

Research Article

Automatic Production Technology of Data News Based on Machine Learning Model

Zheng Li 

Xijing University, Xi'an, Shaanxi 710123, China

Correspondence should be addressed to Zheng Li; lizheng@xijing.edu.cn

Received 3 January 2022; Revised 19 January 2022; Accepted 20 January 2022; Published 11 February 2022

Academic Editor: Xin Ning

Copyright © 2022 Zheng Li. 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.

Data news does not appear out of nowhere. It is the result of the press's ongoing exploration and innovation as part of the ongoing development of information technology. It is based on early computer-assisted news reporting, statistical news and precise news. In terms of theory and technology, it is a game-changer. The media industry is most sensitive to new technologies and is most willing to apply new technologies in the entire process of news reporting, such as information acquisition, content production and management, and content release, in an era of great changes in the popularity of the Internet and the in-depth application of data processing technology in all walks of life. The data news intelligent production platform investigated and designed in this paper, when combined with current emerging technologies, such as big data, cloud computing, and machine learning algorithms, realizes intelligent and automatic news production, ushering traditional news into an era of intelligence, knowledge, and digitization. Machine learning has made significant progress in the automatic production of data news technology and will continue to do so in the future information society.

1. Introduction

New technologies such as information technology, automation technology, and artificial intelligence [1] have been widely applied to various industries in recent years, providing favorable technical conditions for the development of various industries. The media industry is most sensitive to new technologies and most willing to apply new technologies in the entire process of news reporting, such as information acquisition, content production and management, and content release, in this era of great change due to the popularity of the Internet and the deep application of data processing technology in all walks of life [2]. News does not “appear out of nowhere,” but rather is the result of the press's constant exploration and innovation as part of the ongoing development of information technology. It is based on early computer-aided news reporting, statistical news, and accurate news, and it represents a breakthrough and innovation in theory and technology when compared to previous ones [3]. At the moment, major media outlets in the United States and abroad are scrambling to set up data

news production channels or columns. The methods of production are more traditional. The depth and timeliness of data collection and visual presentation, manuscript editing, and final news reporting are not good, and it takes a lot of manpower [4]. The introduction of automatic production technology in data news can increase data news volume efficiency, improve overall production quality, lower labor costs, and promote the development of automatic data news production technology toward full automation [5].

The data news intelligent production platform studied and designed in this paper combines the current emerging technologies, such as big data technology, cloud computing, and machine learning algorithm [6] realizes the intelligent and automatic production of news, and brings the traditional news to an era of intelligence, knowledge, and digitization. Machine learning, also known as statistical machine learning, is an important part of AI field. Its basic idea is to use data to build a model and then use this model to analyze and predict data. Due to the limitation of computer technology, it is difficult for people to effectively process huge data. Now, due to the rapid development of computer

technology, it is possible for people to process and analyze such huge data [7]. The essence of machine learning is a technology that enables computers to have learning ability like people and find useful knowledge from a large amount of data. With regard to the automatic production technology of data news, we pay more attention to the efficiency and correct classification of the automatic production technology of data news. However, in text classification, due to the particularity of Chinese text, text word segmentation is complex and requires a lot of lengthy time [8]. This paper mainly studies how to improve the efficiency of text word segmentation and feature dimensionality reduction, uses machine learning algorithms to fit the effect of model test, and compares the fitting effects of several machine learning algorithms on this data set [9, 10]. In view of the development of data news, new technologies are introduced into the practice of automatic production technology of data news [11, 12]. Taking data news as the object, the relevant knowledge Atlas and its knowledge base are constructed. Using machine manuscript technology, data visualization, digital animation, and intelligent speech synthesis technology, a set of technologies that can realize resource integration, news production practice, business innovation, and intelligent production release-integrated support platform with multirelease, fine management, and iterative development.

Machine learning is the cross-cutting of many disciplines of data news automatic production technology, including information control theory, logic science, mathematical statistics, neuroscience, and computational science [13]. In short, machine learning learns some inherent information in the automatic production technology of data news through computers and obtains new knowledge to improve the intelligence of the automatic production technology of data news of computers, which can make computers think and judge like people. Machine learning has spawned many research branches in the automatic production technology of data news, such as speech recognition, data mining, pattern recognition, intelligent control of robots, and remote sensing information security. To sum up, machine learning has made great progress in the automatic production technology of data news and will make more remarkable achievements in the future information society [14, 15].

2. Related Work

Nam et al. [16] pointed out that at present, automated news production has been widely used and developed in Chinese traditional media and new media. The news writing robot launched by Tencent published China's first automated news, CPI rose 2% year-on-year in August, a 12-month high, officially setting off a wave of domestic automated news production. Literature [17] through the big data analysis method, the changes brought by technological innovation to data news have never stopped. In the environment of big data, with the continuous development of artificial intelligence technology, the mode of news production has changed, the news production process has been reshaped, and the news communication format has undergone

changes. Nam et al. [16] show that foreign automated news writing software mainly carries out programmed reporting based on data. However, for the news editing room, as the news editing room lacks the technology to develop automated writing software, automated news mainly carries out automated news production cooperation with technology companies through news media companies abroad. Utomo and Leu [17] proposed that in China, CCTV cooperated with Baidu to launch the digital news report of according to "Spring Festival transportation" and then launched similar programs. Other portals have also launched data news channels, such as Sohu's "digital way," Sina's "graphic news," Caixin's "digital theory," and Xinhua's "data news." Literature [18] through the big data analysis method, the "robot" to realize automatic news production is actually an automatic manuscript generation software. Through data integration, and based on the data read by the computer, according to the calculation rules provided by relevant industry experts, the judgment rules automatically form news manuscripts in the news template. Khader et al. [19] show that in China, *Tencent finance* and economics channel first released a report with the automatic news writing robot Dreamwriter. Then, *Xinhua news agency*, *First Finance and Economics*, and *Today's Headlines* launched their own writing robots to realize the automatic writing of manuscripts. Kamburugamuve et al. [2] point out that *Tencent Finance*, *Xinhua News Agency*, *First Finance*, and *Today's Headline* are the leaders in the development of automated news production in China. With the development of artificial intelligence technology, machine manuscript writing technology has also received attention at home and abroad through Big Data analysis method [20]. In foreign countries, manuscript-writing robot companies including ARRIA and NARRATIVESCIENCE have written tens of thousands of news articles for many well-known media such as Associated Press. Taninaga et al. [21] show that for personalized manuscripts, through data analysis of users' reading habits and hobbies, the same information can be organized into manuscripts of different angles and styles in a targeted way, and personalized push can be carried out. Indira and Kavithadevi [22] put forward that for automatic news production, it is basically a mode of "manual template plus automatic data filling."

This paper investigates the use of a machine learning algorithm to automate the production of data news. As content creation requires human wisdom and thought, the degree of automation in news content creation is always low. At this point, an automatic writing program will never be able to think independently like a human brain, but for some specific products and under certain conditions, some manual functions in content production can be replaced by automatic means using existing technology. Functional diversification, design standardization, and application integration are gradually becoming the core development trends of data news automation technology in the modern development background of data news technology innovation. The introduction of automation technology into the data news industry can not only improve and optimize traditional technology but also effectively improve the

industry's overall operation level and encourage data news to adapt to the new economic normality.

3. Machine Learning Algorithm

3.1. Principle and Model of Machine Learning Algorithm. Machine learning is an interdisciplinary subject involving statistics, convex analysis, probability theory, algorithm complexity theory, approximation theory, and so on. Machine learning algorithm, which aims at machine learning, is the core of artificial intelligence and mainly studies computer simulation or ways to realize human learning behavior. As the core of artificial intelligence technology, machine learning algorithm is the fundamental way to ensure that computers can have intelligent characteristics. The basic precondition of machine learning algorithm is algorithm training, which predicts the specific value in a certain frequency range by inputting specific data information. Through the application of machine learning algorithm, computers can be made to run intelligently, and new knowledge and skills can be continuously collected and summarized. Gradually form a new knowledge structure to further improve the computer performance. Gradually improve the previous data classification system, form a new computer data classification algorithm, and then achieve the purpose of optimizing data classification standards.

The process of constructing machine learning algorithm is also called predictive modeling of machine learning algorithm. On the premise of effectively mastering this machine learning algorithm model, people can directly analyze the original type of machine learning algorithm, apply it to the analysis of new data, and mine the valuable information. Due to many uncertain factors in information acquisition and processing, the data classification accuracy at this stage cannot meet the practical application needs of thematic information extraction and environmental monitoring. Machine learning method is a method that computer obtains some models using existing data and uses these models to predict the future. In fact, this process is very similar to human learning process, except that machines can carry out large-dimensional data analysis and are tireless, as shown in Figure 1.

In the process of prediction, it is necessary to calculate the distance between each predicted object and all data sets. The calculation of this distance is generally realized by finding Euclidean distance, which can also be Manhattan distance or Chebyshev distance. Euclidean distance can describe the true distance between two points in N -dimensional space. Given that two N -dimensional vectors $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_n)$, Euclidean distance formula of these two vectors can be expressed as

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (1)$$

The results are sorted, and k points closest to the predicted data are selected to form a decision set. In the decision set, compare the number of data points in each category and

select the category with the largest number as the final category of prediction data points.

The linearly separable training set is defined as

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}. \quad (2)$$

$x_i = R^n, y_i \in \{+1, -1\}, i = 1, 2, \dots, N$
 $x_i = R^n, y_i \in \{+1, -1\}, i = 1, 2, \dots, N$, x_i are attribute variables, and y_i is the label variable of x_i . Then the decision function for solving this data set can be equivalently transformed into the following convex quadratic programming problem.

$$w^* \cdot x + b^* = 0. \quad (3)$$

At the same time, its classification decision function can also be obtained

$$f(x) = \text{sign}(w^* \cdot x + b^*). \quad (4)$$

Next, the concept of function interval and geometric interval are introduced. For training data set T and hyperplane (w, b) , the functional interval of (w, b) with respect to single data (x_i, y_i) can be defined as

$$\hat{\gamma} = y_i (w \cdot x_i + b). \quad (5)$$

Then the function interval of hyperplane (w, b) for data set T is defined as follows:

$$\hat{\gamma} = \min_{i=1, \dots, N} \hat{\gamma}_i. \quad (6)$$

Only the function interval is not enough because according to the mathematical expression, if (w, b) is changed by a certain multiple, the decision function we are looking for will remain unchanged, while $\hat{\gamma}$ will change accordingly. At this time, it is necessary to convert the function interval into geometric interval.

$$\gamma = \frac{w}{\|w\|} \cdot x_i + \frac{b}{\|w\|}. \quad (7)$$

Then the function interval of hyperplane (w, b) for data set T is defined as follows:

$$\gamma = \min_{i=1, \dots, N} \gamma_i. \quad (8)$$

Knowing these two concepts, we can introduce the maximum separation hyperplane, which can be intuitively understood as dividing data sets into two categories according to sufficient certainty. It can be transformed into the following constrained optimization problem.

$$\begin{aligned} & \max_{w, b} \gamma \\ & \text{s.t. } y_i \left(\frac{w}{\|w\|} \cdot x_i + \frac{b}{\|w\|} \right) \geq \gamma, \quad i = 1, 2, \dots, N. \end{aligned} \quad (9)$$

The algorithm of machine learning for data classification task is called classifier, and the key evaluation index of this classifier is accuracy. In short, it is the proportion of the correct data obtained after using this classifier for data

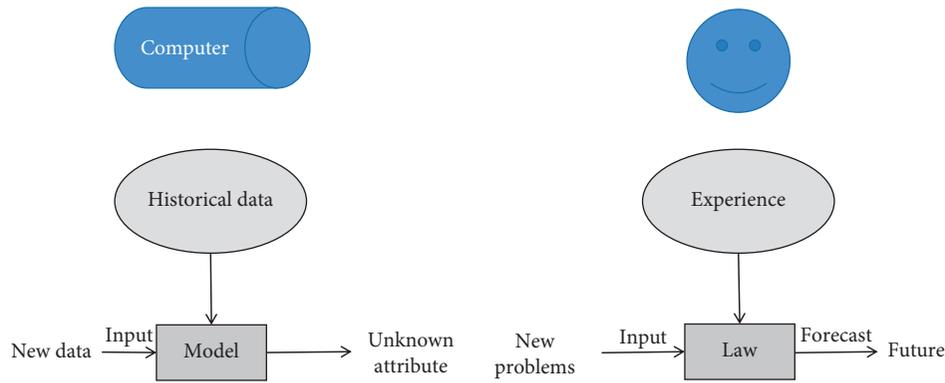


FIGURE 1: Machine learning is analogous to human learning.

classification. Machine learning algorithms include dozens of scene algorithms such as clustering, regression, classification, and text analysis. There are about 30 commonly used algorithms, and there are many deformations. Generally, people divide machine learning into four kinds: supervised learning, unsupervised learning, semisupervised learning, and reinforcement learning.

3.2. Classification Problems in Machine Learning

3.2.1. Supervising Learning Problems. Supervised learning is a type of behavior in which the machine aids the entire learning process by providing labeled data and, in the end, achieves the learning goal. The effectiveness of this machine learning method has been demonstrated after many experts, and scholars have tried it out one after another. However, labeled data are a necessary factor for learning, which directly increases the cost input of the entire machine learning process. Starting with some observation and training samples and looking for data classification rules that cannot be obtained through principle analysis, accurate prediction of future data behaviors and trends can be achieved. Logical regression helps solve the problem of limited sample data classification by balancing the relationship between the validity of data classification learning results and the interpretability of data classification models. The most obvious benefit of this type of machine learning is that the machine's original generalization ability can be fully utilized, allowing the data classification and regression problem to be effectively solved. If only one classifier predicts positively during the training period, the corresponding category is marked as the final data classification result; if more than one classifier predicts positively, the data that have undergone logical regression can be used as the standard, and the data classification result with the larger numerical value is preferred. Neural networks and logistic regression are two of the most commonly used algorithms in this learning method, and feature labeling is always at the heart of the computer training process. Supervision can assist machines in learning how to divide various foods and predict data using rules and regularity. However, as artificial intelligence technology advances, the quality of automated news will improve day by day, and it will become more mature in terms of news sensitivity, topic discovery, emotion expression, expressing opinions, and in-depth reporting.

3.2.2. Unsupervised Learning Problem. In the process of machine learning, unsupervised learning is a more advanced way. This way is to complete the whole learning process independently according to the program setting with the help of the original automatic function of the computer, and at the same time, it can improve the process of absorbing knowledge points on its own on the premise of effective use of all kinds of data. Owing this, the cost input of this kind of learning method is completely controllable, and it does not need high capital cost input and has good economy. Support vector machine (SVM) model is a new machine learning algorithm based on statistical learning theory, which can solve the optimization problem in high-dimensional feature space and find the optimal classification hyperplane, thus fundamentally solving the complex data classification problem. Under the premise of unlabeled sample data, the machine usually does not enter the training state. The most outstanding advantage of this method is that the machine can automatically learn the knowledge it has learned like human beings, ensuring that its knowledge acquisition behavior is highly targeted. Compared with the traditional artificial neural network algorithm, support vector machine uses the principle of minimizing structural risk, which can effectively solve many classification problems such as remote sensing data classification in reality. The relatively frequently used algorithms in this learning method include automatic encoder, deep execution network, and so on. This machine learning method itself can achieve good results in the process of solving the data clustering problem. Then, by using the feature of automatic classification of support vector machine in dimension feature space, various types of information are jointly applied to achieve the goal of information fusion and classification. Through the rational application of data from different sources and auxiliary data, complex remote sensing data can be effectively distinguished, and the accuracy of remote sensing data analysis can be ensured. Relying on the huge database and artificial intelligence technology, the data analysis ability of automated news is stronger, and the calculation accuracy and production efficiency are higher, which effectively solves the efficiency problem of news production.

4. Design of Automatic Production Technology of Data News

4.1. Automatic Production Technology of Data News Based on Machine Learning Algorithm. The core of machine learning algorithm is prediction. The prediction spirit brings new opportunities for the development of automated news. Machine learning algorithm analysis technology can reasonably predict the development law and trend of the research object, and data visualization technology can comprehensively and deeply predict the news, so as to enhance the scientificity and humanization of the report. The promotion and application of data news automation technology also lies in gradually replacing the traditional operation mode of previous manual operation, improving the accuracy and quality level of each processing and production link, standardizing each operation process with automatic process, and building a unified production and processing system according to the characteristics and actual needs of data news automation production technology. News production management system involves news topic selection management, production task management, and whole process monitoring; data management system; data news knowledge Atlas and knowledge base construction system; intelligent generation system of news copy; intelligent generation system of news chart material; audio and video intelligent generation; and fusion and final release subsystem. The framework of news intelligent production business system is shown in Figure 2.

Automated news production nowadays mostly uses automated data to fill in templates and reports news on a large scale based on machine learning algorithms' logic. This model is more solidified, has a less emotional color palette, and the ability to express position and make suggestions is limited. A semantic network is represented by an attribute graph model, and a knowledge map is essentially a semantic network. Edges represent various semantic relationships between entities or concepts, while nodes represent entities or concepts. This system primarily serves as an auxiliary in the automatic processing and production of data news. It can assist with industrial design, production analysis, and NC analysis, among other things. Promoting the widespread use of integrated systems in the automatic production of data news not only maximizes the use of various resources in the production and processing of data news, but it also improves the efficiency of cooperation between various work departments to a degree. "Robot news" is a common type of automated news. A machine learning algorithm is actually a piece of software that generates manuscripts automatically. It is increasingly being used in the media, which has resulted in significant manpower savings and increased media efficiency. In a machine learning algorithm, the construction of a knowledge map is divided into three parts: knowledge entity recognition, entity relationship extraction, and knowledge fusion and representation. The construction process of news knowledge map knowledge base is shown in Figure 3.

Based on news knowledge base, news copy is automatically generated through semantic interaction and intelligent reasoning among template base, fact base and rule base. Machine learning algorithm uses learning sorting method to extract live reports, and constructs news together with news copy obtained by semantic reasoning to optimize and enhance readability and vividness of news. In the whole process of data news and automatic production technology, automatic manufacturing is one of the most critical links, which is in the main position. Besides covering the automatic machining center, it also involves the application of automatic assembly technology, numerical control machine tool technology, or production and processing robots, which comprehensively improves the efficiency and quality level of data news. The fourth is the automated production technology system. The timeliness of news is a decisive yardstick to measure news and an important requirement for media to enhance competitiveness. Using the high-speed computing power of robot learning algorithm, according to the pre-established data model, news reports are automatically generated. This report-generation method reduces the workload of news editors and improves the efficiency of news production.

4.2. Experimental Results and Analysis. This section compares and analyzes the four methods of Nb, LR, RF, and SVM in ① text vectorization; ② on the basis of text vectorization, automatic production technology is integrated, and three experiments are carried out. The experimental results in eight cases are shown in Figures 4–6.

The experimental results show that NB at 20 to 30 is 0.2 and 0.5 higher than LR at 20 to 30, LR at 30 to 40 is 0.3 and 0.4 higher than RF, RF at 10 to 20 is 0.4 and 0.5 lower than SVM, and SVM is 0.6 and higher than NB at 40 to 50. The experimental results in Figure 5 show that NB at 10 to 20 is 0.3 and 0.4 higher than LR at 20 to 30, LR at 20 to 30 is 0.2 and 0.4 higher than RF, RF at 40 to 50 is 0.3 and 0.5 lower than SVM, and SVM is 0.4 and higher than NB at 60 to 70. The experimental results in Figure 6 show that NB at 30 to 40 is 0.6 and 0.7 higher than LR at 20 to 30, LR at 40 to 50 is 0.3 and 0.4 higher than RF, RF is 0.5 and 0.6 lower than SVM at 50 to 60, and SVM is 0.2 and 0.0 higher than NB at 40 to 50.

The experimental results of integrating data automatic news production technology based on machine learning algorithm are better than the experimental results of text vectorization. As data news plays an important role in automatic production technology, the weight of data news can achieve better classification results. Based on the machine learning algorithm, this paper conducted four experiments for comparison, as shown in Figures 7 to 10.

The experimental results show that the horizontal axis represents the number of news, and the vertical axis represents the data rate. The data show that the automatic production technology of data news in Experiment 1 for the first time is closer to the expert mark value than that in

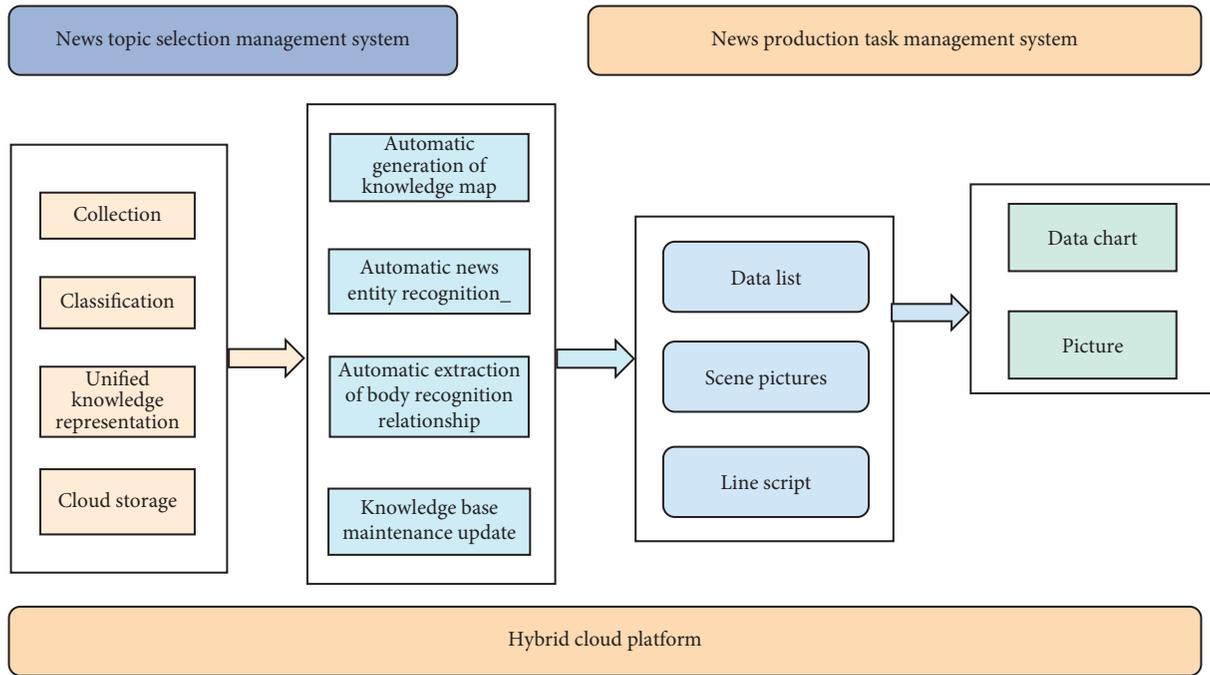


FIGURE 2: Framework diagram of business system of data news intelligent production platform.

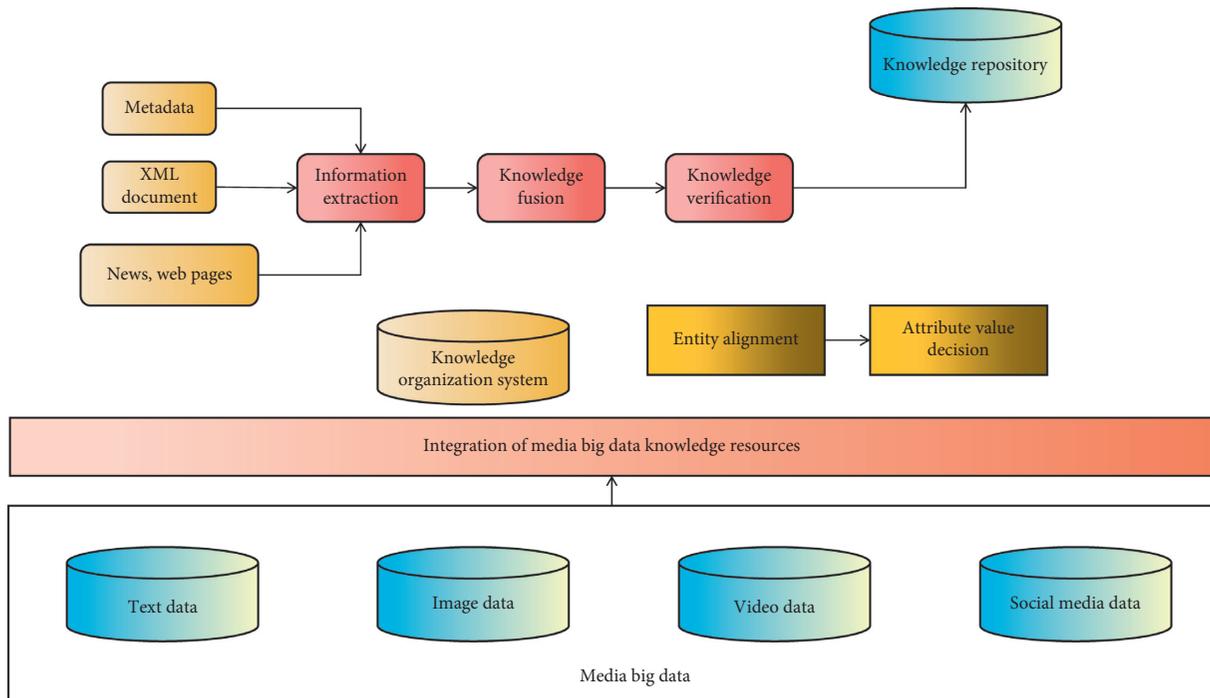


FIGURE 3: News knowledge map and knowledge base construction process.

Experiment 2. It can be seen that the Y value of Experiment 2 is smaller than that of Experiment 1, which shows that the analysis result of Experiment 2 is closer to the automatic production technology of data news marked by experts. On the contrary, it shows that integrating the characteristics of foreign exchange news into the calculation of automatic production technology of data news is more ideal than only considering the degree word. The data show that the

automatic production technology of data news in Experiment 3 is not as close to the expert mark value as that in Experiment 4. The Y value of Experiment 3 is less than that of Experiment 4, indicating that the analysis result of Experiment 4 is closer to the automatic production technology of data news as defined by experts. This situation can be drastically improved by applying flexible automated production technology to data news based on machine learning

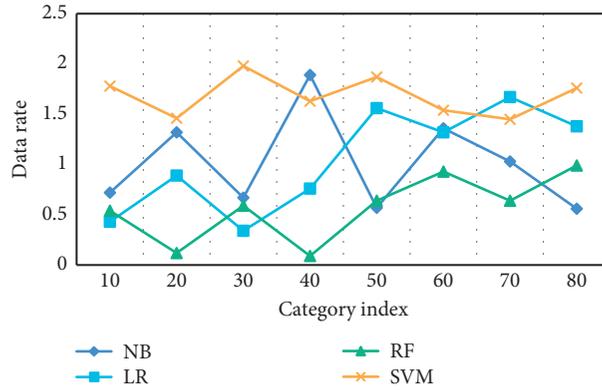


FIGURE 4: Experimental results of the first data news automated production technology.

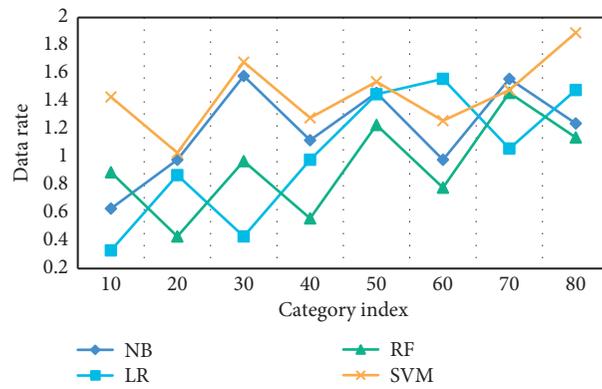


FIGURE 5: Experimental results of the second data news automated production technology.

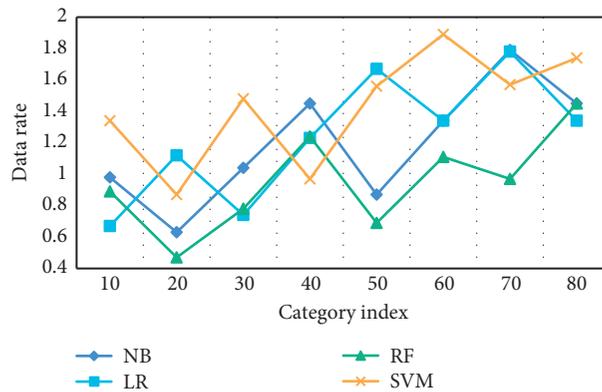


FIGURE 6: Experimental results of the third data news automated production technology.

algorithms. Businesses can create an intelligent management system based on the characteristics and current needs of product processing and production, collect equipment operating parameters and required data information, and then analyze and process it using computers. When combined with the current situation, scientific and feasible control measures can be implemented to provide effective guidance for future production operations. Expanding automated production technology can also provide strong technical

support for data updates and cross-departmental communication, create a more user-friendly operation interface, and improve the interactivity of production equipment. News can have intelligent characteristics after a series of processes. By referring to their own needs and preferences, users can adjust equipment functions and change their appearance to meet the personalized needs of products and services. This is the key to reflecting the interactive aspects of data news design and production in a natural way.

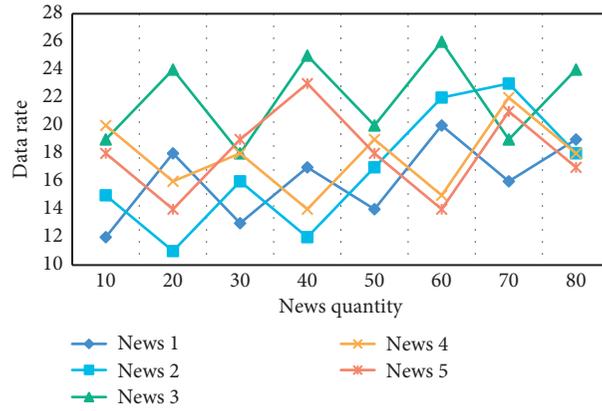


FIGURE 7: Comparative results of the first data news experiment based on machine learning algorithm.

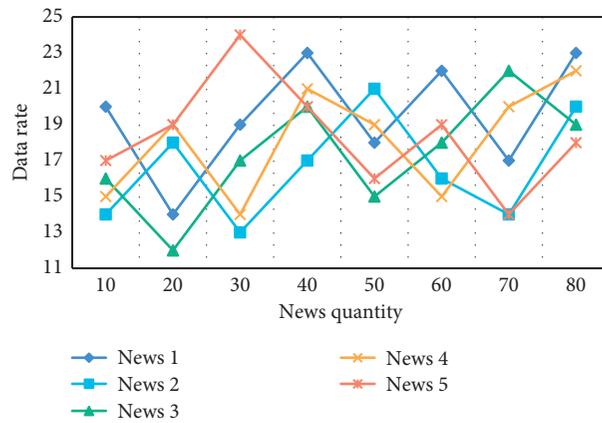


FIGURE 8: Comparative results of the second data news experiment based on machine learning algorithm.

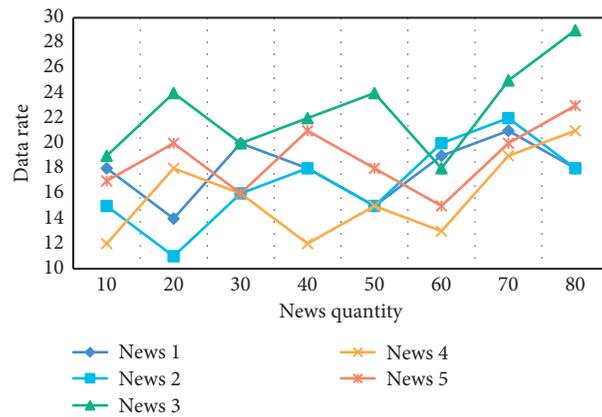


FIGURE 9: Comparative results of the third data news experiment based on machine learning algorithm.

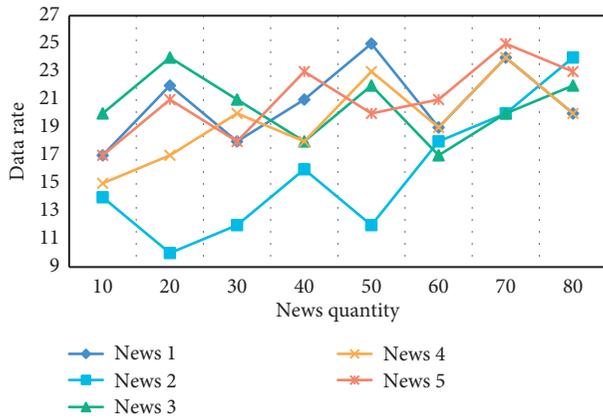


FIGURE 10: Comparative results of the fourth data news experiment based on machine learning algorithm.

5. Conclusion

The solution for automatic data news production technology realizes the automatic data news generation method, the data news knowledge map method, and the data news short video generation method, and constructs the data news automatic production technology. In the network era, the birth and continued development of machine learning algorithms is critical to large-scale data information processing. Based on machine learning, this paper investigates the emotional tendency and intensity of foreign exchange news. Naive Bayes, logistic regression, random forest, and support vector machine are used to mine foreign exchange news emotions in terms of emotional tendency. In the future, it will be necessary to continue to investigate automation technology so that it can be better applied in the machinery manufacturing industry and China's machinery manufacturing level can be improved. More description templates and special events will be added in a later stage to improve the quality of the final news generated, and the system will be applied to other sports events, continue to explore cross-application business scenarios of short videos and other emerging technologies, improve the functions of the entire news generation system, and meet the needs of various sports fans and news audiences in other fields.

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 does not have any possible conflicts of interest.

Acknowledgments

This study was supported without any funding.

References

- [1] X.-B. Jin, W.-Z. Zheng, J.-L. Kong et al., "Deep-learning temporal predictor via bidirectional self-attentive encoder-decoder framework for IOT-based environmental sensing in intelligent greenhouse," *Agriculture*, vol. 11, no. 8, p. 802, 2021.
- [2] S. Kamburugamuve, P. Wickramasinghe, S. Ekanayake, and G. C. Fox, "Anatomy of machine learning algorithm implementations in MPI, Spark, and Flink," *International Journal of High Performance Computing Applications*, vol. 32, no. 1, pp. 61–73, 2018.
- [3] K. Yan, L. Liu, Y. Xiang, and Q. Jin, "Guest editorial: AI and machine learning solution cyber intelligence technologies: new methodologies and applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6626–6631, 2020.
- [4] L. A. Lopes, V. P. Machado, R. A. L. Rabêlo, R. A. S. Fernandes, and B. V. A. Lima, "Automatic labelling of clusters of discrete and continuous data with supervised machine learning," *Knowledge-Based Systems*, vol. 106, pp. 231–241, 2016.
- [5] Y. Zhou, H. Wu, Q. Luo, and M. Abdel-Baset, "Automatic data clustering using nature-inspired symbiotic organism search algorithm," *Knowledge-Based Systems*, vol. 163, pp. 546–557, 2019.
- [6] J. Kong, C. Yang, J. Wang et al., "Deep-stacking network approach by multisource data mining for hazardous risk identification in iot-based intelligent food management systems," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1194565, 16 pages, 2021.
- [7] A. Bustillo, D. Y. Pimenov, M. Mia, and W. Kapłonek, "Machine-learning for automatic prediction of flatness deviation considering the wear of the face mill teeth," *Journal of Intelligent Manufacturing*, vol. 32, pp. 895–912, 2020.
- [8] P. O'Donovan, K. Leahy, C. Gallagher, and D. T. J. O'Sullivan, "A comparison of fog and cloud computing cyber-physical interfaces for industry 4.0 real-time embedded machine learning engineering applications," *Computers in Industry*, vol. 110, pp. 12–35, 2019.
- [9] X. Zhou and P. Jiang, "Variation source identification for deep hole boring process of cutting-hard workpiece based on multi-source information fusion using evidence theory," *Journal of Intelligent Manufacturing*, vol. 28, no. 2, pp. 255–270, 2017.
- [10] Y. Xu, S. Thomassey, and X. Zeng, "Machine learning-based marker length estimation for garment mass customization," *International Journal of Advanced Manufacturing Technology*, vol. 113, pp. 3361–3376, 2021.
- [11] A. Hürkamp, S. Gellrich, A. Dér, C. Herrmann, K. Dröder, and S. Thiede, "Machine learning and simulation-based surrogate modeling for improved process chain operation," *International Journal of Advanced Manufacturing Technology*, vol. 117, pp. 2297–2307, 2021.
- [12] B. Chen, J. Wan, Y. Lan, M. Imran, D. Li, and N. Guizani, "Improving cognitive ability of edge intelligent IIoT through machine learning," *IEEE Network*, vol. 33, no. 5, pp. 61–67, 2019.
- [13] D. Mccarthy, "Collaborate with computers to improve asset management," *Control Engineering*, vol. 64, no. 12, pp. 61–62, 2017.
- [14] D. Mccarthy, "Collaborate with computers to improve asset management," *Control Engineering*, vol. 64, no. 12, pp. 19–20, 2017.

- [15] T. Qiu, H. Wang, K. Li, H. Ning, A. K. Sangaiah, and B. Chen, "SIGMM: a novel machine learning algorithm for spammer identification in industrial mobile cloud computing," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2349–2359, 2019.
- [16] J. G. Nam, J. N. Witanto, S. J. Park, S. J. Yoo, J. M. Goo, and S. H. Yoon, "Automatic pulmonary vessel segmentation on noncontrast chest CT: deep learning algorithm developed using spatiotemporally matched virtual noncontrast images and low-keV contrast-enhanced vessel maps," *European Radiology*, vol. 31, no. 12, pp. 9012–9021, 2021.
- [17] V. Utomo and J.-S. Leu, "Automatic news-roundup generation using clustering, extraction, and presentation," *Multi-media Systems*, vol. 26, no. 2, pp. 201–221, 2020.
- [18] S. Ioannis and P. Steve, "An introduction to infinite HMMs for single-molecule data analysis," *Biophysical Journal*, vol. 112, no. 10, pp. 2021–2029, 2018.
- [19] S. Khader, N. Anuraj, F. Xavier Romero Duran, and H. González-Díaz, "Editorial: improving neuropharmacology using big data, machine learning and computational algorithms," *Current Neuropharmacology*, vol. 15, no. 8, pp. 1058–1061, 2017.
- [20] X. Luo, J. Liu, D. Zhang, and X. Chang, "A large-scale web QoS prediction scheme for the industrial internet of things based on a kernel machine learning algorithm," *Computer Networks*, vol. 101, no. 4, pp. 81–89, 2016.
- [21] J. Taninaga, N. Yu, K. Fujibayashi et al., "Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: a case-control study," *Scientific Reports*, vol. 9, no. 1, pp. 1–9, Article ID 12384, 2019.
- [22] K. Indira and M. K. Kavithadevi, "Efficient machine learning model for movie recommender systems using multi-cloud environment," *Mobile Networks and Applications*, vol. 24, no. 6, pp. 1872–1882, 2019.