International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 10 Issue 5, Sep-Oct 2026 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Sentiment Classification and Trend Analysis of Chinese Basketball Online Public Opinion Based on Machine Learning - A Case Study of the Chinese Men's National Basketball Team at the 2023 FIBA Basketball World Cup

Chen Guanlin, Cao Jingxuan, Hu Yilin

Beijing Wuzi University, Beijing, China

ABSTRACT

Sports-event online public opinion is characterized by strong event-driven dynamics, rapid emotional fluctuations, and frequent use of community slang and sarcasm. Taking the five official games of the Chinese men's basketball team at the 2023 FIBA Basketball World Cup as the event thread, this paper studies the sentiment classification of basketball online public opinion and its changes across match milestones. Considering the limited scale of basketball texts that can be traced back one by one and uniformly manually reviewed, this paper separates the machine-learning experiments from the event-application analysis: first, the public ASAP Chinese review dataset is used to compare five traditional models (KNN, Decision Tree, linear SVM, Bagging, and AdaBoost) and to select a benchmark model suitable for preliminary three-class screening of texts; the better-performing SVM is then applied to the auxiliary preliminary screening of 28 Hupu World Cup event-text units, and the final manual review is completed by combining basketball-domain terms with the original page context. To expand the scope of case evidence, this study further builds a cross-platform public-source index, organizing 39 Zhihu answer entries, 19 Hupu public pages or entries, and 13 Weibo historical trending-topic records directly related to the five games for issue-level cross-verification; indexed records for which complete Chinese source texts are unavailable are not countedas training samples or in the sentiment index. In the benchmark experiment, SVM achieves an Accuracy of 0.6688 and a Macro-F1 of 0.5531; its direct preliminary-screening accuracy on the 28 Hupu event texts is 46.43%, indicating that the generic model can only serve as an auxiliary tool. Based on the manually reviewed results, the sentiment indices of the five games are -0.667, -0.500, -1.000, 0.600, and -0.833 in sequence; sentiment clearly turns positive after the win over Angola and turns negative again after the loss to the Philippines. The correlation coefficient between point margin and sentiment index across the five games is about 0.741, but the sample size is small and the correlation result is treated as descriptive analysis only. Issue induction shows that discussion gradually extends from single-game outcomes and player performance to coach rotation, fundamentals, CBA competitive intensity, and youth training. The study shows that traditional machine learning can improve the preliminary-screening efficiency of sports public-opinion texts, but basketball slang, sarcasm, and multi-object evaluation still require manual calibration in combination with specific match contexts. This paper also provides an empirical basis for the subsequent construction of a multi-platform basketball-specific corpus, domain sentiment models, and automated public-opinion systems.

KEYWORDS: *Chinese men's basketball team; online public opinion; machine learning; sentiment analysis; sports events; cross-platform retrieval; manual review.*

How to cite this paper: Chen Guanlin | Cao Jingxuan | Hu Yilin "Sentiment Classification and Trend Analysis of Chinese Basketball Online Public Opinion Based on Machine Learning - A Case Study of the Chinese Men's National Basketball Team at the 2023 FIBA Basketball World Cup" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-10 | Issue-5, October 2026, pp.99-117, URL: www.ijtsrd.com/papers/ijtsrd142207.pdf

IJTSRD142207

Copyright © 2026 by author (s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0)

1. Introduction

1.1. Research Background and Significance

1.1.1. Research Background

Mobile Internet has changed the way sports events are watched and discussed. Online discussion of major matches usually runs through pre-match prediction, in-match interaction, and post-match review; player performance, coaching adjustments, and match results can all cause public-opinion changes within a short time. The core task of text sentiment analysis is to identify the attitude and emotional tendency in the text [1]; when this method is applied to sports events, the text must also be aligned with specific match milestones so that the reasons for emotional changes can be explained.

During the 2023 FIBA Basketball World Cup, the Chinese men's basketball team played five games, finishing with one win and four losses. The five games were, in order: a 63:105 loss to Serbia, a 69:89 loss to South Sudan, an 89:107 loss to Puerto Rico, an 83:76 win over Angola, and a 75:96 loss to the Philippines. Consecutive losses, a brief rebound, and another loss form a clear sequence of event milestones, and provide case conditions for comparing fan sentiment across different match nodes. Hupu is a vertical community where Chinese basketball discussion is relatively concentrated; its public match posts, post-match discussions, and rating pages can provide textual material directly related to match milestones.

This paper is positioned as an applied study of basketball public opinion. At present, there are 28 Hupu World Cup event texts that can be traced back one by one and undergo manual semantic review; if these samples were directly divided into training, validation, and test sets, the sample size of individual subsets and categories would be too small, and the resulting machine-learning metrics would lack stability. Therefore, this paper keeps the algorithm benchmark experiment on public Chinese sentiment data and uses the model for auxiliary preliminary screening of basketball texts; the labels actually used to compute event trends are determined by the researcher through manual review in combination with the original basketball context. To enhance the breadth and verifiability of data sources, the study subsequently conducts expanded retrieval of public content on Zhihu, Hupu, and Weibo, establishing a traceable cross-platform source index. The expanded index is used for issue-level cross-verification and future corpus expansion, and is not mixed with the 28 manually reviewed samples to compute the sentiment index.

1.1.2. Research Significance

(1) Theoretical significance. Existing research on sports online public opinion often discusses emotional tendency, dissemination characteristics, or hot issues separately, and rarely examines emotional changes in combination with specific match milestones. Taking five consecutive games as a timeline, this paper compares changes in sentiment direction, evaluation targets, and discussion issues within the same case, so that post-match emotional fluctuations can correspond to real match outcomes, providing a relatively intuitive research approach for the event-based analysis of basketball public opinion.

(2) Practical significance. The operation and communication of basketball events rely heavily on fan feedback. By jointly organizing event milestones, sentiment direction, and evaluation targets, event operators can more quickly identify fans' concentrated reactions to issues such as player performance, coach rotation, controversial calls, and match experience, providing references for post-match communication, content planning, and public-opinion response. After long-term accumulation, the relevant metrics can also be linked with interactive activities and content dissemination effects, providing data support for event operations.

In terms of commercial communication, there is a strong association among teams, players, and sponsor brands. If brand-related discussion is further collected in the future, changes in brand reputation along with event developments can be observed on the basis of the existing sentiment identification and target identification. This paper does not use brand commercial data, so this part only discusses possible directions for transforming the results and is not treated as a verified conclusion of this paper.

In terms of fan culture governance, both rational technical discussion and emotional criticism and personal attacks exist in basketball communities. Distinguishing negative emotion, evaluation targets, and specific issues helps platforms and event organizers identify the problems that genuinely require response, avoiding conflating ordinary post-match complaints with persistent malicious expression. The sentiment model in this paper only serves an auxiliary screening role, and formal judgment still needs to be combined with text context.

1.2. Literature Review

1.2.1. Research on Sports Online Public Opinion

Domestic research on sports online public opinion has gradually shifted from macro dissemination mechanisms to data-driven analysis. Wang Xiaochen and Hu Yuze studied the dissemination and governance of online public opinion under multi-

subject participation using the Beijing Winter Olympics as an example [2]; Wan Xiaohong and Chen Jiabao discussed the characteristics of online public opinion on sports events from the two levels of structural analysis and sentiment analysis [3]; Lei Yueqiu and Li Yuanyuan used Beijing Winter Olympics Weibo texts to examine social media sentiment dissemination in major sports events [4]. These studies indicate that public opinion on major events has clear stages, and both event milestones and disseminating subjects affect discussion content and emotional expression.

Foreign research began earlier in using social media data to observe the relationship between fan emotion and match events. Yu and Wang analyzed real-time tweets during the 2014 World Cup and found that goals scored and conceded significantly change fans' emotional expression [5]; Chang analyzed tweets during the Super Bowl and also found that scoring by one's own team and by the opponent cause positive and negative emotional fluctuations respectively [6]; Gong and Wang, taking the 2019 NBA Finals as an example, further showed a clear correspondence between match results and key events and changes in fan attitudes [7]. This type of research provides a direct basis for organizing sports public-opinion data by match events, but is relatively insufficient with respect to slang, sarcasm, and multi-object evaluations in Chinese basketball communities.

As for Chinese basketball communities, Hupu has been used in sports natural language processing research. The LiveQA dataset constructed by Liu et al. comes from live information of more than 1,670 NBA games on Hupu [8], indicating that Hupu's public sports content has a relatively clear match timeline and event semantics. Unlike LiveQA, which focuses on question-answering tasks, this paper focuses on sentiment and issue changes within the same event cycle; therefore, five World Cup games are used as the event index, and each text is assigned to the corresponding match stage.

1.2.2. Machine-Learning-Based Text Sentiment Analysis

Machine-learning text sentiment analysis usually includes text cleaning, feature representation, model training, and performance evaluation. TF-IDF measures feature importance using term frequency and inverse document frequency, and is a classic text representation method [9]. For short texts in Chinese forums, character n-grams can reduce the impact of segmentation errors on person names, abbreviations, and Internet slang, and are therefore suitable as the basic text features of this paper.

In terms of classification algorithms, KNN completes classification based on neighborhood similarity, and Decision Tree recursively partitions samples using feature conditions; linear SVM is suitable for handling high-dimensional sparse text, and its maximum-margin idea was systematically proposed by Cortes and Vapnik [10]. Bagging reduces model variance through repeated sampling and ensemble voting [11], and AdaBoost gradually combines weak classifiers by increasing the weights of misclassified samples [12]. This paper uses Scikit-learn for model training and evaluation [13].

Public Chinese sentiment data can be used to compare the basic performance of different classification algorithms under unified conditions. The ASAP dataset released by Bu et al. contains 46,730 Chinese restaurant reviews and provides training, validation, and test set partitions [14]. This paper uses the five-star ratings of this dataset to construct a three-class benchmark experiment, whose purpose is only to compare traditional machine-learning models and select an auxiliary screening tool; the test results on restaurant reviews are not treated as the final performance of the basketball public-opinion model.

The language of basketball communities differs markedly from restaurant reviews; in particular, player nicknames, basketball slang, sarcasm, and comments that evaluate multiple targets simultaneously increase the difficulty of automatic classification. Therefore, this paper does not take "cross-domain transfer" as a research goal; instead, after the algorithm benchmark experiment, it separately examines the model's performance in the preliminary screening of basketball texts. Machine prediction serves only as an initial judgment, and the final sentiment labels are still manually reviewed according to the basketball context. Such handling makes machine learning serve practical public-opinion analysis rather than treating performance differences across domains as the main conclusion of the paper.

Beyond sentiment direction, sports public opinion also needs to answer "what fans are discussing." Considering that the number of World Cup texts that can be verified one by one in this paper is small, the main text does not use automatic topic models that require a large corpus; instead, it organizes the main issues through manual reading and keyword induction. This part is case-based issue induction, mainly used to assist in explaining emotional changes.

In summary, existing research has shown that social media sentiment changes with match results and key events, and has also provided a mature

methodological foundation for traditional machine-learning sentiment classification. On this basis, this paper places its research focus on Chinese basketball event applications: on the one hand, comparing the basic performance of traditional classification models on public corpora; on the other hand, organizing Hupu texts by match milestones and analyzing public-opinion changes across the five World Cup games through machine preliminary screening, manual review, and issue induction.

1.3. Research Content and Technical Route

This paper organizes the research according to the logic of "event and text organization — machine-learning benchmark comparison — machine preliminary screening and manual review of basketball texts — event trend analysis — application demonstration." The study focuses on four questions: RQ1, what are the differences in three-class performance among five traditional machine-learning models on a unified Chinese sentiment benchmark; RQ2, what typical misjudgments occur when the selected model is directly used for preliminary screening of basketball texts; RQ3, how the sentiment index, evaluation targets, and discussion issues change across the five World Cup games; and RQ4, how to transform the above analysis process into an auxiliary tool for event public opinion. The research technical route is shown in Figure 1.

(1) Event and text organization. Verify the results of the five games according to FIBA match records; organize pre-match predictions, post-match discussions, review articles, and rating information directly related to the games from Hupu public pages; and manually review the sentiment of 28 traceable event-text units. At the same time, establish a cross-platform public-source index and archive Zhihu answer entries, Hupu public pages, and Weibo historical trending-topic records by event milestone.

(2) Machine-learning benchmark comparison. Construct a three-class experiment using the ASAP public Chinese review data, compare the five models (KNN, Decision Tree, linear SVM, Bagging, and AdaBoost) on character 2-gram and 3-gram TF-IDF features, and focus on Macro-F1 and per-class recognition.

(3) Machine preliminary screening and manual calibration of basketball texts. Apply the relatively stable model from the benchmark experiment to the preliminary screening of 28 Hupu event texts; establish a domain-term list for phenomena such as "战犯" (culprit), "寄了" (done for), player nicknames, sarcasm, and multi-target evaluation; and manually review them in combination with the original page context. At the same time, in combination with the cross-platform public index, manually summarize issues such as player performance, coaching tactics, fundamentals, league competition, and youth training.

(4) Event trend analysis. Taking the five games as event milestones, compute the sentiment index after manual review, compare the synchronized changes between emotion and match results, and use descriptive linear regression to examine the in-sample relationships among event sequence number, point margin, and sentiment index.

(5) Application prototype demonstration. Design functions such as sentiment trend, hot issues, target evaluation, and risk warning according to the event public-opinion analysis process, so that the research results can correspond to practical scenarios such as event review, content operation, and public-opinion communication.

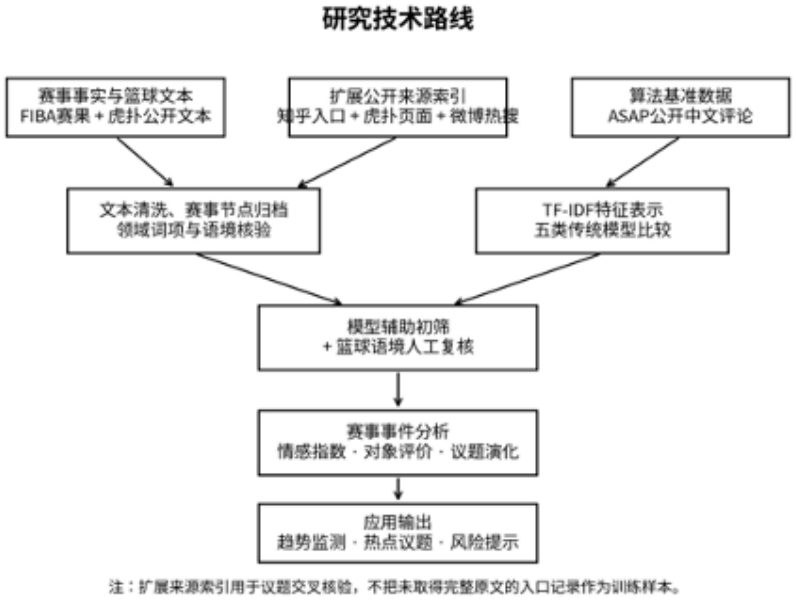


Figure 1 Research Technical Route

1.4. Research Innovations

First, a distinctive research perspective. This paper does not merely count the proportions of positive and negative comments; instead, it maps emotional changes one by one to the five specific games, and simultaneously combines evaluation targets and discussion issues to explain why post-match public opinion changes. For short-format sports events, organizing by match events is more conducive to explaining emotional fluctuations than simply aggregating by calendar date.

Second, methodological improvements in application. In response to the slang, sarcasm, player nicknames, and multi-target evaluations common in basketball forums, this paper adds basketball-domain term checking and manual semantic review after machine classification. Machine learning improves the efficiency of preliminary screening, while manual review handles complex expressions that are difficult for machines to judge, keeping the labels ultimately used for trend analysis consistent with the specific basketball context.

Third, data and application features. This paper organizes official match results, Hupu event texts, and Hupu rating information by the same match milestone, and further establishes a cross-platform public-source index. As of this round of retrieval, the index contains 39 Zhihu answer entries, 19 Hupu public pages or entries, and 13 Weibo event trending-topic records. Different sources are used respectively to verify match results, analyze text sentiment, supplement target evaluation, and cross-verify issues; for indexed records for which complete Chinese source texts have not yet been obtained, this paper does not count them as training samples or in the sentiment index. Based on this traceable data-organization approach, the paper further produces four types of output — sentiment trend, hot issues, target evaluation, and risk warning — providing references for event operation and post-match public-opinion review.

2. Related Theories and Technical Foundations

2.1. Overview of Online Public Opinion Sentiment Analysis

The goal of online public-opinion sentiment analysis is to identify evaluation polarity, emotional intensity, and evaluation targets from user texts, and further aggregate them into a group-level attitude structure. This paper focuses on the three-class polarity task, i.e., dividing texts into negative, neutral, and positive. Compared with binary classification, the three-class scheme retains neutral information such as "explanatory discussion, suggestions, and factual statements," and is suitable for the mixed expressions in sports forums that combine emotional venting with technical review.

2.1.1. Text Preprocessing Techniques

Text preprocessing first removes URLs, HTML tags, consecutive whitespace, and page-template noise, while retaining Chinese, English, numbers, and basic punctuation that can carry tone. For Hupu texts, an additional basketball-domain term list is maintained, including "China men's basketball team, World Cup, Li Kaier (Kyle Anderson), Zhao Rui, Zhou Qi, Hu Jinqiu, Djordjevic, three-pointer, rebound, pick-and-roll, zone defense, turnover," so as to keep evaluation targets intact during manual review and error analysis.

It should be noted that although "Chinese word segmentation" appears in the table of contents of this paper, the core features of the classifier do not rely on traditional word-level segmentation; instead, character-level 2-grams and 3-grams are adopted. The reason is that forum texts contain player abbreviations, Internet expressions, English abbreviations, and irregular punctuation; if forced word segmentation is applied, proper names or colloquial phrases such as "Li Kaier" and "Hu Jinqiu" may be cut apart. Character n-grams still belong to a granularity design of Chinese text preprocessing and segmentation, which reduces the direct impact of custom dictionary quality on classification results while ensuring the robustness of short texts. The word-level domain list mainly serves interpretability checking rather than replacing the model's primary features.

2.1.2. Text Vectorization

TF-IDF determines weights through the extent to which a term appears in the current document and the scarcity of that term across the entire corpus. Let TF(t, d) denote the term frequency of character fragment t in document d, df(t) the number of documents containing t, and N the total number of documents; its basic form is:

$$TF - IDF(t, d) = TF(t, d) \times log((N + 1)/(df(t) + 1))$$

Combining character n-grams with TF-IDF produces high-dimensional sparse text vectors. The specific n-gram range, minimum document frequency, number of features, and dimensionality-reduction parameters are experimental settings, and are uniformly described in Chapter 3.

2.1.3. Basketball Domain Language Adaptation and Topic Identification

Basketball online texts are highly context-dependent. Fans often express attitudes through player nicknames, tactical abbreviations, slang, and sarcasm;

for example, terms such as "战犯" (culprit) and "寄了" (done for) usually carry an obvious negative tendency, while some nicknames may themselves simultaneously carry teasing and approving meanings. Technical terms such as "zone defense," "pick-and-roll," and "rotation" usually do not directly indicate positive or negative sentiment, and need to be judged in combination with evaluative words and match results.

Therefore, at the application stage for basketball texts, this paper retains player names, teams, tactics, and event-related terms, and lists texts containing slang, sarcasm, or multiple evaluation targets as key review targets. This term list does not directly determine the final sentiment labels; it only helps researchers locate texts that the machine model is prone to misjudge.

2.2. Principles of Machine Learning Models

2.2.1. K-Nearest Neighbors (KNN)

KNN is an instance-based supervised learning method. For a sample to be classified, it computes the distance or similarity to the training samples, selects the K nearest neighbors, and completes classification according to the neighbors' labels. Text vectors often use cosine similarity to measure closeness in direction, whose expression is:

$$cos(x,y) = (x \cdot y)/(||x|| \cdot ||y||)$$

2.2.2. Decision Tree

Decision Tree recursively selects features and partition conditions to gradually assign samples to different leaf nodes. Classification trees often use the Gini index to measure node purity; let p_k denote the proportion of class-k samples in a node, then:

$$Gini(D) = 1 - \Sigma p_k^2$$

2.2.3. Support Vector Machine (SVM)

SVM completes classification by finding the maximum-margin hyperplane; its core is to balance the classification margin against the training error. Linear SVM is especially suitable for high-dimensional sparse text features [10]. For multi-class tasks, multiple linear discriminant functions can be trained via the one-vs-rest strategy. The specific penalty parameter and class weights are given in the experimental settings of Chapter 3.

2.2.4. Ensemble Learning

Bagging trains multiple base learners through Bootstrap resampling of the training samples and merges their results by voting to reduce the variance of a single model [11]. The specific base learner and number of estimators are described in the experimental part of Chapter 3.

AdaBoost increases the weights of misclassified samples during iteration, making subsequent weak classifiers pay more attention to hard-to-classify samples, and combines multiple weak classifiers by weighting [12]. The specific weak learner, number of iterations, and learning rate are described in the experimental part of Chapter 3.

2.2.5. Linear Regression

Linear regression is used to describe the linear relationship between explanatory variable x and continuous dependent variable y; its basic form is:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Chapter 4 uses the event sequence number and China's point margin as explanatory variables, and the manually reviewed sentiment index as the dependent variable for descriptive regression. Since only five games are included, the regression results are used only to compare in-sample relationships, without causal inference.

2.3. Time-Series Trend Analysis Methods

2.3.1. Quantification of Sentiment Scores

To facilitate comparing the overall sentiment direction at different match stages, this paper constructs a "match sentiment index." For a match window t, let P_t denote the number of positive texts, U_t the number of neutral texts, N_t the number of negative texts, and M_t = P_t + U_t + N_t the total number of samples; then:

$$S_t = (P_t - N_t)/M_t,\ S_t \in [-1,1]$$

The closer the index is to 1, the higher the proportion of positive texts in the sample; the closer to -1, the higher the proportion of negative texts; a value near 0 indicates that positive and negative tendencies offset each other or that neutral discussion is more common. This metric is a custom statistic used for case comparison in this paper, and does not represent the overall attitude of all platform users.

2.3.2. Time-Series Decomposition and Trend Analysis

The Chinese team had only five games in the World Cup, which does not meet the conditions for seasonal time-series decomposition. Therefore, this paper does not use the traditional trend–seasonal–random decomposition; instead, it treats the five games as consecutive event milestones and compares the sentiment-index changes at each milestone.

In addition to the event sequence number, this paper further uses the point margin to represent the strength of the match result. The event sequence number is used to observe whether sentiment shows a stable change as the games progress, while the point margin is used to observe whether sentiment changes in

synchrony with win/loss and score differences; both are used only for descriptive comparison.

2.4. Evaluation Metrics for Classification and Trend Prediction

2.4.1. Evaluation Metrics for Sentiment Classification

The classification models are evaluated using Accuracy, Precision, Recall, and F1-score, defined as follows:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

$$Precision = TP/(TP + FP)$$

$$Recall = TP/(TP + FN)$$

$$F1 = 2 \times Precision \times Recall/(Precision + Recall)$$

Because the ASAP test set has an imbalanced class distribution, this paper focuses on using Macro-F1 to compare the balanced recognition ability across the three sentiment classes, and simultaneously uses Accuracy and the confusion matrix to analyze specific misjudgments.

2.4.2. Evaluation Metrics for Regression Trend Prediction

Regression analysis uses mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2) to evaluate the fitting results, calculated as follows:

$$MAE = (1/n)\Sigma \mid y_i - \hat{y}_i \mid$$

$$RMSE = \sqrt{[(1/n)\Sigma(y_i - \hat{y}_i)^2]}$$

$$R^2 = 1 - [\Sigma(y_i - \hat{y}_i)^2/\Sigma(y_i - \bar{y})^2]$$

R^2 reflects the extent to which the model explains sample variation, but it is easily affected by individual observations when the sample size is very small; therefore, Chapter 4 also reports the correlation coefficient, MAE, and RMSE, and limits the scope of interpretation.

3. Empirical Study on Machine-Learning-Based Sentiment Classification of Chinese Basketball Public Opinion

3.1. Data Collection and Preprocessing

3.1.1. Data Sources

This paper uses three types of data. The first is the ASAP public Chinese review dataset, used for the benchmark comparison of five traditional machine-learning algorithms; the second is Hupu public event texts related to the Chinese men's basketball team at the 2023 World Cup, used for basketball public-opinion case analysis and manual review; the third is a cross-platform public-source index, used for expanded retrieval and issue cross-verification. In addition, official FIBA match results are used to verify match facts, and the Hupu World Cup rating page for the Chinese team is used to supplement target-level evaluation of players and coaches. The three types of data serve different tasks: ASAP does not participate in the computation of the World Cup sentiment index; the 28 Hupu event texts are the direct textual basis for the trend analysis in Chapter 4; and the cross-platform index is used to verify whether issues recur across different communities and dissemination arenas. The official results of the five games are shown in Table 1, and the composition of the expanded public sources is shown in Table 2 and Figure 2.

Hupu event texts have been retrieved and parsed from public mobile pages since August 10, 2026. Around the five World Cup games, the study selects pre-match predictions, post-match discussions, review articles, and highlighted replies directly related to the Chinese men's basketball team, and records them in the format "page ID — publication time — match stage — text unit." After cleaning and semantic segmentation, a total of 28 event-text units that can be traced back one by one are retained: 3 pre-match, 3 for Serbia, 6 for South Sudan, 4 for Puerto Rico, 5 for Angola, 6 for the Philippines, and 1 post-match summary. Among them, the core samples corresponding to the five games total 24. These 28 texts are not used to train the classifier, but for checking the machine preliminary-screening effect, manual review, and event trend analysis.

Table 1 Official Results of the Five Games of the Chinese Men's Basketball Team at the 2023 World Cup

Date	Opponent	China's Score	Opponent's Score	Point Margin	Result
2023-08-26	Serbia	63	105	-42	Loss
2023-08-28	South Sudan	69	89	-20	Loss
2023-08-30	Puerto Rico	89	107	-18	Loss
2023-08-31	Angola	83	76	7	Win
2023-09-02	Philippines	75	96	-21	Loss

Table 1 lists the dates, scores, point margins, and win/loss results of the five games, all verified against the official FIBA 2023 World Cup match records.

Table 2 Composition and Use of the Cross-Platform Public-Source Index

Source	Count	Data Form	Use in This Paper
Zhihu	39 entries	Answer page / question page index	Cross-platform issue comparison; entry for future Chinese corpus expansion
Hupu	19 pages or entries	Event posts, highlighted replies, news pages, rating pages	Original-context verification, object evaluation, source of expanded cases
Weibo	13 trending-topic records	Event-related topics, time on the trending list, and popularity values	Auxiliary verification of event popularity and controversy milestones

Table 2 shows the source composition of the expanded retrieval, and Figure 2 visually displays the quantities of the three types of public sources. The study organizes a total of 58 verifiable public-source entries, of which 39 come from Zhihu and 19 from Hupu; in addition, 13 Weibo historical trending-topic records directly related to the five games are recorded. It must be emphasized that a "source entry" is not equivalent to a comment-by-comment sample: some Zhihu indices can currently only stably obtain answer entries or translation indices, and Weibo trending-topic records do not contain complete comment text, so these records do not enter machine-learning training or the sentiment-index statistics of Chapter 4. The relevant URLs, event milestones, preliminary-screening summaries, issues, and source-text availability notes are archived separately. The answer counts displayed on public question pages are used only to judge future expansion potential; answers that have not been fetched, cleaned, and deduplicated one by one are not counted in the actual sample size.

The original ASAP data is divided into a training set of 36,850 samples, a validation set of 4,940 samples, and a test set of 4,940 samples. This paper constructs three-class labels based on the five-star ratings: 1–2 stars are negative, 3 stars are neutral, and 4–5 stars are positive. To reduce class imbalance during training, 2,000 samples are randomly drawn from each of the negative, neutral, and positive classes in the training set, forming a balanced training set of 6,000 samples; the test set of 4,940 samples keeps the original distribution. The original validation set remains independent and is not merged with the training or test sets. Random downsampling discards some samples from the majority classes, so this experiment is used to compare the relative performance of five traditional algorithms and to select an auxiliary preliminary-screening model, not as a final performance conclusion for a basketball-specific classifier.

The cross-platform source index is stored separately as a project data file (2023 China Men's Basketball Public Opinion – Public Index Collection – Continuous Expansion Edition.xlsx), with fields including platform, event milestone, date, source type, source URL, preliminary-screening summary, preliminary-screening sentiment, issue, confidence, and remarks on source-text availability. The "preliminary-screening sentiment" and "issue" fields are completed with AI assistance and are used only for retrieval management, sample filtering, and manual-review prioritization; they are neither manual annotation results nor used to compute the model test metrics, sentiment proportions, or regression variables in this paper.

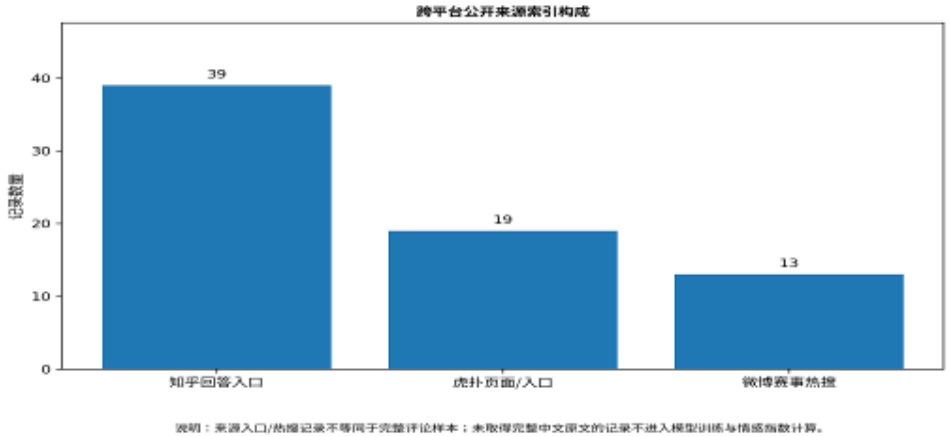


Figure 2 Composition of the Cross-Platform Public-Source Index

3.1.2. Data Cleaning and Chinese Word Segmentation

Text cleaning includes removing URLs, HTML tags, and consecutive whitespace, removing irrelevant information such as page navigation, templated prompts, and image placeholders, and segmenting text by the unit of "a statement with an independent sentiment theme." If a main post contains both factual statements and evaluative conclusions, only the sentence segments directly related to the research topic are retained in the sentiment unit, and an entire long article is not treated as a single unexplainable over-long sample. The cleaning and segmentation rules are formulated uniformly before annotation and kept consistent throughout the annotation process.

Chinese segmentation adopts a "two-layer strategy." The first layer is character-level 2-gram and 3-gram segmentation for model computation, requiring no external segmenter; the second layer is the retention of domain terms for manual auditing and error analysis, identifying player names, teams, and tactical terms as complete units. For example, "李凯尔表现太棒了" (Li Kaier played amazingly well) produces local fragments such as "李凯," "凯尔," "表现," and "太棒" at the character-feature layer, while at the manual-review layer "李凯尔" is still treated as a complete evaluation target. This strategy balances model robustness with semantic interpretability.

3.1.3. Construction of Sentiment Labels

Sentiment labels are determined by manual annotation. The judgment principles are: a text is labeled positive when its core semantics clearly praise the team, players, or match improvement, or express the joy of victory; negative when it clearly expresses disappointment, criticism, pessimism, negation, or strong dissatisfaction with technique and tactics; and neutral when it mainly consists of factual statements, technical suggestions, or rational explanations without an obvious praising or disparaging direction. For texts containing mixed positive and negative semantics, the main argument, evaluation target, and concluding sentence serve as the basis for labeling.

The 28 Hupu event-text units are manually reviewed one by one by the researcher according to the above unified rules. For texts involving basketball slang, sarcasm, multiple targets, and mixed sentiment, the original page context, evaluation target, and concluding sentence are examined simultaneously during review. Due to project-cycle and sample-size constraints, this study does not organize dual independent annotation, and therefore does not report consistency metrics such as Cohen's Kappa; this limitation is stated truthfully in Section 5.3.

Because the scale of the 28 Hupu samples is small, this paper does not further split them into training, validation, and test sets. If split by common ratios, each subset and some sentiment categories would retain only a few samples, and model performance would be highly dependent on individual texts. Therefore, the machine-learning benchmark evaluation uses ASAP's official data partition, and the Hupu samples serve only for application testing and case analysis.

3.2. Feature Engineering and Vectorization

TfidfVectorizer is fitted on the 6,000 balanced training texts with settings analyzer="char", ngram_range=(2,3), min_df=3, max_features=6000, sublinear_tf=True, and norm="l2". The vectorizer is fitted only on the training samples, and both the validation and test sets are transformed through the same vectorizer. Random sampling and models involving random processes uniformly set random_state=42.

For Bagging and AdaBoost, to reduce the computational overhead of tree models directly handling the 6,000-dimensional sparse matrix, TruncatedSVD is applied after TF-IDF to reduce dimensionality to 50; the 50 components have a cumulative explained variance of about 10.26%. This compression loses some sparse-feature information, so the ensemble models' results are affected by both the classifier and the dimensionality reduction.

KNN, Decision Tree, and SVM use the same TF-IDF representation, while Bagging and AdaBoost use the same 50-dimensional SVD representation. All models learn on the same 6,000 balanced training samples and uniformly report results on the ASAP test set, to ensure consistent conditions for model comparison.

3.3. Empirical Study Based on Classification Models

The five-model experiment in this section is used to compare traditional classification algorithms and select an auxiliary preliminary-screening tool; the test metrics come from the official ASAP test set and are not treated as the final performance of a basketball review classifier. The basketball event conclusions still follow the manually reviewed labels described in Section 3.1.3.

3.3.1. KNN Model Construction and Tuning

KNN sets K=11 and uses cosine distance, distance weights, and brute-force search. On the ASAP test set, Accuracy is 0.5850, Macro-Precision is 0.4764, Macro-Recall is 0.5845, and Macro-F1 is 0.4817.

3.3.2. Decision Tree Model Construction and Tuning

Decision Tree sets max_depth=30 and min_samples_leaf=5, and uses balanced class weights. On the ASAP test set, Accuracy is 0.5113, Macro-Precision is 0.4278, Macro-Recall is 0.5184, and Macro-F1 is 0.4111.

3.3.3. SVM Model Construction and Tuning

LinearSVC sets C=1.0 and class_weight="balanced". On the ASAP test set, Accuracy is 0.6688, Macro-Precision is 0.5280, Macro-Recall is 0.6502, and Macro-F1 is 0.5531. In the per-class results, the negative class has Precision=0.4219, Recall=0.7988, and F1=0.5521; the neutral class has Precision=0.2381, Recall=0.4547, and F1=0.3126; and the positive class has Precision=0.9239, Recall=0.6970, and F1=0.7946.

SVM's negative Recall reaches 0.7988, indicating that it can identify many clearly negative samples, but its negative Precision is only 0.4219, so it still misclassifies some neutral or complex texts as negative. At the same time, the neutral class has an F1 of only 0.3126, making it the hardest of the three classes to identify. Therefore, SVM is suitable for preliminary screening, but is not suitable for directly generating basketball public-opinion conclusions without manual review.

3.3.4. Construction of Two Ensemble Models (Bagging and Boosting)

Bagging uses decision trees with maximum depth 12 and minimum leaf-node sample size 5 as base learners, sets 8 estimators, max_samples=0.7, and max_features=0.9; its test Accuracy is 0.6308 and Macro-F1 is 0.5188. AdaBoost uses shallow decision trees of depth 2, sets n_estimators=30 and learning_rate=0.6; its test Accuracy is 0.6109 and Macro-F1 is 0.5421.

Table 3 Performance Comparison of Five Machine-Learning Models on the ASAP Benchmark Test Set

Model	Accuracy	Macro-P	Macro-R	Macro-F1	Weighted-F1
KNN	0.5850	0.4764	0.5845	0.4817	0.6371
Decision Tree	0.5113	0.4278	0.5184	0.4111	0.5703
SVM	0.6688	0.5280	0.6502	0.5531	0.7080
Bagging	0.6308	0.5053	0.6079	0.5188	0.6766
AdaBoost	0.6109	0.5429	0.6395	0.5421	0.6625

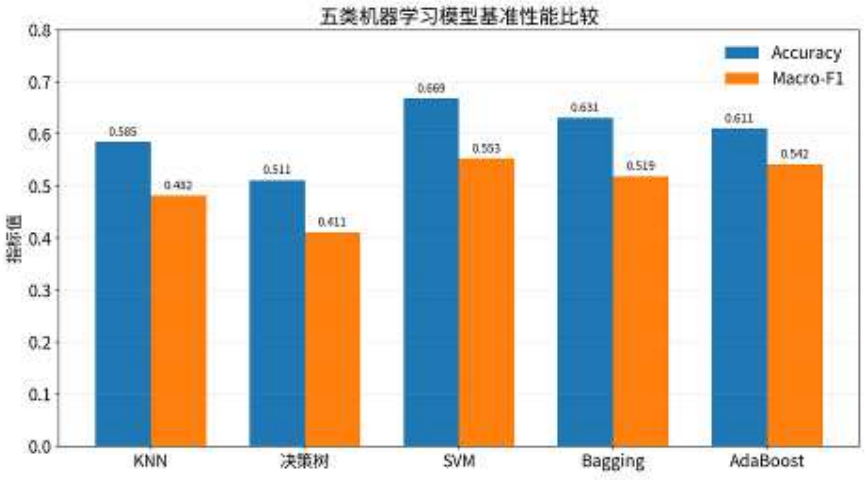


Figure 3 Comparison of Benchmark Performance of Five Machine-Learning Models



Figure 4 Confusion Matrix of SVM on the Benchmark Test Set

3.3.5. Model Performance Comparison and Error Case Analysis

The performance of the five models on the ASAP test set is shown in Table 3, and the intuitive comparison of Accuracy and Macro-F1 is shown in Figure 3. If Macro-F1 is used to measure the balanced recognition ability across the three sentiment classes, SVM performs best among the five models: its Macro-F1 is 0.0714 higher than KNN, 0.1420 higher than Decision Tree, 0.0343 higher than Bagging, and 0.0110 higher than AdaBoost. Because the test set has an imbalanced class distribution, this paper uses Macro-F1 as the primary model-selection metric, and judges whether a model is suitable for auxiliary screening in combination with Accuracy and per-class metrics.

The confusion matrix in Figure 4 further illustrates the error structure of SVM: of 338 negative samples, 270 are correctly identified; of 717 neutral samples, 326 are correctly identified, among which 180 are misclassified as negative and 211 as positive; of 3,885 positive samples, 2,708 are correctly identified, with another 987 misclassified as neutral. Neutral comments usually contain factual statements, suggestions, or weak evaluations, and the class boundary is relatively blurry.

Subsequently, SVM is directly applied to the 28 Hupu World Cup event texts described earlier and compared with the manually reviewed labels. The machine preliminary screening correctly judges 13 texts, with an accuracy of 13/28 = 46.43%. Misjudgments mainly occur in basketball slang such as "寄了" (done for) and "战犯" (culprit), sarcastic expressions, player nicknames, and cases where a single text evaluates both a player and the coach. This result is only used to assess the usability of machine preliminary screening on basketball texts, and the ASAP model is not regarded as an already-completed basketball-specific sentiment classifier.

Based on the above results, this paper adopts the processing strategy of "machine preliminary screening + manual review": SVM first gives an initial judgment of positive, neutral, or negative, and the researcher then returns to the original post context to check slang, sarcasm, and multi-target evaluations one by one. The match sentiment indices in Chapter 4 are all computed from the manually reviewed labels, thereby avoiding treating machine misjudgments directly as real fan public opinion.

3.4. Basketball Domain Adaptation and Issue Structure Analysis

3.4.1. Basketball Domain Terms and Irony Handling

According to the misjudgment situations in the 28 World Cup texts, this paper divides the basketball terms requiring key review into five categories: responsibility attribution, outcome expectation, individual nickname, technique and tactics, and long-term issues. For example, "战犯" (culprit) and "拉路" (underperforming) usually carry obvious negative evaluations; "寄了" (done for) and "没戏" (no hope) express pessimistic expectations; technical terms such as "pick-and-roll," "zone defense," and "rotation" usually have no fixed sentiment; and player nicknames must be judged in combination with context. The specific handling is shown in Table 4.

Table 4 Basketball Public-Opinion Domain Terms and Handling Strategies

Type	Example Terms	Main Risk	Handling Strategy
Responsibility Attribution	战犯 (culprit), 背锅 (scapegoat), 拉路 (underperforming)	High negative intensity, clear target	Retain the evaluation target and mark as a high-risk negative candidate
Outcome Expectation	寄了 (done for), 没戏 (no hope), 出线无望 (no chance to advance)	Colloquial, weak literal sentiment	Judge in combination with event milestones and negation words
Nickname Teasing	大锤 (Big Hammer), 人类蠕动精华 (the essence of human slowness)	Teasing and praise coexist	Never determine polarity from the term alone
Technique and Tactics	pick-and-roll, zone defense, rotation, open layup	Most terms are neutral in themselves	Identify sentiment in combination with evaluative and outcome words
Institutional Issues	CBA, youth training, separation of sports and education, development system	Often spill over from a single-game loss	Treat as topic labels rather than pure sentiment words

The terms in Table 4 do not serve as automatic rules to directly determine positive or negative labels; instead, they help researchers quickly locate texts that the machine model is prone to misjudge. During manual review, the evaluation target, negation words, contrastive relations, and match context are checked simultaneously before determining the final sentiment category.

3.4.2. Exploratory Issue Structure and Cross-Platform Comparison

This paper does not run an automatic topic model on the 28 World Cup texts; instead, it organizes the discussion content through manual reading and keyword induction. Based on the Hupu texts at each match stage, the main issues can be summarized into five categories: single-game athletic performance, individual player evaluation, coaching and tactical choices, CBA league competition and fundamentals, and youth training and the development system. The typical keywords and occurrence stages of each issue are shown in Table 5. To reduce the interpretive bias caused by a single-platform case, the study further uses the cross-platform source index shown in Table 2 for non-metric cross-verification: it only judges whether similar issues recur in other public arenas, and does not convert the index's preliminary-screening labels into sentiment proportions.

Table 5 Main Issues in World Cup Public Opinion and Their Evolutionary Implications

Issue	Typical Keywords	Main Stages of Occurrence	Public-Opinion Implication
Single-Game Athletic Performance	score, shooting, rebounds, turnovers	Appeared in all five games	Immediate feedback on match outcomes
Individual Player Evaluation	Li Kaier (Kyle Anderson), Zhao Rui, Hu Jinqiu, Wang Zhelin	Prominent after the Angola and Philippines games	Divergence between overall sentiment and individual evaluation
Coaching and Tactics	rotation, pick-and-roll, zone defense, timeouts, foreign coach	Rose after consecutive losses	Responsibility attribution expands from players to the decision-making level
League and Fundamentals	CBA, shooting, intensity, open layup	Prominent after the South Sudan and Philippines games	Spills over from national-team performance to league quality
Youth Training and Development System	youth training, separation of sports and education, talent development	Prominent in the late and summary stages of the event	Short-term failure interpreted as a long-term structural problem

From the case texts of this paper, the discussion focus is characterized by a gradual expansion from "match outcome — players — coaching tactics — long-term development issues." The early group stage centers more on scores, shooting, and player performance; after consecutive losses, discussion of rotation, fundamentals, and league intensity increases; after the loss to the Philippines, some posts further discuss youth training and the development system. The expanded retrieval provides similar issue clues: discussions such as coaching personnel selection, CBA intensity, fundamentals, talent selection, and youth training repeatedly appear in the relevant Zhihu answer entries; Weibo historical trending topics include controversial subjects such as "why not play Hu Jinqiu" and "Djordjevic's tactics" at the South Sudan and Puerto Rico match milestones. Because the data forms of these expanded records are not uniform, this paper treats them only as auxiliary evidence for the existence of issues, without cross-platform sentiment-proportion comparison.

4. Time-Series-Based Trend Analysis of Chinese Basketball Public Opinion

4.1. Quantification of Public Opinion Trends

Chapter 4 uses the manually reviewed labels of the 28 Hupu event-text units, including 3 pre-match and 1 post-match summary, with the core samples corresponding to the five games totaling 24. For each game, the numbers of positive, neutral, and negative texts are counted separately, and the match sentiment index defined in Section 2.3.1 is computed. The sample structure of the five games is shown in Table 6.

Table 6 Manually Reviewed Sentiment Structure of the Five Games

Match Window	Sample Count	Positive	Neutral	Negative	Sentiment Index	Point Margin
Serbia	3	0	1	2	-0.667	-42
South Sudan	6	0	3	3	-0.500	-20
Puerto Rico	4	0	0	4	-1.000	-18
Angola	5	4	0	1	0.600	7
Philippines	6	0	1	5	-0.833	-21

As can be seen from Table 6, the sentiment indices in the first three losing games are all negative, reaching -1.000 for the Puerto Rico game; after the first win over Angola, the index rises to 0.600; and after the final loss to the Philippines, it falls to -0.833. This result indicates that the emotional changes in this paper's samples are clearly synchronized with match results. To intuitively compare the sentiment structure and event-sequence changes across the five games, Figure 5 and Figure 6 are drawn respectively.

4.2. Decomposition of Public Opinion Trends Across Key Match Periods

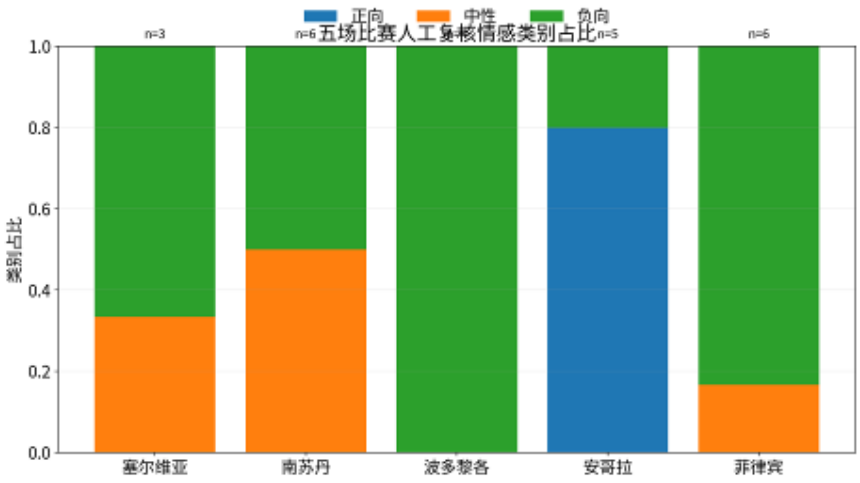


Figure 5 Proportions of Sentiment Categories Across the Five Games

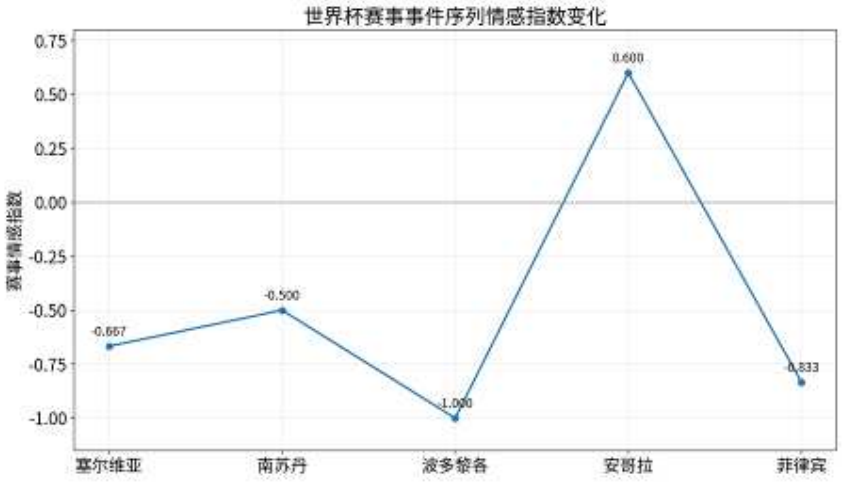


Figure 6 Changes in the Sentiment Index over the World Cup Event Sequence

The first stage is pre-match expectation and the impact of the opening game. Pre-match discussion contains both optimistic and pessimistic judgments; after the 63:105 loss to Serbia, among the 3 game-related texts, 0 are positive, 1 is neutral, and 2 are negative, with a sentiment index of -0.667. Because the opponent is relatively strong, some discussion still takes the strength gap as the main explanation, retaining a certain space for neutral analysis.

The second stage is the consecutive losses to South Sudan and Puerto Rico. After the 69:89 loss to South Sudan, among the 6 texts, 3 are neutral and 3 are negative, with a sentiment index of -0.500; discussion begins to extend from single-game technique and tactics to rotation, shooting, and fundamentals. Subsequently, after the 89:107 loss to Puerto Rico and a three-game losing streak in the group stage, all 4 related texts are reviewed as negative, and the sentiment index drops to -1.000.

The third stage is the emotional recovery after the first win over Angola. After China's 83:76 win over Angola, among the 5 event texts, 4 are positive and 1 is negative, and the sentiment index rises to 0.600, the only positive value across the five games. The performances of players such as Li Kaier (Kyle Anderson) and Hu Jinqiu receive relatively many positive evaluations.

The fourth stage is the renewed turn to negative after the loss to the Philippines. After the 75:96 loss, among the 6 samples, 5 are negative and 1 is neutral, and the sentiment index falls to -0.833. Discussion further extends from player and rotation issues to shooting fundamentals, league development, and the youth-training system.

The Hupu rating page provides target-level supplementary information. Poor overall results do not mean that all players are evaluated identically: Li Kaier (Kyle Anderson) scores 9.8, Zhao Rui 9.6, and Hu Jinqiu 9.6, while some players and coaches receive markedly lower ratings. The related target ratings and participation volume are shown in Figure 7. This indicates that event public-opinion analysis should distinguish between the evaluation of the team's overall performance and the evaluation of specific players and coaches.

Combining the issue induction in Table 5, it can be seen that sentiment direction and discussion content are not fully synchronized. When Angola was defeated, positive discussion concentrated mainly on individual player performance; after the Philippines loss, negative discussion involved more about coaching, fundamentals, the league, and youth training. Although both are negative public opinion, criticism of a single-game mistake and questioning of the long-term development system differ in how they should actually be responded to. The cross-platform source index shown in Table 2 also retrieves similar coaching, league, and development issues; however, because the expanded index is not a complete review corpus in a uniform format, all sentiment proportions, sentiment indices, and regression results in Chapter 4 still use only the 24 match-period samples among the 28 manually reviewed Hupu texts, avoiding the mixing of different data specifications.

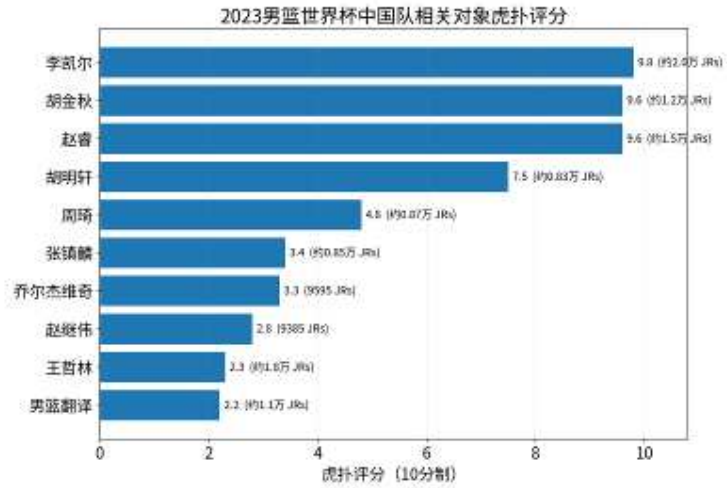


Figure 7 Hupu Ratings and Participation Volume for China-Related Subjects at the World Cup

4.3. Linear-Regression-Based Trend Prediction and Robustness Check

First, the event sequence number is used as the explanatory variable. Let x = 1, 2, 3, 4, 5 correspond in order to Serbia, South Sudan, Puerto Rico, Angola, and the Philippines, and let y be the manually reviewed sentiment index. The fitting result is:

$$y = -0.710 + 0.0767x$$

This model has R^2=0.0368, MAE=0.424, and RMSE=0.555, indicating that the sentiment indices of the five games do not show a stable linear upward or downward trend with the event sequence number.

Second, the point margin is used as the explanatory variable. The point margins of the five games are -42, -20, -18, +7, and -21 in sequence, and the Pearson correlation coefficient between point margin and manually reviewed sentiment index is about 0.741. Simple linear regression yields:

$$情感指数 = 0.0259 + 0.0269 \times 净胜分$$

This model has R^2=0.549, MAE=0.334, and RMSE=0.380, and fits the five-game sample of this tournament better than the event-sequence-number model. However, with n=5, the statistical inference power is very weak, and the two-sided p-value of the correlation test is about 0.152; therefore, this cannot be used to conclude that a stable and significant relationship exists between point margin and sentiment index, still less to make a causal interpretation. The in-sample relationship between point margin and sentiment index is shown in Figure 8.

Sensitivity checking shows that this correlation is relatively sensitive to the only win, against Angola: after removing Angola, the correlation coefficient between point margin and sentiment index for the remaining 4 games is about -0.309. Therefore, the relationship shown in Figure 8 can only be treated as a descriptive phenomenon in this World Cup case, and cannot be extrapolated as a stable statistical law.

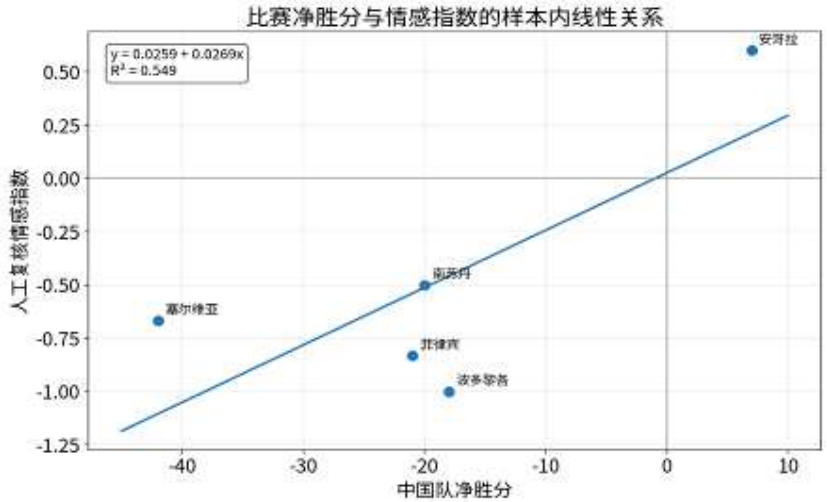


Figure 8 In-Sample Linear Relationship Between Point Margin and Manually Reviewed Sentiment Index

Combining the two sets of regressions, the event sequence number itself can hardly explain emotional changes, while win/loss and score difference show more obvious synchronized changes with sentiment direction in this case. Since there are only 5 observation points, this paper does not predict a sixth game, nor does it extrapolate the regression results beyond this World Cup.

5. Conclusions and Recommendations

5.1. Research Conclusions

Taking the five games of the Chinese men's basketball team at the 2023 FIBA Basketball World Cup as a case, this paper uses traditional machine learning to complete a sentiment-classification benchmark experiment, and combines machine preliminary screening, manual review, the cross-platform public index, and event analysis of Hupu event texts to examine changes in fan sentiment and discussion content during the World Cup. The study insists on distinguishing the data specifications of "complete text samples," "public source entries," and "trending-topic records," and does not count content that has been retrieved but whose Chinese source texts have not yet been obtained one by one as training samples. The main conclusions are as follows.

(1) In the ASAP benchmark experiment, SVM achieves an Accuracy of 0.6688 and a Macro-F1 of 0.5531, the highest Macro-F1 among the five models. SVM's Recall on negative samples is 0.7988, but the neutral class has an F1 of only 0.3126. This result is used to select an auxiliary preliminary-screening tool for basketball texts, and is not equivalent to the final classification performance on basketball corpora.

(2) When SVM is directly applied to the 28 Hupu World Cup event texts, the machine preliminary-screening accuracy is 46.43%. The main misjudgments are related to basketball slang, sarcasm, nicknames, and multi-target evaluation, so the final event labels adopt the manually reviewed results. The role of machine learning in this paper is mainly to improve preliminary-screening efficiency rather than to replace human judgment.

(3) The manually reviewed sentiment indices of the five games are -0.667, -0.500, -1.000, 0.600, and -0.833 in sequence. Sentiment turns from negative to positive after the win over Angola and turns negative again after the loss to the Philippines, indicating that public opinion in short-format events changes rapidly with match results.

(4) The correlation coefficient between point margin and sentiment index across the five games is about 0.741 with $R^2=0.549$, while the event-sequence-number model has an R^2 of only 0.0368; however, $p\approx0.152$, and the correlation direction changes after removing Angola. Therefore, this result can only indicate a certain synchronization between match results and emotional changes in this case, and cannot serve as a stable law or a causal conclusion.

(5) Fans' attitudes toward different evaluation targets differ markedly. When the overall results are poor, some players still receive high ratings; meanwhile, post-match discussion gradually extends from scores and player performance to issues such as coaching tactics, fundamentals, CBA competition, and youth training. This round of expanded retrieval establishes 39 Zhihu answer entries, 19 Hupu public pages or entries, and 13 Weibo event trending-topic records; the cross-platform clues are strongly consistent with the Hupu cases on issues such as coaching personnel selection, league intensity, fundamentals, and youth training, but because the data forms differ, this paper does not further compare sentiment proportions across platforms.

(6) Based on the data-processing workflow of this paper, a public-opinion analysis prototype for event operation can be formed, outputting sentiment trends, hot issues, target evaluation, and risk warnings. This prototype is positioned as decision support, and manual review is still required for formal use.

5.2. Recommendations

First, establish a real-time monitoring mechanism of "event milestones + target nodes" for event operation. For short-format events such as the World Cup and the Asian Cup, the 24 hours before a match, the 2 hours after a match ends, and the 24 hours after a match can be set as key event windows to separately count the sentiment direction and discussion heat of teams, coaches, core players, substitute players, referees, and technique-and-tactics issues. When certain calls, rotations, or player performances trigger a concentrated rise in negative sentiment, operators can prioritize providing rule explanations, technical data, and post-match explanations, reducing the emotional spread caused by information vacuum. After long-term accumulation, public-opinion feedback can also be linked with event scheduling, interactive activities, content dissemination, and the viewing experience, forming an auxiliary decision-making system that upgrades from experience-based operation to data-driven operation.

Second, for clubs and event communication, players, coaches, and match events can be treated as key monitoring targets. For sudden controversies, both the sentiment direction and the specific issues should be examined simultaneously, avoiding judgments based solely on the negative proportion. If brand-related data is obtained in the future, the relationship between player events and brand reputation can be further analyzed.

Third, for fan culture governance, personal attacks, labeling expressions, and persistent extreme negative comments can be treated as key review targets; technical review, rational discussion, and constructive suggestions should be distinguished from general negative sentiment. The sentiment model can only be used for preliminary screening and cannot replace the platform's manual moderation.

Fourth, the processing workflow of this paper can be made into a visual application prototype. The input side consists of public event texts, interaction data, and event-milestone information; the analysis side completes cleaning, machine preliminary screening, manual review, issue induction, and event summarization; and the output side generates sentiment trends, hot issues, target evaluation, and risk warnings. The specific functions are shown in Table 7, and the overall functional framework is shown in Figure 9.

Fifth, after the data volume expands in the future, the simple positive, neutral, and negative classification can be further extended to target-level and aspect-level sentiment analysis, for example distinguishing compound views such as "approving the players' performance but criticizing the coach's rotation."

Table 7 Functional Modules and Outputs of the Application Prototype

Module	Core Input	Core Output	Main Users
Real-Time Public-Opinion Monitoring	Public event texts, engagement volume, event milestones	Popularity curves, negative spikes, target rankings	Event organizers, media
Sentiment Trend Analysis	Sentiment labels, targets, time	Positive/neutral/negative proportions, sentiment index, polarization signals	Clubs, governing bodies
Hot-Issue Identification	Topic words, entities, co-occurrence relations	Issue ranking, issue migration, target-issue matrix	Media, brand parties
Risk Warning and Push	Negative intensity, growth rate, sensitive issues	Tiered warnings, event summaries, recommended handling	Event organizers, brands, platforms
Report Output	All of the above results	Daily reports, weekly reports, special-topic reports	Managers and researchers

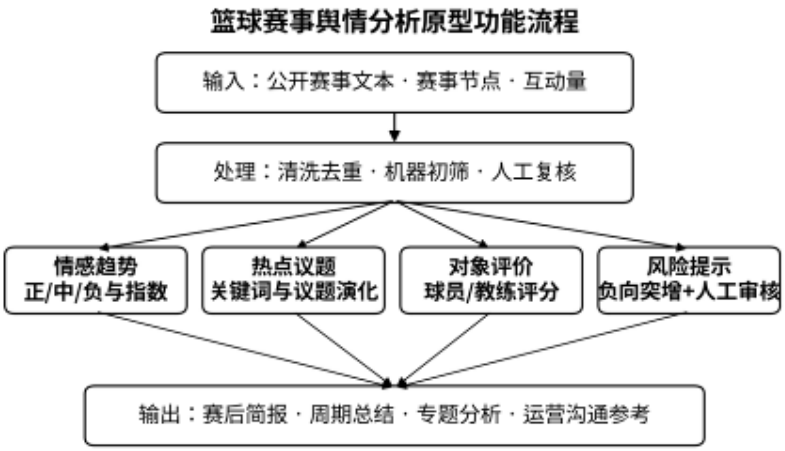


Figure 9 Functional Framework of the Chinese Basketball Event Public-Opinion Analysis Prototype (Conceptual Design)

5.3. Limitations and Future Directions

First, the number of complete basketball text samples that can enter the sentiment-index computation is still small. Although this paper has extended the public retrieval scope to Zhihu, Hupu, and Weibo, and established 58 public-source entries and 13 Weibo event trending-topic records, only the 28 Hupu event texts have completed semantic segmentation and

manual review under a unified specification; therefore, the sentiment indices and regression in Chapter 4 still belong to small-sample case analysis.

Second, ASAP restaurant reviews are used only for the algorithm benchmark experiment and do not share the same language distribution as basketball texts. SVM's direct preliminary-screening accuracy on the 28 basketball texts is low, so this paper does not use machine predictions directly for the final trend computation. To build a genuine basketball-specific sentiment classification model in the future, basketball review samples should be expanded and manually annotated, and training, validation, and test sets should be partitioned independently for modeling.

Third, the manual review of the basketball texts in this paper was completed by a single researcher, so subjective judgment bias may still exist. This paper reduces the risk of mislabeling through unified judgment rules and review in combination with the original context; in the future, two or more annotators can annotate independently, and consistency testing can improve label reliability.

Fourth, the trend analysis includes only five games, and the regression results are easily affected by a single game, so this paper only provides descriptive interpretation. In the future, more events such as the World Cup, the Asian Cup, and the CBA can be extended, increasing the number of games and adding variables such as opponent strength, pre-match expectations, and key player performance.

Fifth, this paper has carried out cross-platform public-source retrieval, but has not yet completed a cross-platform comment-level dataset in a uniform format. Some Zhihu entries still need further access to the Chinese source texts, and Weibo currently provides mainly historical trending topics and limited public content; therefore, at this stage only issue-level cross-verification can be performed, and different platforms cannot be directly merged into a training set or compared by sentiment proportion. In the future, original Chinese comments should be continuously expanded under the premise permitted by platform rules, and uniformly annotated after cleaning and deduplication.

Sixth, the current application part completes only the functional-framework design, and the cross-platform source index is still in the traceable-collection and expansion stage. In the future, if a larger-scale basketball review corpus whose Chinese source texts can be obtained one by one is formed, the approach of "AI pre-annotation + manual sampling/review" can be adopted to reduce annotation cost, and training, validation, and test sets can be independently partitioned after deduplication to build a basketball-specific model. The real-time performance, false-alarm rate, and actual decision value of a formal online system still need to be verified through real operation.

Data availability statement: The cross-platform public-source index newly added in this study has been archived separately with the project, including verifiable URLs and retrieval-purpose descriptions. Considering platform rules, user-content protection, and data integrity, this paper does not reproduce long passages of user original text, nor does it write data that only has indices or trending-topic records as obtained comment samples.

Fund project: This study was funded by the project "Undergraduate Scientific Research and Entrepreneurship Action (2026)" (Project No. 202601040K021).

References

[1] ZHAO Yanyan, QIN Bing, LIU Ting. Text sentiment analysis [J]. Journal of Software, 2010, 21(8): 1834-1848.

[2] WANG Xiaochen, HU Yuze. Research on the dissemination and governance of online public opinion of large-scale sports events under multi-subject collaboration [J]. Journal of Shenyang Sport University, 2023, 42(2): 123-129.

[3] WAN Xiaohong, CHEN Jiabao. Characteristics and governance of online public opinion on sports events in the new communication ecology: A dual perspective based on structural analysis and sentiment analysis [J]. China Sport Science, 2021, 41(4): 23-31.

[4] LEI Yueqiu, LI Yuanyuan. Exploring social media sentiment dissemination in major sports events: An empirical analysis based on Weibo text data of the Beijing Winter Olympics [J]. Omnimedia Exploration, 2022(12): 44-46.

[5] YU Y, WANG X. World Cup 2014 in the Twitter World: A big data analysis of sentiments in U.S. sports fans' tweets[J]. Computers in Human Behavior, 2015, 48: 392-400.

[6] CHANG Y. Spectators' emotional responses in tweets during the Super Bowl 50 game[J]. Sport Management Review, 2019, 22(3): 348-362.

[7] GONG X, WANG Y. Exploring dynamics of sports fan behavior using social media big data:

A case study of the 2019 National Basketball Association Finals[J]. Applied Geography, 2021, 129: 102438.

[8] LIU Q, JIANG S, WANG Y, et al. LiveQA: A Question Answering Dataset over Sports Live[C]//Proceedings of the 19th Chinese National Conference on Computational Linguistics. 2020: 1057-1067.

[9] SALTON G, BUCKLEY C. Term-weighting approaches in automatic text retrieval[J]. Information Processing & Management, 1988, 24(5): 513-523.

[10] CORTES C, VAPNIK V. Support-vector networks[J]. Machine Learning, 1995, 20: 273-297.

[11] BREIMAN L. Bagging predictors[J]. Machine Learning, 1996, 24(2): 123-140.

[12] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.

[13] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, et al. Scikit-learn: Machine Learning in Python[J]. Journal of Machine Learning Research, 2011, 12: 2825-2830.

[14] BU J, REN L, ZHENG S, et al. ASAP: A Chinese Review Dataset Towards Aspect Category Sentiment Analysis and Rating Prediction[C]//Proceedings of NAACL-HLT 2021. 2021: 2069-2079.