

Research Article

The Main Influencing Factors of Artificial Intelligence Big Data Sports News Coverage

Zhiyu Leng,¹ Siyuan Li ,² and Jie Li¹

¹Hoseo University, Asan City, Chungcheng-Namdo, Republic of Korea

²College of Physical Education, Adamson University, Manila 0900, Philippines

Correspondence should be addressed to Siyuan Li; siyuan.li@adamson.edu.ph

Received 18 April 2022; Revised 15 June 2022; Accepted 29 June 2022; Published 12 July 2022

Academic Editor: Chia-Huei Wu

Copyright © 2022 Zhiyu Leng et al. 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.

Through systematic analysis and research on the reporting methods of various sports events and the status quo of data news under the background of big data, it is found that sports news reports under the background of big data have presented new reporting forms and contents. Combining the application of open source data with sports news can make good use of relevant technical forces to improve the reading efficiency of sports news. This can consolidate old audiences, attract new audiences, and maximize the promotion of various sports events, allowing more audiences to participate. After distributing 336 questionnaires, 303 were recovered. A total of 303 answer sheets met the requirements. The successful recovery rate of the questionnaire was 90.18%. Among them, 172 were male, accounting for 56.77% of the valid questionnaires; 131 were female, accounting for 43.23% of the valid questionnaires. Full-time students accounted for a relatively high proportion of the survey respondents, with a total of 130 people, accounting for 42.90% of the valid questionnaires. It can be seen that most of the people who pay attention to sports news reports are young and middle-aged groups.

1. Introduction

With the rapid development of information technology, data collection and storage are becoming more and more convenient. All traces left by humans through electronic devices can be used as a source of information in devices and networks. McKinsey & Company first introduced the concept of big data: “Data has penetrated into multiple industries and has become an important factor in production.” The expansion of data into journalism has created a new reporting method and product format—data journalism. Data journalism is where journalists process data, obtain valid data, and express news content visually in the form of graphs or charts.

With the development of data development technology and data journalism technology, there are more and more applications of visual reports in sports news reports. The sports news center has inevitably changed from text to data. Using information technology to organize seemingly messy data into effective information, which is visually displayed on the website, can help viewers to intuitively understand

event information. Visualizations in sports news stories are no longer simple statistics of players or teams but include temporal and spatial data. In recent years, visual reporting has been widely used in large-scale news and has become one of the main ways of sports reporting.

Technological innovation is not only an interdisciplinary research field, but also a major systematic project studied by many scholars from the perspectives of economics, sociology, and management. Therefore, we need to systematically consider the influencing factors of sports science and technology innovation and fully recognize the complex characteristics of sports science and technology innovation, such as multisubjectivity, synergy, cross-industry, and interdisciplinary. This paper comprehensively grasps sports technology innovation from the overall environment composed of internal and external.

2. Related Work

Hassabis et al. investigate the historical interaction between the fields of artificial intelligence and neuroscience and

highlight current advances in artificial intelligence that are inspired by research on neural computing in humans and other animals [1]. Li et al. introduce the basic concepts in AI and discuss the relationship between AI and candidate technologies in 5G cellular networks [2]. Liu et al. believe that artificial intelligence (AI) technology is getting more and more attention from academia and industry. However, artificial intelligence methods face great challenges under different practical operating conditions [3]. Thrall et al. believe that global interest in packaged artificial intelligence (AI) applications including imaging is high and growing rapidly, driven by the availability of large datasets (“big data”), significant advances in computing power, and new deep learning algorithms [4]. The above studies have carried out a detailed analysis of the application of data mining technology and pattern recognition. It is undeniable that these studies have greatly promoted the development of the corresponding fields. Caviglione et al. use two detection methods based on artificial intelligence tools such as neural networks and decision trees to discover malware secretly exchanging data [5]. Bui et al. proposed and validated a new hybrid artificial intelligence method called particle swarm optimization neuro-fuzzy (PSO-NF) for spatial modeling of tropical forest fire susceptibility [6]. Lemley et al. argue that a variety of long-standing problems in machine learning, artificial intelligence, and computer vision have significantly improved, which often enough to break long-standing performance barriers [7]. Wodecki considers artificial intelligence (AI) to be a fascinating concept whose origins can be found in the mid-20th century [8]. According to Price and Flach, state-of-the-art tools from machine learning and artificial intelligence are gradually automating parts of the peer review process. However, there are still many opportunities for further improvement [9]. Burton et al. provide helpful case studies and links to resources for AI educators. They also provide specific advice on how to integrate AI ethics into AI-in-general courses and how to teach stand-alone AI ethics courses [10]. Bin and Kumbier believe that artificial intelligence (AI) is inherently data-driven, which requires the application of statistical concepts through human-machine collaboration in the process of data generation, algorithm development, and outcome evaluation. They discuss how this human-machine collaboration is achieved [11].

3. Methods of Artificial Intelligence and Big Data

3.1. Artificial Intelligence. Artificial intelligence is a branch of computer science, a new technology that studies the laws of human intellectual activities, and is used to extend, expand, and simulate human intelligence. Artificial intelligence is a synthesis of computing, cyberspace, information theory, mathematics, psychology, and other disciplines. This is a new marginal subject developed on the basis of mutual penetration. AI has been widely used in fingerprint and facial recognition, intelligent search, evidence theory, expert systems, games, automatic programming, and space [12].

Since the 1990s, the amount of data has exploded. Especially after 2012, the rapid construction of the Internet, the explosive growth of online geometric data, and the richness of machine learning training materials have increased significantly. By learning from data, the machine continuously accumulates experience and optimizes decision parameters. The research of artificial intelligence has changed from a single agent to an artificial intelligence that spreads according to the network environment.

Other major research in artificial intelligence is information representation, automatic search, learning and acquisition of data, computer processing systems, understanding of natural language, computer vision, automatic programming, intelligent robots, etc. At present, the three focuses of artificial intelligence research are agents and multiagent systems, intelligent user interfaces, and intelligent computers. Intelligent interface technology is the study of how to enable people to use computers conveniently and naturally.

Computer vision: artificial intelligence is mainly used in the field of computer vision, which is to simulate the human visual system, so that the computer can have the ability to determine the position, motion status, and recognition of objects

Natural language processing: the natural language process is an important part of artificial intelligence applications, which can effectively realize human-computer interaction. Algorithmic learning based on big data and parallel computers has developed many natural language processes [13]

In terms of intelligent robots: intelligent robots are robots that can independently control themselves and can recognize and interact effectively with the external environment. Intelligent robots also include thinking elements, which can partially receive external information to process and respond accurately

Other applications: various methods obtained in AI research can also be used in other fields such as auxiliary education, auxiliary planning, and disease diagnosis. For example, computer doctors can diagnose many patients at the same time, quickly and accurately assess the details of the disease, recommend appropriate drugs and dosages, and monitor for 24 hours. “Computer teachers” can teach according to the specific situation of students. “Computer secretary” can accurately and comprehensively record all conversations, telephone calls, etc., and quickly provide the required information

3.2. Big Data. Although the term “big data” is well known, there is no unified definition in the industry [14]. Gartner, a big data research organization, defines “big data” as a large, high-quality, and diverse source of information. It requires a new process model to achieve more powerful effects than the decision-making, perception, discovery, and optimization processes.

Regarding intensive data science, the famous database scientist and Turing Award winner proposes a fourth paradigm in the book. Intensive data science has found that it is necessary to separate intensive data science from data

science, given the unique needs and technical support characteristics of intensive data science research. This new approach to scientific research is known as the fourth paradigm or data exploration. 4Paradigm provides platform-centric AI solutions and uses core technologies to develop end-to-end enterprise-level AI products. The meanings of the four scientific paradigms are shown in Table 1:

The strategic importance of big data lies not in controlling it, but in “processing” it, which means realizing greater value [15]. Big data is widely used in various fields of social production, which is also a big data analysis process that allows users to make decisions and understand potential value. The comprehensive analysis model adopts nontraditional data management systems according to the characteristics of large changes in data, which are difficult to analyze and process through distributed platforms. Traditional data cannot be used to store datasets, nor can they be used to store the output/response of datasets. The big data architecture of the model is shown in Figure 1.

3.3. The Relationship between Big Data and Artificial Intelligence. The main difference between AI and big data is that big data needs to be structured, integrated, and cleaned, while AI is the output [16], which creates a fundamental difference between the two. They are also used differently, with big data typically being used for insights, allowing research and recommendations for movies or TV shows to viewers based on what people are watching. Taking into account customer habits and what they like, the conclusion is that customers tend to feel the same way.

Although they are different, this is a good fit for big data, because artificial intelligence needs data to build, especially machine learning. For example, a machine learning imaging application can look at tens of thousands of aircraft images to learn and identify what an aircraft is made of.

The more data an AI application has, the more accurate the results will be. In the past, AI took time to process data, so it did not work well. Without advanced sensors and the internet like today, it is difficult to provide real-time data. Today, people need fast processors, input devices, networks, and large datasets. Artificial intelligence requires a large amount of data as the basis for “thinking” and “decision-making”; big data requires artificial intelligence technology for data value operations, so without big data, there is no artificial intelligence.

3.4. Category Construction. Use the “frame list” method to find news frames, which provides a “frame list” with 11 reference indicators. The 11 reference metrics are title, subtitle, image, image description, lead, news source, and citation. Often, we select and cite the special series icon, date and graph, and epilogue or short comment to which the article belongs. Although the “frame list” can be used for quantitative statistics, it is somewhat insufficient for systematic classification [17]. In the research direction of “discourse structure” proposed by the researchers, four analytical dimensions are given to analyze news texts, namely, syntactic structure, manuscript structure, topic structure, and rhetorical structure. The research direction of the researchers

proposes the “tertiary structure” method of media authenticity construction. In this research direction, the construction of high-level media certification is represented by introductions, titles, etc.; the construction of intermediate-level media certification includes manifestations such as main events, history, attribution, and criticism; the real construction of low-level media is expressed through rhetoric and grammar. Although researchers have conducted dimensional systematic analysis of texts, they are more suitable for qualitative research. Therefore, this paper draws on the advantages of each research orientation and establishes a research category as shown in Figure 2.

In the past two years, cover media has moved towards AI + media and has continuously adjusted and optimized according to strategic positioning, data analysis, and user needs and upgraded from version 1.0 to version 4.0. (1) News first, and immersive experience increases user stickiness. (2) The listening position is prominent, and listening to the news enhances the user experience. The specific operation structure is shown in Figure 3.

4. Methods and Results of Sports News Reports

4.1. Methods

4.1.1. Literature Method. According to the research content, this article has consulted a large number of related monographs in the fields of sports, linguistics, discourse, journalism, and communication; this article refers to journals, master’s thesis, and doctoral dissertations on news language in the “China Academic Journals” database, classified for reference [18].

4.1.2. Case Analysis Method. This article selects words, phrases, or complete texts that can reflect the language characteristics of sports news reports as cases and analyzes their use to support the views expressed in this article.

4.1.3. Comparative Analysis. This article compares and analyzes the characteristics of sports news discourse analysis that are different from grammatical positions, different from phrase pattern selection, and different topic selections.

4.1.4. Mathematical Statistics. This paper deals with the data of selected print media and uses Office soft equipment to perform statistical analysis on the data according to statistical principles.

4.1.5. Questionnaire Survey Method. The research object is the top 20 urban newspapers in the 2006 China Newspaper Industry Competitiveness Monitoring Report.

4.2. Data Sources and Visualization Forms. For news, data quality determines news quality. The analysis of data sources shows that there are 18 articles without data sources and 95 articles with cross data sources (referring to the collection of different data sources). Among the personal data, the data of 17 articles are from the Xinhua News Agency’s internal news agency database, 8 are from the government and government websites, 4 are from authoritative reports and third-party databases, and 4 are from the Internet and other media

TABLE 1: Four scientific paradigms.

Time	Scientific paradigm	Basic content
Thousands of years ago	Experimental science	Describe natural phenomena
In recent hundreds of years	Theoretical science	Using models and scientific induction
In recent decades	Computational science	Simulate complex situations
Nowadays	Data exploration	Comprehensive unification of theory, experiment, and simulation; collecting and integrating data through equipment or simulator; use software to process data; use computer to store information; scientists call databases to analyze data

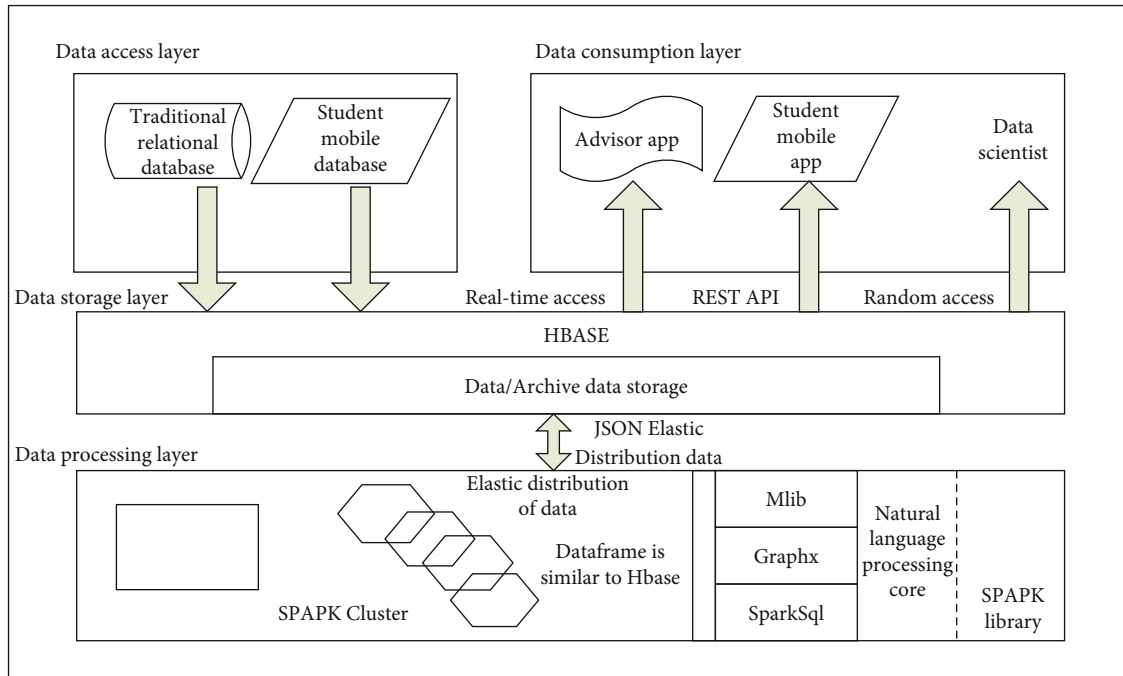


FIGURE 1: Big data architecture.

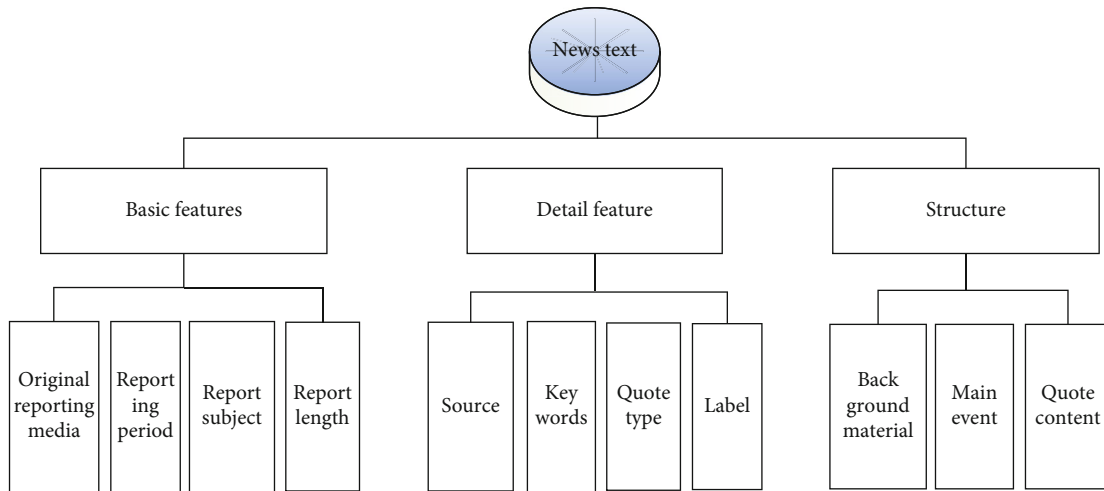


FIGURE 2: Schematic diagram of category construction.

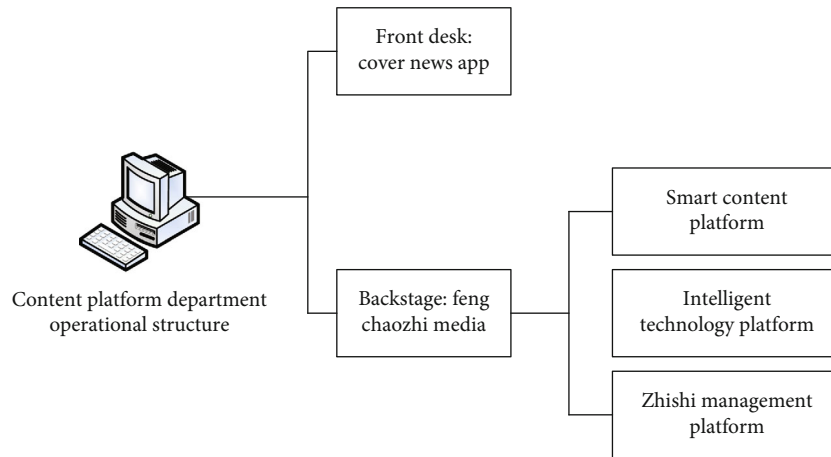


FIGURE 3: Operational architecture of cover news content platform.

databases [19]. The specific data sources are shown in Figure 4.

Since data news is presented directly to the public, visual form is an important criterion for measuring the quality of data news. “Visual form” is an independent aesthetic and artistic concept originating from the art history and aesthetic theory of German-speaking countries in the late 19th and early 20th centuries. Through the analysis and statistics of 154 new visual form instances, it can be seen that the visual form is most used in the column form, followed by the graphic form. Among them, dynamic visual news accounts for about 10%, and the rest is visual static news. The specific statistics are shown in Figure 5.

From this, it can be seen that Xinhua’s “data journalism” is able to handle the skills of images and texts and has edited more content according to media conditions and column requirements, even complete interactive visual dynamic data journalism with high technology [20]. However, the current level of visual form is still in a transitional stage that continues to move closer to the mature stage.

4.3. The Total Amount of Sports News In-Depth Reports in Sports News Reports. According to the location of the event, it is divided into Japanese sports reports and Shanghai sports reports. It is divided into school sports and social sports according to the different people in the sports world. The statistics of sports-related reports in Shanghai Daily News are shown in Table 2.

It can be seen that sports news in Shanghai is mainly reported by “Shanghai Daily News,” with a total of 565 articles, accounting for 50% of the total reports. Second, 164 major international incidents were reported, accounting for 14.1% of the total. The number of local sports reports in Japan was 261, accounting for 23% of the total. In Japanese local sports reports, school sports reports are basically the same as social sports reports; in Shanghai sports reports, there are 493 social sports reports and 72 school sports reports, of which social sports reports account for the largest proportion of all kinds of reports. It can be seen that the focus of “Shanghai Daily Sports News” is the sports events in Shanghai [21].

From the point of view of the public at the time, sports coverage formed a clear staging of seasonal changes. Taking 1933 as an example, the change in the number of reports with the seasons is shown in Figure 6.

There is a clear trend in the number of reports over time. During the coldest period from January to March, there are fewer sports reports. As temperatures rise, so does the amount of sports coverage, peaking between July and October. With cooler temperatures in November and December, the number of sports coverage is trending downward.

This section mainly analyzes and researches the reports of China Sports News, Sports Weekly, and Liaoshen Evening News during the NBA08-09 season finals, NBA09-10 season finals, and NBA10-11 season finals. And it analyzed in-depth coverage of sports news as a percentage of total sports coverage. The statistics of the number of sports news in-depth reports in sports news reports are shown in Table 3.

The statistical results in Table 3 show that Sports Weekly’s coverage of the NBA Finals shows a fluctuation curve from 333 in 2009 to 212 in 2011. In 2009, the aftermath of the Beijing Olympic Games did not completely subside, nor did the attention of the outside world completely subside. The whole society’s attention to sports events is still at a relatively high stage, which is why the reports on sports events in 2009 have attracted more and more attention from a wide audience. In 2010, with the opening of the World Cup in South Africa, media and public attention to the NBA Finals were affected. During the year, several newspapers reported relatively little coverage of the NBA Finals. In 2011, with the advent of the era of all media, the competition of print media became fierce, and emerging media such as the Internet and mobile phones were embattled on all sides, and the public was seriously divided. Print media is facing a crisis of marginalization, and the public’s attention to sports events is often rational [22].

Through the analysis of a large number of corpus, the characteristics of sports news reports in terms of vocabulary and rhetoric, this paper selects all 180 reports in the sports pages of “Sports Weekly,” “China Sports News,” and “Liaoshen Evening News” on November 22, 2010, for quantitative analysis and research. Figure 7 shows the difference

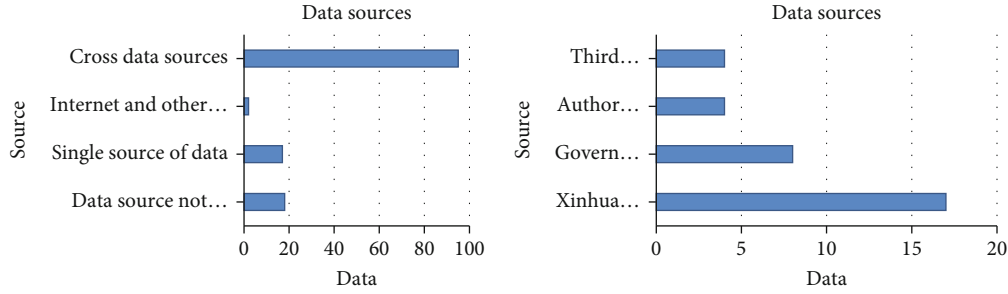


FIGURE 4: Sources of data in the “data news” column of Xinhuanet in the first half of 2017.

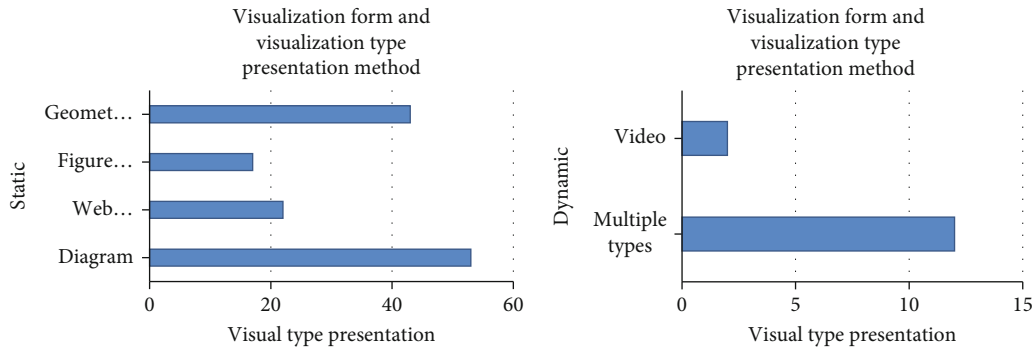


FIGURE 5: Stacked bar chart of the visualization form of the “data news” column of Xinhuanet in the first half of 2017.

TABLE 2: Statistics of the number of sports-related reports in Shanghai Daily News.

Years	Japanese school sports	Japanese social sports	Sports at Lu school	Social sports at Lu	Major events	Other
August-November 1932	44	1	24	39	39	29
January-December 1933	50	50	26	192	52	64
January-March 1934	4	15	8	45	25	17
July 1935	3	3	0	39	7	10
January, March, May, June, July 1937	36	55	14	136	41	37
Subtotal	137	124	72	493	164	157
Percentage	12%	10.9%	6.4%	42.2%	14.1%	13.7%

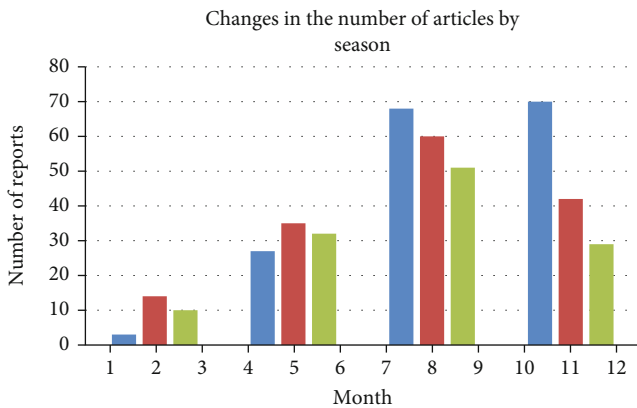


FIGURE 6: Changes in the number of reports with seasons.

in vocabulary usage between headlines and body texts in sports news reports.

It can be clearly seen from Figure 7 that the color words, rhetorical words, and colloquial words in the title are significantly higher than their proportions in the text, which has a lot to do with the layout position of the title in the news. Because the audience is very random in reading news reports, its layout advantage determines its leading role, so a headline that can attract audiences is particularly important in sports news [23].

5. Artificial Intelligence Big Data Sports News Reports

5.1. Algorithms of Artificial Intelligence and Big Data. Naive Bayes is mainly used for classification, such as news classification, text classification, and patient classification. The

TABLE 3: Statistics on the number of sports news in-depth reports in sports news reports (unit: article).

	08-09 season		09-10 season		10-11 season	
	Report	In-depth report	Report	In-depth report	Report	In-depth report
China Sports News	38	17	22	11	20	16
Sports Weekly	333	108	157	90	212	106
Liaoshen Evening News	14	6	7	4	13	6
Total	385	133	186	103	245	128

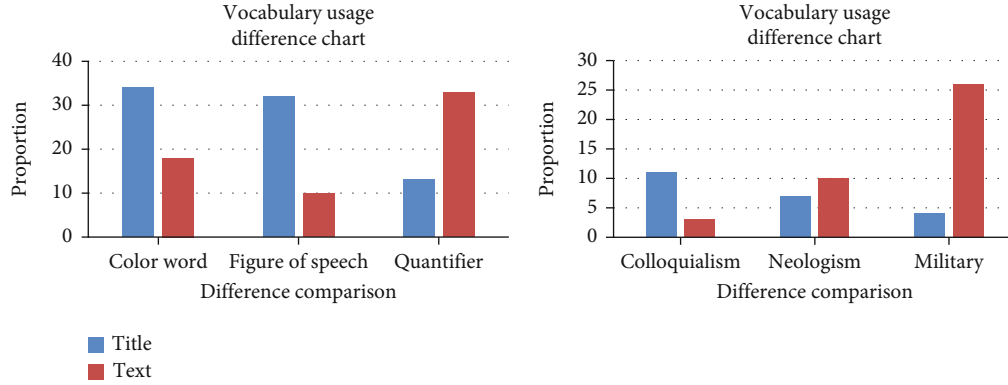


FIGURE 7: Differences in vocabulary usage between headlines and body texts in sports news reports.

naive Bayes formula based on the Bayes formula is as follows:

$$p(n|m) = \frac{p(m|n)p(n)}{p(m)}, \quad (1)$$

where $p(n|m)$ means that m determines the probability of n occurrence when it has occurred. For example, $p(\text{cold/sneeze, fever})$ means the probability of a cold when sneezing and fever symptoms occur; $p(n)$ means the probability of n occurrence when there is no precondition. The maximum entropy model is based on the concept of entropy; the definition of entropy is derived from physical thermodynamics and is used to express the uncertainty of a material in the molecular state. The entropy mentioned in the artificial intelligence algorithm is "Shannon information entropy," and the formula is as follows:

$$H(Y) = -\sum_y p(y) \log p(y) = -\sum_j^m p(y_j) \log p(y_j), \quad (2)$$

where y represents a random variable and $p(y)$ represents the probability distribution of y . The greater the uncertainty of the probability distribution, the greater the entropy. The maximum entropy model has been successfully applied to natural language processing, such as machine translation, partial speech tagging, word segmentation, and text classification.

The ratings of all users for all items are recorded in the database, represented by a rating matrix R . $R_{ij} \in \{0, 1\}$ express dislike; $R_{ij} \in \{1-5\}$ express like. The higher the

score, the higher the preference. The scoring matrix is shown in Table 4.

The Pearson correlation coefficient is a method used to describe the degree to which two vectors are related. The larger the value is, the greater the correlation degree is, and the smaller the value is, the smaller the correlation degree is. Taking user-based CF as an example, then set the vector

$$I(a) = (R_{a1}, R_{a2}, R_{a3} \cdots R_{an}), \quad (3)$$

where $I(a)$ represents the set of ratings for all items by user a , that is, the row corresponding to user a in the rating matrix. The formula for calculating the similarity between user a and user b by the Pearson correlation coefficient is as follows:

$$\text{sim}(a, b) = \frac{\sum_{n \in I(a) \cap I(b)} (R_{an} - \bar{R}_a)(R_{bn} - \bar{R}_b)}{\sqrt{\sum_{n \in I(a) \cap I(b)} (R_{an} - \bar{R}_a)^2} \sqrt{\sum_{n \in I(a) \cap I(b)} (R_{bn} - \bar{R}_b)^2}}. \quad (4)$$

Among them, \bar{R}_a and \bar{R}_b are the average scores of user a and user b for all items.

Cosine similarity is used to calculate the similarity between texts. The correlation between two vectors is described by calculating the angle between them. The smaller the angle, the more similar the two vectors; the larger the angle, the less similar. In the rating matrix, each row represents the rating vector of each user, and the formula for calculating the similarity between user a and user b by cosine

TABLE 4: Scoring matrix.

	I_1	I_2	I_3	I_4
A_1	4	3	4	3
A_2	4	4		2
A_3	5	5	4	
A_4	4	2	3	4
A_5	2			5

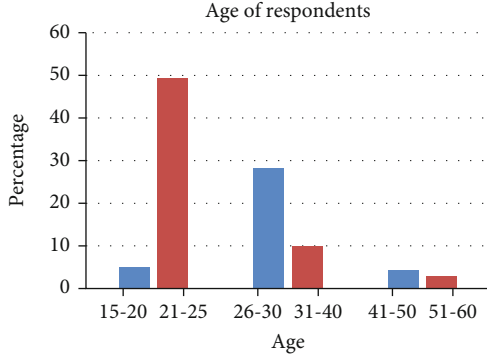


FIGURE 8: Age group of respondents.

similarity is as follows:

$$\text{sim}(a, b) = \cos(a, b) = \frac{a \times b}{|a| \times |b|} = \frac{\sum_{i=1}^n R_{ai} R_{bi}}{\sqrt{\sum_{i=1}^n R_{ai}^2} \sqrt{\sum_{i=1}^n R_{bi}^2}}. \quad (5)$$

The formula for calculating the similarity between user a and user b through the modified cosine similarity is as follows:

$$\text{sim}(a, b) = \frac{\sum_{i=1}^n (R_{ai} - \bar{R}_a)(R_{bi} - \bar{R}_b)}{\sqrt{\sum_{i=1}^n (R_{ai} - \bar{R}_a)^2} \sqrt{\sum_{i=1}^n (R_{bi} - \bar{R}_b)^2}}. \quad (6)$$

Recommendation prediction, predicting the rating of item i by user a , can be calculated by the items that have been rated by neighbor users. The calculation formula is as follows:

$$P(a, i) = \frac{\sum_{b \in K_{\text{neighbor}} \cap A(i)} \text{sim}(a, b) \times R_{bi}}{\left| \sum_{b \in K_{\text{neighbor}} \cap A(i)} \text{sim}(a, b) \right|}, \quad (7)$$

where $P(a, i)$ represents the predicted rating of item i by user a , $U(i)$ represents the set of users who rated item i , and $\text{sim}(a, b)$ represents the similarity between user a and user b . The above formula does not take into account the difference between user ratings, so the average rating of users is

added, and the calculation formula is as follows:

$$P(a, i) = \bar{R}_a + \frac{\sum_{b \in K_{\text{neighbor}} \cap A(i)} \text{sim}(a, b) \times (R_{bi} - \bar{R}_b)}{\left| \sum_{b \in K_{\text{neighbor}} \cap A(i)} \text{sim}(a, b) \right|}. \quad (8)$$

When evaluating the quality of recommendation algorithms, two methods are often used: mean absolute deviation (MAE) and root mean square error (RSME), calculated as follows:

$$\text{MAE} = \frac{\sum_{i \in I_z} |P_{z,i} - r_{z,i}|}{N}. \quad (9)$$

In the calculation results, the smaller the MAE value, the higher the accuracy of the algorithm.

The FCM algorithm is a method that can automatically classify data samples. In the FCM algorithm, the objective function 1 is determined by the following formula:

$$J_i[U, v] = \sum_{m=1}^n \sum_{j=1}^c \mu_{mj}^i \cdot d_{mj}^2 \quad i \in (1, \infty). \quad (10)$$

For fuzzy databases $D^T = \{t^1, \dots, t^m\}$, if $X = \{x_1, \dots, x_p\}$ is $A = \{a_1, \dots, a_p\}$, then $Y = \{y_1, \dots, y_p\}$ is $B = \{b_1, \dots, b_q\}$.

This means that if there is a record that satisfies this rule antecedent, the rule consequent has a high probability. The calculation formulas of fuzzy support and fuzzy confidence are defined as follows:

$$\text{FS}_{Z,C} = \frac{\sum_{i=1}^n \prod_{j=1}^m \mu(t^i[(z_j, c_j)])}{|D^T|}, \quad (11)$$

$$\text{FC}_{((X,A),(Y,B))} = \frac{\text{FS}_{(Z,C)}}{\text{FS}_{X,A}}.$$

Data normalization transformation: scaling data so that it is concentrated in a small range. It removes data cell boundaries and converts data to pure values with no dimensions, allowing metrics for different commands or cells to be compared or weighted. Min-max normalization: it is to linearly transform the original data so that the result is mapped between [0-1]. The conversion function is as follows:

$$x^* = \frac{x - \min}{\max - \min}. \quad (12)$$

Weight value of frequent itemsets: in the sequence database S , if P is a frequent sequence and $P = (S_1 S_2 \dots S_l)$, then the weight value of the frequent itemset P is equal to the cumulative sum of the weights of the frequent itemset/the length of the frequent itemset.

$$\text{Weight}(P) = \frac{\sum_{n=1}^{\text{length}(P)} \text{Weight}(P)}{\text{length}(P)}. \quad (13)$$

Sequential pattern value $\text{VSP}(P)$: the sequential pattern

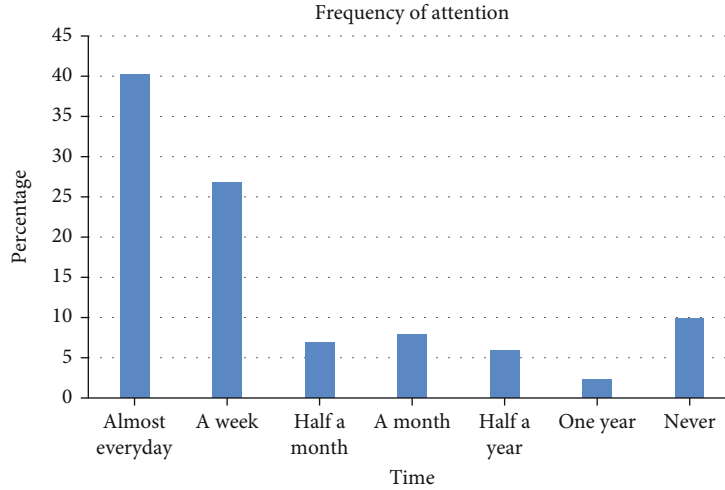


FIGURE 9: The frequency of audience's attention to sports news reports.

value is used to measure the importance of frequent sequences and is mainly obtained with reference to the weight value and the support degree. The sequence state value VSP is equal to the weight value * support.

$$\text{VSP}(P) = \text{Weight}(P) * \text{Support}(P). \quad (14)$$

In the association rule algorithm, three important concepts are defined for the association rule AB:

Definition 1 Support (support).

$$\text{support}(A \Rightarrow B) = P(A \cap B). \quad (15)$$

It represents the probability that event A and event B occur at the same time in the database transaction set D .

Definition 2 Confidence.

$$\text{confidence}(A \Rightarrow B) = P(B|A) = \frac{P(A \cap B)}{P(A)}. \quad (16)$$

It represents the probability of event B occurring when event A has already occurred in database transaction set D .

Definition 3 Frequent itemsets. Let $I = \{a_1, a_2, a_3, \dots, a_m\}$ be the set of items and $D = \{T_1, T_2, T_3, \dots, T_n\}$ the set of database transactions, where for any $T_i \in D$, there is a $T_i \subseteq I$. For any subset I_j of I , it is called an itemset. If I_j contains K elements, I_j is called a K itemset.

If itemset I_j satisfies the minimum support min_support given by the user, then I_j is called frequent itemset.

5.2. The Main Influencing Factors of Sports News Coverage.

Since the reform of China's press and publication system, media and publication sources, such as sports information sources, author sources, professional editors, and special reporters, have been reorganized through the market com-

petition mechanism. The rapid development of modern technology and transportation has weakened the natural benefits of urban newspapers. With the development of modern sports in the direction of specialization and stardom, the content and volume of sports news in Metropolis Daily have increased rapidly, and the trend of marginalization seems to meet the needs of leisure and entertainment readers as the main reading destination. Compared to professional sports news, theory, science, technology, knowledge, and originality, sports news has a much lower marginal cost of content, convincing coverage. Therefore, the cost of comparative reporting has become another important factor affecting the reporting of China Metropolis Daily [24].

The operation of China's sports science and technology innovation system is essentially the regulation of the process of sports science and technology innovation. At present, the problems to be solved urgently in China's sports science and technology innovation are as follows: the transformation of sports science and technology innovation achievements is insufficient; the independent innovation ability of sports science and technology innovation is weak; the content of sports scientific research is uneven; the internationalization of sports scientists is progressing slowly. It is necessary to strengthen the construction of innovative ideas and innovative culture in the sports industry and to develop and improve the market mechanism and competition mechanism. It requires changing concepts, eliminating equality and dependence, attaching importance to psychology, forming a sports culture of courage and innovation, and having a leadership position in creating an internal environment of scientific innovation and sports-led technology. It encourages exploration, encourages risk taking, accepts failure, and respects different academic viewpoints. It adheres to different research directions of sports, adheres to the basic principle of combining mass sports and competitive sports scientific research, and strengthens sports industry scientific research. In terms of factors at the media agency level, there has been very little analysis at this level for a long time. It is a relatively new perspective, but it is now gradually catching

the attention of scientists. A media organization is generally considered a production organization, and its conditions and production constraints are specifically studied. Impact of media ownership: regarding the influence of media ownership on reported content, as a social system, the development of media is limited by its own political and economic characteristics. Market environment impact: as a production organization, the media is not a closed system but must respond to its own market environment. Generally speaking, market demand factors such as the size of the city where the media is located, the education level of the population, the level of family income and market competition, and the number of immigrants in international news will affect the dissemination of international news. Media power influence: due to the high cost of international news, its operation will be influenced by media power. Power factors include the size of the media, the availability of news agency sources, the number of international information staff, the number of international agencies, and the number of international journalists. Influence of media culture: media culture refers to the internal atmosphere, shared concepts, and work practices specific to the media organization.

Now the word “data news” is very popular, “data journalism,” also known as “data-driven journalism,” is a new way of news production explored by the journalism industry in the era of big data constantly adapting to changes in the media environment. But the use of data news basically appears in some people’s livelihood economic news, while sports news is rarely used.

This type of technology can help sports fans understand sports and attract more sports fans, so that sports can be promoted in depth, breadth, and science. Therefore, if it want to be performed well in sports news, it must pay attention to the improvement of sports data news. In particular, the preview of the event needs the help of data to be more rational and efficient. In addition, the gaming industry is booming, especially in the World Cup, Champions League, and other events; a large number of fans will pay attention to the preview of the event. At this point, sports data journalism seems to be all the more important.

5.3. Questionnaire Survey. The survey was conducted by stratified random sampling method, and the age and occupation of the respondents were basically stratified and classified to reflect the situation of the research population to the greatest extent possible. A total of 336 questionnaires were distributed, and 303 were recovered. After review, a total of 303 answer sheets met the requirements, and the successful recovery rate of the questionnaire was 90.18%. Among them, 172 were male, accounting for 56.77% of the valid questionnaires; 131 were female, accounting for 43.23% of the valid questionnaires. The age of the respondents is shown in Figure 8.

It can be seen from the questionnaire survey that most of the people who pay attention to sports news reports are young and middle-aged groups. Affected by age and hobbies, the attention of the surveyed population to sports news reports varies. There are sports fans who pay attention to sports news every day, and nonsports fans who do not

receive such news information all year round. In general, the audiences that pay more attention to sports news coverage are full-time students and professionals engaged in related sports. The frequency of audience’s attention to sports news reports is shown in Figure 9.

In the questionnaire survey, the sports news coverage of the Summer Olympics is the most concerned by the audience, with a proportion of 76.58% (233 people) of the valid questionnaires. Judging from Xinhua News Agency’s sports news reports on the London Olympics, the use of hot topics, high-end interviews, project analysis, viewpoint comments, Olympic figures, essays, and other reporting, coupled with the launch of Xinhua Sports’ official Weibo and “Xinhua Xuanwen,” forms constitute a high-capacity, multiplatform sports news information dissemination channel, while ensuring the timeliness of the report and the effectiveness of meeting the high standards of the audience.

6. Conclusions

Today, when the sports industry is in full swing, we need to make reasonable use of big data. It requires the selection of visual charts and data analysis appropriate to the content of sports coverage and the addition of professional sports journalism skills. It must strengthen the horizontal connection between different print media and at the same time accelerate the integration with new media such as radio, TV media, online media, and mobile media. And it needs to carry out comprehensive, detailed, and objective news in-depth reports, in order to achieve a greater breakthrough in the in-depth reports of print media sports news. When reporting competitive sports activities, we promote the spirit of solidarity, love, cooperation, and never giving up in sports, so that the audience can truly understand the essence of sports.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors state that this article has no conflict of interest.

References

- [1] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, “Neuroscience-inspired artificial intelligence,” *Neuron*, vol. 95, no. 2, pp. 245–258, 2017.
- [2] R. Li, Z. Zhao, X. Zhou et al., “Intelligent 5G: when cellular networks meet artificial intelligence,” *IEEE Wireless Communications*, vol. 24, no. 5, pp. 175–183, 2017.
- [3] R. Liu, B. Yang, E. Zio, and X. Chen, “Artificial intelligence for fault diagnosis of rotating machinery: a review,” *Mechanical Systems & Signal Processing*, vol. 108, pp. 33–47, 2018.
- [4] J. H. Thrall, X. Li, Q. Li et al., “Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success,” *Journal of the American College of Radiology*, vol. 15, no. 3, pp. 504–508, 2018.

- [5] L. Caviglione, M. Gaggero, J. F. Lalande, W. Mazurczyk, and M. Urbanski, "Seeing the unseen: revealing mobile malware hidden communications via energy consumption and artificial intelligence," *IEEE Transactions on Information Forensics & Security*, vol. 11, no. 4, pp. 799–810, 2016.
- [6] D. T. Bui, Q. T. Bui, and Q. P. Nguyen, "A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area," *Agricultural and Forest Meteorology*, vol. 233, pp. 32–44, 2017.
- [7] J. Lemley, S. Bazrafkan, and P. Corcoran, "Deep learning for consumer devices and services: pushing the limits for machine learning, artificial intelligence, and computer vision," *IEEE Consumer Electronics Magazine*, vol. 6, no. 2, pp. 48–56, 2017.
- [8] A. Wodecki and A. Wodecki, "Artificial intelligence methods and techniques," in *Artificial Intelligence in Value Creation*, pp. 71–132, Palgrave Macmillan, Cham, 2019.
- [9] S. Price and P. A. Flach, "Computational support for academic peer review," *Communications of the ACM*, vol. 60, no. 3, pp. 70–79, 2017.
- [10] E. Burton, J. Goldsmith, S. Koenig, B. Kuipers, N. Mattei, and T. Walsh, "Ethical considerations in artificial intelligence courses," *AI Magazine*, vol. 38, no. 2, pp. 22–34, 2017.
- [11] Y. U. Bin and K. Kumbier, "Artificial intelligence and statistics," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 1, pp. 6–9, 2018.
- [12] A. Jeavons, "What is artificial intelligence?," *Research World*, vol. 2017, no. 65, pp. 75–75, 2017.
- [13] H. Lu, Y. Li, and C. Min, "Brain intelligence: go beyond artificial intelligence," *Mobile Networks and Applications*, vol. 23, no. 7553, pp. 368–375, 2017.
- [14] S. Makridakis, "The forthcoming artificial intelligence (AI) revolution: its impact on society and firms," *Futures*, vol. 90, no. jun., pp. 46–60, 2017.
- [15] P. Glauner, J. A. Meira, P. Valtchev, R. State, and F. Bettinger, "The challenge of non-technical loss detection using artificial intelligence: a survey," *International Journal of Computational Intelligence Systems*, vol. 10, no. 1, pp. 760–775, 2017.
- [16] T. Yang, A. A. Asanjan, E. Welles, X. Gao, S. Sorooshian, and X. Liu, "Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information," *Water Resources Research*, vol. 53, no. 4, pp. 2786–2812, 2017.
- [17] D. Berry, "Sports nutrition evolution," *Food Business News*, vol. 15, no. 9, p. 52, 2019.
- [18] J. Bryson and A. Winfield, "Standardizing ethical design for artificial intelligence and autonomous systems," *Computer*, vol. 50, no. 5, pp. 116–119, 2017.
- [19] A. Agrawal, J. S. Gans, and A. Goldfarb, "What to expect from artificial intelligence," *MIT Sloan Management Review*, vol. 58, no. 3, pp. 23–26, 2017.
- [20] "Society news," *The American Journal of Sports Medicine*, vol. 46, no. 5, pp. 1258–1259, 2018.
- [21] L. Li and J. Zhang, "Research and analysis of an enterprise E-commerce marketing system under the big data environment," *Journal of Organizational and End User Computing (JOEUC)*, vol. 33, no. 6, pp. 1–19, 2021.
- [22] D. R. Kanniga, M. Gurusamy, and P. Vijayakumar, "An efficient cloud data center allocation to the source of requests," *Journal of Organizational and End User Computing (JOEUC)*, vol. 32, no. 3, pp. 23–36, 2020.
- [23] J. Hou, Q. Li, Y. Liu, and S. Zhang, "An enhanced cascading model for E-commerce consumer credit default prediction," *Journal of Organizational and End User Computing (JOEUC)*, vol. 33, no. 6, pp. 1–18, 2021.
- [24] S. Liao and C. Ho, "Mobile payment and mobile application (app) behavior for online recommendations," *Journal of Organizational and End User Computing (JOEUC)*, vol. 33, no. 6, pp. 1–26, 2021.