiment features that are irrelevant to the topic or unimportant to the whole sentence. Since both the convolutional operation and the gating unit can be computed in parallel, the model is trained more efficiently.

Figure 4 shows the classification accuracy and macro-F1 of different algorithms on Task 1 with bold and add

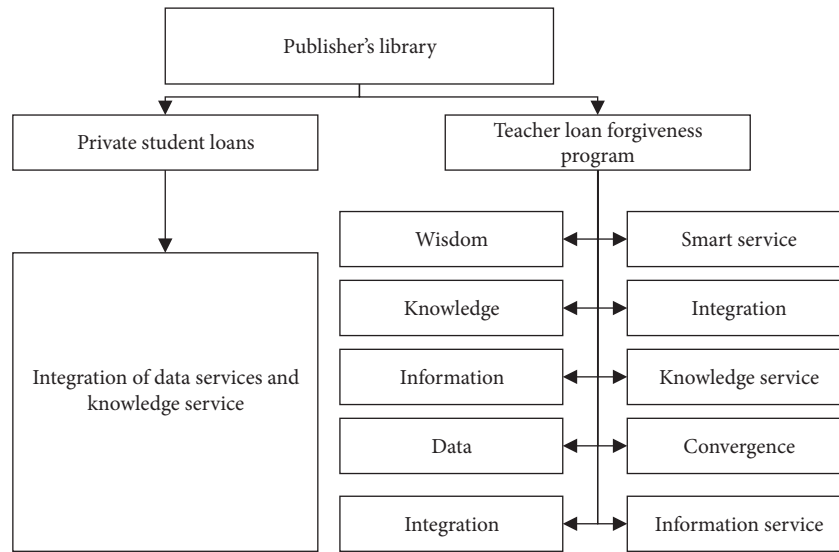


FIGURE 2: Hierarchy of information integration of publishing organizations in the digital age.

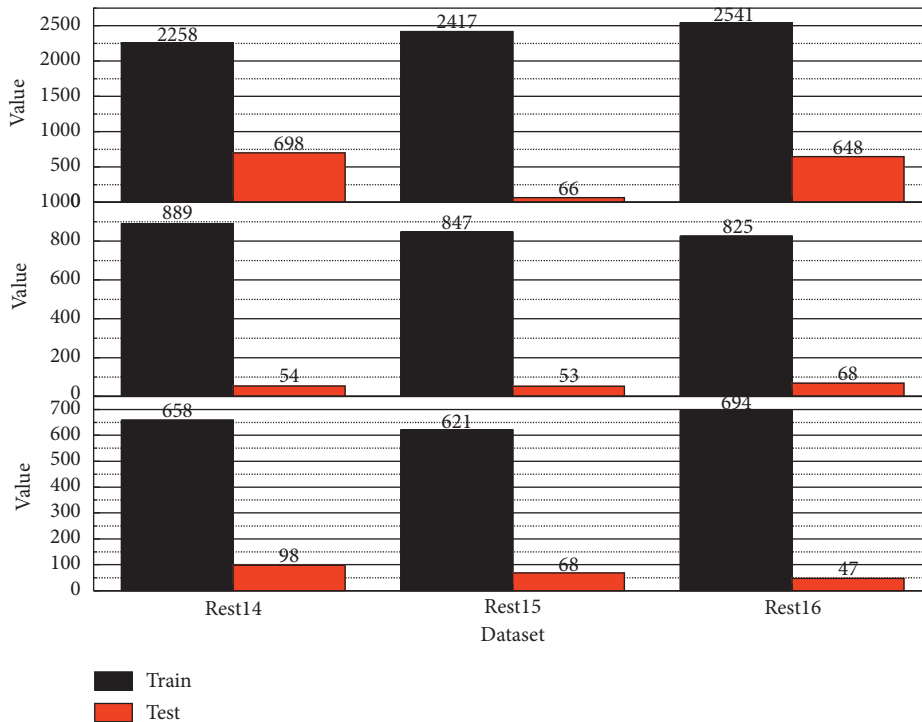


FIGURE 3: Statistical information of the data set.

underline operations for the best experimental results and the suboptimal results under each evaluation metric, respectively. As can be seen from the table, the SSFTM model proposed in this paper achieves the best classification results on all three datasets, and the accuracy is improved by 0.92%, 0.62%, and 0.66%, and the F1 scores are improved by 0.65%, 0.51%, and 0.56%, respectively, compared with the suboptimal results.

Meanwhile, the performance of GCAE based on convolutional neural networks is significantly lower than that of other models based on recurrent neural networks,

such as CAN and AS-Capsules, indicating that despite higher computational efficiency, convolutional neural networks cannot have captured long-distance dependencies between word sequences and discourse order features, which are not conducive to extracting the thematic and sentiment information embedded in sentences. Moreover, we observed that despite the same use of attentional mechanisms to construct topic-related sentence representations.

To make full use of the attention mechanism to extract topic-related attribute features and sentiment features in

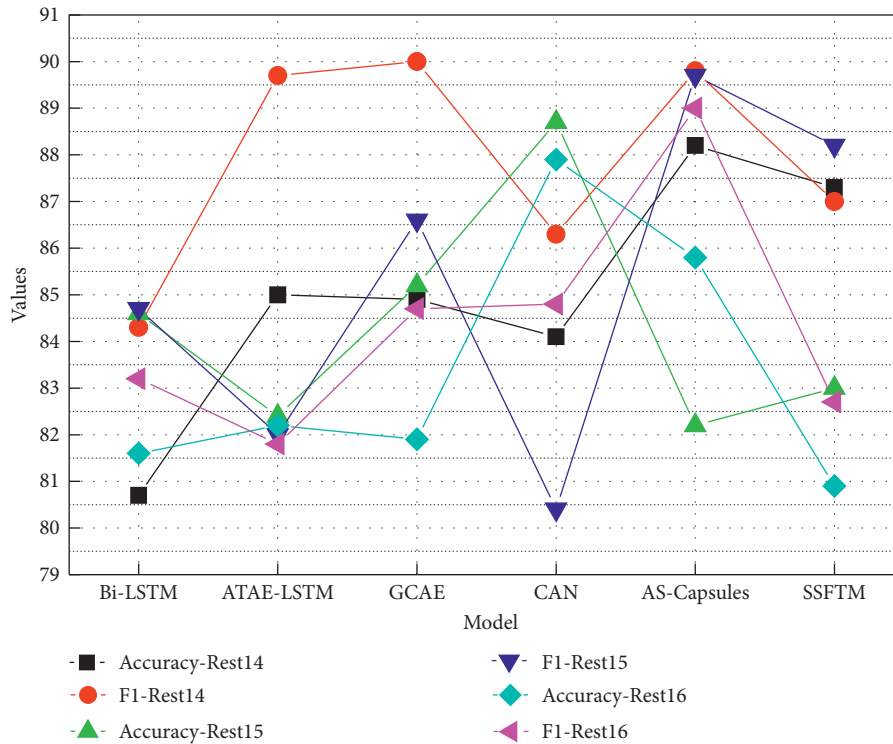


FIGURE 4: Classification results of different algorithms on Task 1.

sentences and, at the same time, improve the model’s adaptability to complex sentences containing multiple topics or negative structures, the SSFTM model uses different modules to process sentences sequentially in a multilevel and all-round way, with a clear division of labor and close cooperation among the modules. To verify the effectiveness of the above-proposed improvement methods on the topic sentiment classification task, a set of comparison experiments were designed to remove one of the improvement methods separately and explore the degree of contribution of each improvement method to improve the classification results.

Figure 5 shows the classification results of the different comparison models on the sentiment classification task for a given topic. The performance of all the comparison models undergoes some degree of degradation compared to SSFTM. Among them, the performance of the SSFTM-M model, which eliminates the hierarchical attention network, undergoes a significant degradation, proving that the pure reliance on RNNs cannot adequately model the semantic association between disjoint words and that the single-layer attention network is limited in its ability to extract attribute words and sentiment words, thus demonstrating the effectiveness of the hierarchical attention structure in mining indirect dependencies between words. Meanwhile, the performance of SSFTM-C is also significantly weaker than the other models, indicating that attribute words and sentiment words have certain correlations and complementarities, and making full use of these correlations helps the model to accurately capture insignificant attribute and sentiment information in sentences. The classification

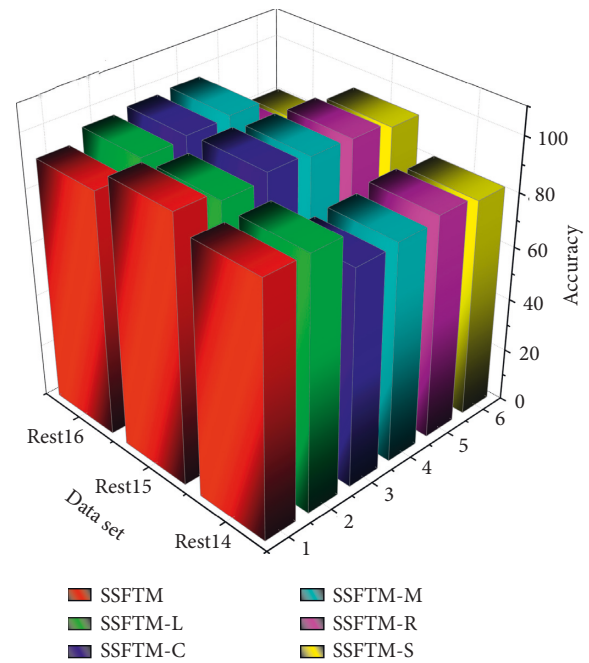


FIGURE 5: Classification accuracy of the contrast model on the sentiment classification task for a given topic.

accuracy of the SSFTM-S model on both datasets occurs in equal magnitude, indicating that sparse regular terms in most cases are can effectively improve the attention mechanism and exclude the interference of irrelevant noise information.

4.2. Analysis of Information Fusion Results of Digital Industry Chain. The sample data taken for the population analysis based on e-book reviews are still analysed with the data of travel e-books in Jingdong, and the analysis is carried out specifically on the time when the works were released (in years) and the situation of the works being reviewed. In the actual operation process, excluding some works with no marking of time, the remaining 975 e-books are organized as shown in Figure 6. Among them, the time of works is the generation time of e-books; the number of reviewed works indicates the number of reviewed works; the number of works by year indicates the number of all works in a certain publication year, and the number of works from 2011 to 2020 indicates the total number of works published in that year that were reviewed from 2011 to 2020. This is because the generation of reviews may occur each year, and may not occur each year, or may occur consecutively. Therefore, the calculation of the number of reviewed works is taken as the mean value. Of course, the analysis of the life cycle of the digital publishing industry is also very necessary. Only by properly understanding, the development dynamics of the digital publishing industry can effective policies be formulated promptly. From the national level, mastering the development status of the digital publishing industry can provide a theoretical basis for the national development planning of the digital publishing industry and the introduction of supporting policies and provide a good external environment for the development of the digital publishing industry.

From the perspective of enterprises, understanding the development status of the whole industry enables digital publishing enterprises to adjust their development strategies and formulate their long-term plans promptly. Finally, studying the life cycle of the digital publishing industry can also provide a basis for the life cycle management and evaluation of digital publishing enterprises. The purpose of this paper is to investigate whether the life cycle of digital publications matches the development cycle of the digital publishing industry and whether the current eco-efficiency matches the development trend of the digital publishing industry, as shown in Figure 7.

As can be seen from Figure 8, the weights of each primary index for evaluating the performance of integration of knowledge services between publishing organizations and libraries in the digital era, such as integration cost, integration quality, integration effect, and integration sustainability, as well as the weights of the secondary indexes under each primary index, are basically in line with the current actual situation. The establishment of the index system for evaluating the performance of knowledge service integration between publishing organizations and libraries in the digital era and the determination of the weights of each index in this paper have positive significance for the evaluation of the performance of knowledge service integration between publishing organizations and libraries in the real situation, aiming to provide a certain reference for evaluating the performance of knowledge service integration between publishing organizations and libraries in the digital era and promote, as shown in Figure 8. The purpose is to provide



FIGURE 6: Sample data of e-book review population analysis.

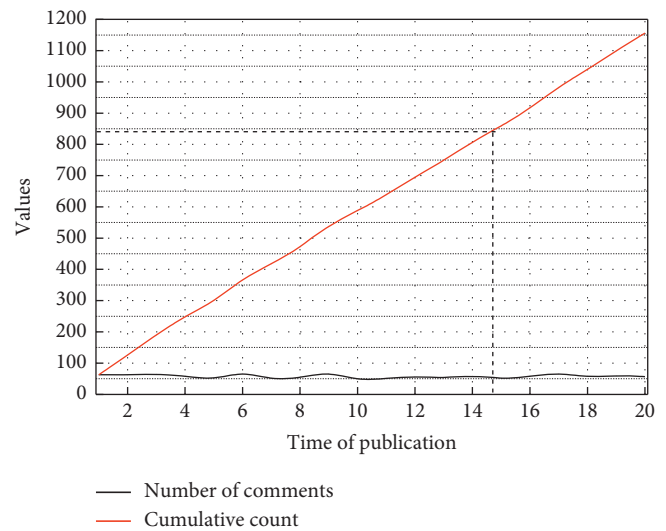


FIGURE 7: Half-life of e-books.

some reference for evaluating the performance of the integration of knowledge services between publishers and libraries in the digital era and to promote the integration of knowledge services between publishers and libraries in the digital era.

To analyse the effect comprehensively, it is necessary to include both risk and performance in the study of the effect of knowledge service integration between publishers and libraries in the digital era. This section first analyses the composition of knowledge service integration effects, i.e., including knowledge service integration risks and knowledge service integration performance. Then, the importance of analysing the risk of knowledge service integration and evaluating the performance of knowledge service integration is pointed out. Then, the game relationship between

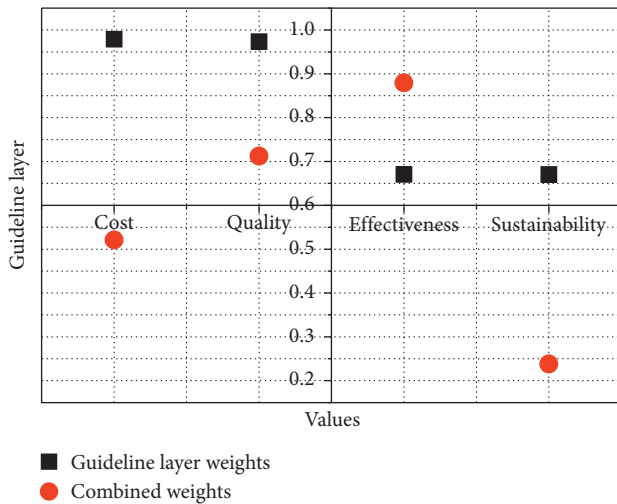


FIGURE 8: Integration performance evaluation index weights.

publishers and libraries in the digital era is analysed using the mathematical method of game theory, and it is learned that the knowledge service integration between publishers and libraries in the digital era is a non-zero-sum cooperative game, and the main determinants of the game relationship are analysed. These include the external risks of policy constraints, potential competition, and unexpected situations, and the internal risks of trust, adverse selection, spill over effects, and benefit distribution.

5. Conclusion

This paper clarifies the concept of digital publishing ecological chain and the types of participating subjects, solving the previous controversial and confusing definitions, and at the same time, it divides the structure of digital publishing ecological chain into basic and derivative structures, distinguishing digital publishing activities from nondigital publishing interactions. At the same time, this paper also analyses the flow of resources in digital publishing, studies the eco-efficiency of each subject in the digital publishing ecosystem in digital publishing activities, and discovers hidden problems, which not only provides new research ideas and research methods for subsequent studies but also reveals the problems of digital publishing activities. Finally, this paper analyzes the various kinds of imbalances involved in digital publishing, analyses the specific reasons for their generation, and puts forward constructive opinions on the regulation and optimization measures of the digital publishing ecological chain, which provides a good solution to the current ecological problems. Moreover, considering that the local topic information in sentences helps guide the attention mechanism to get better attention distribution, the model defines a local topic-aware module to extract the topic information unique to each sentence and dynamically update the global attribute vector and sentiment vector adaptively. Meanwhile, to exploit the semantic association of attribute words and sentiment words under the same topic and synergistically enhance the model's ability to extract

topic-related and sentiment-related information, the model uses a tensor neural network to portray the correlation between the two from multiple perspectives and achieves the propagation and interaction of the two complementary information through a coupled multilayer attention network to capture the long-range syntactic dependencies between words. The fusion algorithm I studied has high accuracy and efficiency and combines the advantages of multiple algorithms. This algorithm can be used in the actual word publishing industry chain, and it has a strong role and significance. The model also incorporates linguistic knowledge to alleviate the problem of affective drift caused by negation structure.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that they have no conflicts of interest regarding the publication of this paper.

References

- [1] P. Grover and A. K. Kar, "Big data analytics: a review on theoretical contributions and tools used in literature," *Global Journal of Flexible Systems Management*, vol. 18, no. 3, pp. 203–229, 2017.
- [2] S. Khanra, A. Dhir, and M. Mäntymäki, "Big data analytics and enterprises: a bibliometric synthesis of the literature," *Enterprise Information Systems*, vol. 14, no. 6, pp. 737–768, 2020.
- [3] E. Nica, K. Janošková, and M. Kovacova, "Smart connected sensors, industrial big data, and real-time process monitoring in cyber-physical system-based manufacturing," *Journal of Self-Governance and Management Economics*, vol. 8, no. 4, pp. 29–38, 2020.
- [4] Z. H. Munim, M. Dushenko, V. J. Jimenez, M. H. Shakil, and M. Imset, "Big data and artificial intelligence in the maritime industry: a bibliometric review and future research directions," *Maritime Policy & Management*, vol. 47, no. 5, pp. 577–597, 2020.
- [5] K. Nisar, "Smart home: multisensor information fusion towards better healthcare," *Advanced Science Letters*, vol. 24, no. 3, pp. 1896–1901, 2018.
- [6] R. Davidson, "Cyber-physical production networks, artificial intelligence-based decision-making algorithms, and big data-driven innovation in industry 4.0-based manufacturing systems," *Economics, Management, and Financial Markets*, vol. 15, no. 3, pp. 16–22, 2020.
- [7] F. Song, M. Zhu, Y. Zhou et al., "Smart collaborative tracking for ubiquitous power IoT in edge-cloud interplay domain," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6046–6055, 2019.
- [8] D. Yang, L. Wu, S. Wang, H. Jia, and K. X. Li, "How big data enriches maritime research - a critical review of Automatic Identification System (AIS) data applications," *Transport Reviews*, vol. 39, no. 6, pp. 755–773, 2019.
- [9] O. Throne and G. Lăzăroi, "Internet of things-enabled sustainability, industrial big data analytics, and deep learning-assisted smart process planning in cyber-physical

- manufacturing systems,” *Economics, Management and Financial Markets*, vol. 15, no. 4, pp. 49–58, 2020.
- [10] X. Zhang and J. S. L. Lam, “A fuzzy Delphi-AHP-TOPSIS framework to identify barriers in big data analytics adoption: case of maritime organizations,” *Maritime Policy and Management*, vol. 46, no. 7, pp. 781–801, 2019.
- [11] W. I. Yudhistyra, E. M. Risal, I.-s. Raungratanaamporn, and V. Ratanavaraha, “Exploring big data research: a review of published articles from 2010 to 2018 related to logistics and supply chains,” *Operations and Supply Chain Management: An International Journal*, vol. 13, no. 2, pp. 134–149, 2020.
- [12] G. Ji, L. Hu, and K. H. Tan, “A study on decision-making of food supply chain based on big data,” *Journal of Systems Science and Systems Engineering*, vol. 26, no. 2, pp. 183–198, 2017.
- [13] H. G. Van der Voort, A. J. Klievink, M. Arnaboldi, and A. J. Meijer, “Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making?” *Government Information Quarterly*, vol. 36, no. 1, pp. 27–38, 2019.
- [14] C. Liu, H. Li, Y. Tang, D. Lin, and J. Liu, “Next generation integrated smart manufacturing based on big data analytics, reinforced learning, and optimal routes planning methods,” *International Journal of Computer Integrated Manufacturing*, vol. 32, no. 9, pp. 820–831, 2019.
- [15] A. Mohamed, M. K. Najafabadi, Y. B. Wah, E. A. K. Zaman, and R. Maskat, “The state of the art and taxonomy of big data analytics: view from new big data framework,” *Artificial Intelligence Review*, vol. 53, no. 2, pp. 989–1037, 2020.
- [16] X. Zhao, Y. Jia, A. Li, R. Jiang, and Y. Song, “Multi-source knowledge fusion: a survey,” *World Wide Web*, vol. 23, no. 4, pp. 2567–2592, 2020.
- [17] J. Liu, C. Wu, G. Wu, and X. Wang, “A novel differential search algorithm and applications for structure design,” *Applied Mathematics and Computation*, vol. 268, pp. 246–269, 2015.
- [18] X. Zenggang, T. Zhiwen, C. Xiaowen et al., “Research on image retrieval algorithm based on combination of color and shape features,” *Journal of Signal Processing Systems*, pp. 1–8, 2019.
- [19] L. Poma, H. A. Shawwa, and E. Maini, “Industry 4.0 and big data: role of government in the advancement of enterprises in Italy and UAE,” *International Journal of Business Performance Management*, vol. 21, no. 3, pp. 261–289, 2020.
- [20] Y. Liu, “Incomplete big data imputation mining algorithm based on BP neural network,” *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 4, pp. 4457–4466, 2019.
- [21] Y. Xiao, C. Li, L. Song, J. Yang, and J. Su, “A multidimensional information fusion-based matching decision method for manufacturing service resource,” *IEEE Access*, vol. 9, pp. 39839–39851, 2021.
- [22] J. Yang, J. Wen, B. Jiang et al., “Multi-sensor information fusion in ocean of things based on improved adaptive dempster-shafer evidence theory,” in *Proceedings of the 2019 IEEE International Conference on Industrial Internet (ICII)*, pp. 399–404, IEEE, Orlando, FL, USA, November 2019.

Copyright © 2021 Haixiang He. This is an open access article distributed under the Creative Commons Attribution License (the “License”), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License. <https://creativecommons.org/licenses/by/4.0/>