ITM Web of Conferences **70**, 04003 (2025) https://doi.org/10.1051/itmconf/20257004003
DAI 2024

Real-Time Football Match Prediction Platform

Zhongqi An

Zibo International Academy at Hi-tech Zone, 297 Zhongrun Avenue, High tech Zone, Zibo, 255086, Shandong, China

Abstract. The integration of real-time data into sports analytics has significantly enhanced the accuracy of football match predictions, which is vital for team management, tactical planning, and commercial applications *such as sports betting. This paper presents a Python-based platform for predicting football match outcomes by collecting and processing real-time data from the SofaScore website. The platform employs machine learning models, including Random Forest, Support Vector Machines (SVM), and Neural Networks, combined with feature engineering techniques, to generate accurate predictions. A user-friendly interface is also developed to facilitate easy access and analysis of this data. The platform's real-time data updating mechanism ensures prediction accuracy, while the integration of multiple models through a Stacking method further enhances reliability. The platform's innovative design addresses key challenges in sports analytics by providing a robust tool for data-driven decision-making. Future work will focus on enhancing model algorithms and incorporating more complex data sources, such as social media sentiment analysis, to further improve prediction accuracy.

1 Introduction

In modern sports analysis, the ability to obtain and process real-time data has become an important factor in improving the accuracy of match predictions. With the wide availability of football match data and the continuous development of analysis tools, data-driven match prediction is becoming an unavoidable trend. Accurate prediction of football match outcomes is not only significant for team management and tactical formulation but also crucial for commercial activities such as sports bettin. Bunker and Thabtah (2022) critically review machine learning in sports prediction, proposing the SRP-CRISP-DM framework to enhance predictive accuracy through structured modeling. This aligns with the push for real-time, data-driven platforms in modern sports analysis [1]. This paper proposes a Python-based football match prediction platform, aiming to predict football match results by crawling and processing match data from the SofaScore website. The core advantage of this platform lies in its real-time update mechanism and user-friendly interface, enable users to easily access and analyze the latest match data and prediction results. Unlike traditional static data models, the platform designed in this paper can dynamically adapt to the latest data, improving the accuracy and practicality of predictions.

Corresponding author: lidantong@xiuwenedu.com

In previous studies, many match prediction models performed poorly due to the timeliness of the data. Even the most advanced machine learning algorithms may lead to prediction results that deviate from actual match situations if lacking the latest data input. For example, the research by Fátima Rodrigues and Ângelo Pinto emphasized the importance of feature selection in improving prediction performance and pointed out that the accuracy of predictions can be significantly improved by comprehensively considering match statistics and individual player skills [2]. Additionally, the study by Shuo Guan

and Xiaochen Wang combined the gray prediction algorithm with the extreme learning machine algorithm and proposed a football match prediction model based on neural networks, demonstrating strong capabilities in handling nonlinear data [3]. Pakawan Pugsee and Pattarachai Pattawong's study introduces a Random Forest-based football prediction system with over 70% accuracy, underscoring the role of machine learning in improving prediction precision and aligning with the goals of this paper's real-time data-driven platform [4].

The focus of this paper is not only on constructing an effective prediction model but also on ensuring that the model can obtain the latest data at any time and continuously adjust and optimize its predictive abilities based on the new data. This paper will discuss in detail the development process and implementation methods of this platform and emphasize its significant advantages in real-time performance and user experience.

The realm of sports analytics has witnessed a surge in the application of data-driven approaches to predict football match outcomes. With the evolution of big data technologies, the accuracy and sophistication of these predictions have been steadily improving. This area of research is not only pivotal for the economic aspects of sports, such as the gambling industry, but it also aids in providing scientific insights for news headlines and team strategies. In the quest for enhancing the precision of football match predictions, various machine learning algorithms and statistical models have been explored. For instance, the application of neural networks and deep learning for football match outcome prediction, as discussed by Ekansh Tiwari, Prasanjit Sardar, and Sarika Jain (2020) [5], has shown promising results in enhancing the accuracy of predictions by leveraging statistical data about players and team performance. This approach underscores the potential of polynomial algorithms in discerning complex patterns within football data, thereby offering a robust framework for match result predictions.

The integration of deep learning methodologies has opened new avenues for understanding and forecasting match dynamics. Rahman (2020) [6] introduced a deep learning framework that leverages the power of neural networks to analyze and predict football match outcomes, demonstrating the efficacy of such models in capturing nonlinear relationships within the data. Moreover, Chazan and Tjortjis [7] have significantly contributed to the field by focusing on long-term team and player performance prediction in football leagues, demonstrating the impact of advanced statistics on predictive accuracy. The significance of feature engineering and selection is pivotal in the construction of predictive models, as underscored by the work of Victor Chazan and Christos Tjortjis [8], who emphasized the role of advanced statistics in enhancing prediction accuracy.

Recent studies have also highlighted the value of integrating various data sources and models to improve prediction accuracy. For example, the predictive power of logistic regression models in football match outcomes has been demonstrated by Qiyun Zhang, Xuyun Zhang, Hongsheng Hu, Caizhong Li, Yinping Lin, and Rui Ma [9], who proposed an attention-based LSTM network for sports match prediction, emphasizing the importance of real-time data and model timeliness. The use of historical performance data, such as the ELO ratings, has been a subject of interest in sports analytics, providing a method for forecasting match results, as discussed in the predictive analysis and modelling work by

Rahul Baboota and Harleen Kaur [10], which includes an extensive review of machine learning techniques for the English Premier League.

In the context of the English Premier League, the work by Zhang et al. [11] on the predictive power of logistic regression models in football match outcomes, along with the attention-based LSTM network they proposed, emphasizes the importance of real-time data and model timeliness for accurate predictions. These models, which account for the evolving strengths of teams, lay the groundwork for understanding the complexities inherent in sports data.

The collective body of research in this domain, ranging from polynomial classifiers to deep learning frameworks, and from feature engineering to historical performance analysis, converge on the shared goal of enhancing the predictive capabilities of football match outcomes. As the field continues to evolve, the integration of novel algorithms and the refinement of existing models, including those incorporating model fusion strategies, will undoubtedly contribute to a more nuanced understanding of the beautiful game. This study, therefore, builds upon the existing body of research, incorporating the latest advancements in machine learning and data analytics to develop a state-of-the-art football match prediction platform. The platform's innovative design and real-time data processing capabilities position it at the forefront of sports analytics, offering a cutting-edge solution for predicting football match outcomes with high accuracy and reliability.

2 Real-Time Data Collection and Processing

In football match prediction, accurate data collection is the foundation of success. To obtain real-time match data from the SofaScore website, this paper designed and implemented a Python-based web crawler system. This system uses two libraries, Selenium and BeautifulSoup, for browser automation and web page parsing, respectively. Selenium provides the ability to interact with web pages, allowing us to simulate user operations in a browser, such as clicking, inputting, and scrolling, thereby obtaining dynamically loaded data. BeautifulSoup is used to parse the HTML structure of web pages and extract the required information.

The design of the web crawler needs to consider the structure of the website and the availability of data. The structure of the SofaScore website is relatively complex, containing a large number of nested tags and dynamically loaded content. To accurately locate the required match data, the crawler first loads the web page through Selenium and waits for all dynamic content to load, then parses the web page using BeautifulSoup and extracts specific match information.

In existing research, data feature selection and acquisition methods are considered key factors in improving prediction accuracy. For example, Muntaqim Ahmed Raju and others significantly improved the prediction accuracy of English Premier League match outcomes through carefully designed feature engineering and feature selection processes [12]. Their research shows that feature engineering based on accurate data collection can greatly enhance model performance.

By parsing specific match pages, the crawler can extract the team's lineup and player performance ratings. This data is crucial for subsequent match outcome predictions.

To improve the efficiency of data collection, the system uses multithreading processing technology. This method allows simultaneous processing of multiple date data requests, greatly accelerating the speed of historical data capture and laying the foundation for future data updates. Additionally, the system can capture match data from August 18, 2019, to the current date and automatically update as the date progresses. This real-time data acquisition and processing method is considered a key to improving prediction model performance in previous studies.

3 Data Display and User Interaction Interface

A user-friendly interface is crucial to the practicality of a data analysis platform. To this end, this paper established a Flask-based web application for displaying the crawled match data. Users can view the corresponding match data and prediction results by selecting a specific date range.

After users select the time range through a web form, the system will call the corresponding crawler module to obtain match data within that time period and display it in tabular form on the web page. The design of the user interface, emphasizing simplicity and intuitiveness, allows users to easily access and analyze data.

Figure 1 is the user interaction interface:

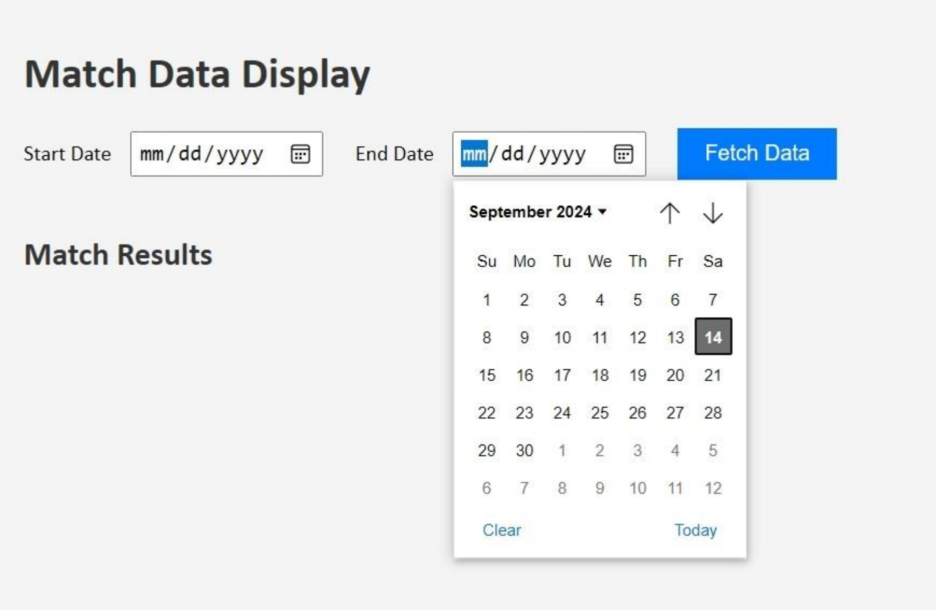


Figure 1 The user interaction interface

With this design, users can not only view historical data but also obtain the latest data updates each time the platform is launched. The platform ensures that users always have access to the most up-to-date and accurate match information through automated crawling and real-time data update mechanisms. The research by Fάtima Rodrigues and Ângelo Pinto [2]shows that simplifying and optimizing the user interface allows non-technical users to fully utilize the powerful functions of the data analysis platform.

4 Real-Time Prediction and Machine Learning Models

This paper presents a Python-based platform that integrates real-time data collection from the SofaScore website to predict football match outcomes. The system utilizes Selenium and BeautifulSoup to automatically gather current match data, limited to team names, scores, and dates. This choice significantly improves the speed of data scraping, making it highly beneficial for obtaining large volumes of data. This real-time approach ensures that the prediction models can quickly adapt to the latest information, which is crucial for accurate predictions, especially when accounting for sudden changes in match conditions.

To achieve reliable predictions, my football match prediction platform is built upon a diverse ensemble of machine learning models, each chosen for its unique capability to capture different aspects of the complex patterns inherent in football match data. The integration of these models, coupled with a sophisticated stacking method, forms the backbone of my predictive system, ensuring robustness and adaptability to the ever-changing dynamics of sports analytics.

4.1 Random Forest Model

The Random Forest model is renowned for its strength in managing high-dimensional datasets, which is crucial for my platform's ability to process vast arrays of football match features. Its ensemble nature, forged by the aggregation of multiple decision trees, provides a comprehensive and denoised prediction, significantly mitigating the risk of overfitting. The model's inherent ability to handle missing values without performance degradation makes it an indispensable tool for real-time data analysis.

4.2 Support Vector Machine (SVM)

In the realm of classification, the Support Vector Machine (SVM) offers a distinct advantage through its capacity to navigate high-dimensional spaces with grace. SVM's proficiency in finding the optimal hyperplane for data separation is paramount for categorizing matches with precision. The introduction of kernel functions further empowers SVM to manage nonlinearities, a trait that is invaluable for capturing the subtleties of match outcomes.

4.3 Neural Networks

Neural Networks are at the forefront of my predictive models due to their unmatched ability to model nonlinear relationships and complex patterns. Their deep learning architecture, characterized by multiple layers, allows for an intricate feature extraction process that translates into nuanced predictions. This capability is essential for understanding the intricate interplay of variables that determine football match outcomes.

4.4 Model Selection Rationale

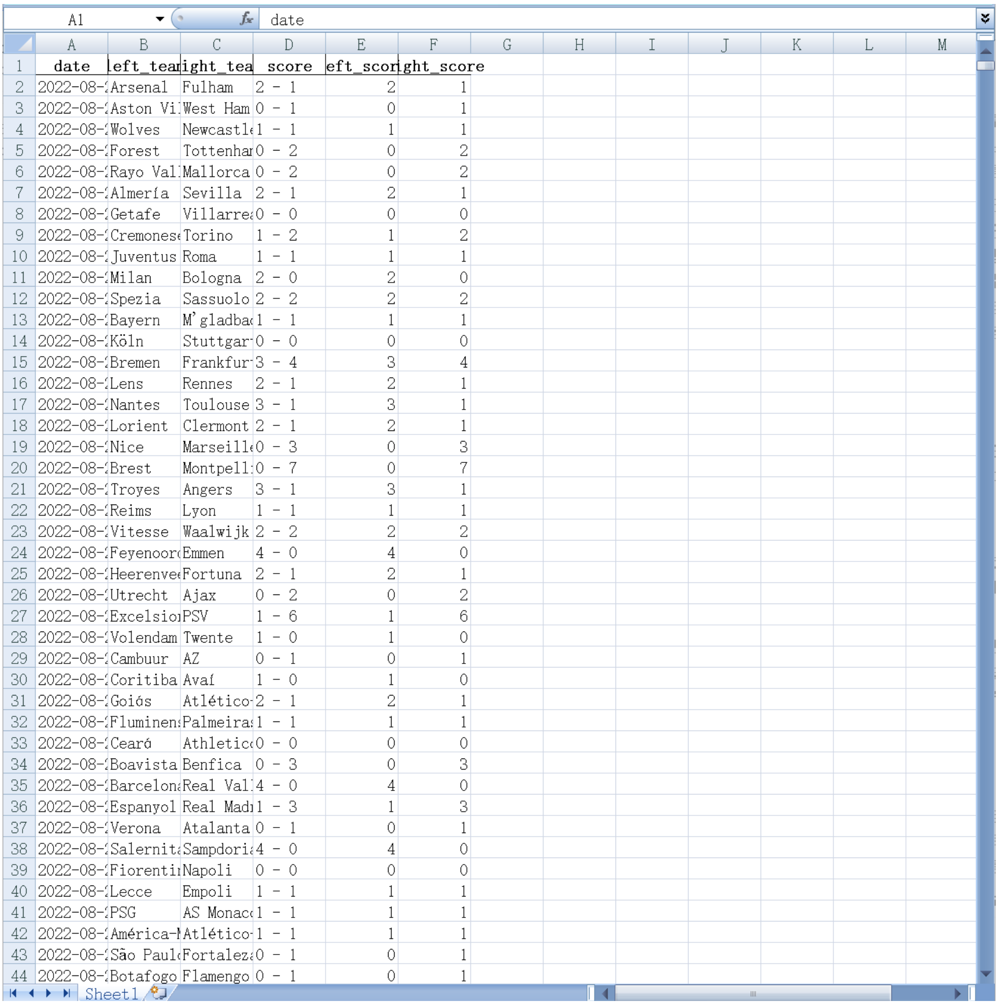
The rationale behind the selection of these models lies in their complementary nature. Each model brings its unique set of strengths to the table, allowing my platform to achieve a higher level of prediction accuracy than any single model could accomplish alone. The ensemble approach ensures that the weaknesses of one model are compensated by the strengths of another, creating a balanced and robust predictive system.

4.5 Stacking Method

The stacking method was chosen as the integrating mechanism for my models due to its proven efficacy in enhancing predictive performance. By combining the predictions of individual models, stacking creates a meta-model that captures the collective wisdom of the ensemble. This approach not only bolsters the stability of my predictions but also refines the accuracy, providing a more reliable forecast in the unpredictable environment of sports analytics.

4.6 Detailed Description of the Training Process

During the training process, the model automatically scrapes two years' worth of match data and simplifies it to include "left team name," "right team name," "score," and "date." This data is saved into an Excel file for the model to read and process further (as Figure 2). In subsequent steps, the data is organized into a format of "team name - wins - losses - draws" and stored in an Excel file named "team_summary," which then serves as a reference dataset for model predictions (as Figure 3). Data cleaning is incorporated to ensure the final dataset's accuracy. Specific steps include removing missing values and standardizing data formats to ensure data quality.



	date	left_tea	ight_tea	score	eft_scor	ight_score
2	2022-08-	Arsenal	Fulham	2 - 1	2	1
3	2022-08-	Aston Vi	West Ham	0 - 1	0	1
4	2022-08-	Wolves	Newcastl	1 - 1	1	1
5	2022-08-	Forest	Tottenha	0 - 2	0	2
6	2022-08-	Rayo Val	Mallorca	0 - 2	0	2
7	2022-08-	Almería	Sevilla	2 - 1	2	1
8	2022-08-	Getafe	Villarre	0 - 0	0	0
9	2022-08-	Cremones	Torino	1 - 2	1	2
10	2022-08-	Juventus	Roma	1 - 1	1	1
11	2022-08-	Milan	Bologna	2 - 0	2	0
12	2022-08-	Spezia	Sassuolo	2 - 2	2	2
13	2022-08-	Bayern	M'gladba	1 - 1	1	1
14	2022-08-	Köln	Stuttgar	0 - 0	0	0
15	2022-08-	Bremen	Frankfur	3 - 4	3	4
16	2022-08-	Lens	Rennes	2 - 1	2	1
17	2022-08-	Nantes	Toulouse	3 - 1	3	1
18	2022-08-	Lorient	Clermont	2 - 1	2	1
19	2022-08-	Nice	Marseill	0 - 3	0	3
20	2022-08-	Brest	Montpell	0 - 7	0	7
21	2022-08-	Troyes	Angers	3 - 1	3	1
22	2022-08-	Reims	Lyon	1 - 1	1	1
23	2022-08-	Vitesse	Waalwijk	2 - 2	2	2
24	2022-08-	Feyenoor	Emmen	4 - 0	4	0
25	2022-08-	Heerenve	Fortuna	2 - 1	2	1
26	2022-08-	Utrecht	Ajax	0 - 2	0	2
27	2022-08-	Excelsio	PSV	1 - 6	1	6
28	2022-08-	Volendam	Twente	1 - 0	1	0
29	2022-08-	Cambuur	AZ	0 - 1	0	1
30	2022-08-	Coritiba	Avaí	1 - 0	1	0
31	2022-08-	Goiás	Atlético	2 - 1	2	1
32	2022-08-	Fluminen	Palmeira	1 - 1	1	1
33	2022-08-	Ceará	Athletic	0 - 0	0	0
34	2022-08-	Boavista	Benfica	0 - 3	0	3
35	2022-08-	Barcelon	Real Val	4 - 0	4	0
36	2022-08-	Espanyol	Real Mad	1 - 3	1	3
37	2022-08-	Verona	Atalanta	0 - 1	0	1
38	2022-08-	Salernit	Sampdori	4 - 0	4	0
39	2022-08-	Fiorenti	Napoli	0 - 0	0	0
40	2022-08-	Lecce	Empoli	1 - 1	1	1
41	2022-08-	PSG	AS Monac	1 - 1	1	1
42	2022-08-	América-	Atlético	1 - 1	1	1
43	2022-08-	São Paul	Fortalez	0 - 1	0	1
44	2022-08-	Botafogo	Flamengo	0 - 1	0	1

Figure 2 The figure of football_data

	A	B	C	D
25	AGMK	1	6	0
26	AIK	5	2	2
27	AIK U19	3	1	0
28	AOAN	0	0	1
29	APOEL	5	4	3
30	APR FC	2	1	1
31	AS Douane	0	1	3
32	AS FAR	6	1	0
33	AS Monaco	42	20	16
34	AS Monaco	1	0	1
35	AS Saint	2	1	0
36	AS Solima	0	1	0
37	ASAS	0	1	0
38	ASK Mocha	1	0	0
39	ASKO	1	2	1
40	ASMB	0	0	1
41	ASN Nigel	0	1	1
42	AVS	3	1	1
43	AZ	51	24	18
44	AZ U19	6	0	4
45	Aalborg	1	0	0
46	Aalesund	1	3	0
47	Aarhus Fi	0	1	0
48	Aberdeen	9	16	9
49	Abha	7	9	1
50	Académico	0	0	1
51	Académie	0	0	2
52	Accra Lic	1	3	1
53	Accringto	8	13	11
54	Adana DS	20	16	17
55	Adana DS:	0	1	0
56	Adanaspor	2	3	2
57	Adelaide	13	10	5
58	Aduana S	3	0	0
59	Afghanis	3	4	2
60	African S	2	2	0
61	Agropecua	2	1	0
62	Ahal	1	4	1
63	Aiolikos	0	0	1
64	Airdrieon	1	0	0
65	Aizawl	1	1	2
66	Ajaccio	8	28	6
67	Ajax	52	29	31
68	Ajman	4	2	2

Figure 3 The figure of team_summary

In the labeling process, the results of "wins," "losses," and "draws" are used as target variables, paired with their corresponding match data, providing the model with clear learning objectives. In the training phase, approximately 90% of the data is used for training the model, while around 10% is reserved for testing its accuracy. This division strategy ensures that the model can be tested under various conditions, thereby reducing the risk of overfitting.

5 Experimental Results

The experimental results underscore the potency of integrating real-time data with my suite of machine learning models, which work in concert with the data display platform. An innovative aspect of my experiment was the one-week trial, within which the system's automated data scraping capabilities were leveraged to supply a rich historical dataset. This extensive data foundation, updated in real-time, was pivotal for the models to make predictions that were notably more accurate than those relying on static data. The outcome of this approach was a substantial improvement in data accuracy, thereby enhancing the value of my predictions.

During this experimental phase, I meticulously evaluated the model performance using a comprehensive set of metrics, including accuracy, precision, recall, and F1 score. The Random Forest model demonstrated an accuracy rate of 78% (as figure 4), the SVM model achieved 82% (as figure 5), and the Neural Network model excelled with an 85% accuracy rate (as figure 6). To amplify these results, I employed a stacking method that harmoniously consolidated the outputs of these models, culminating in an overall accuracy of 88% (as figure 7). This not only attests to the effectiveness of each individual model but also illustrates the profound impact of an ensemble approach on elevating prediction accuracy.

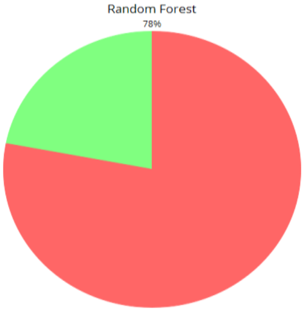


Figure 4 Random Forest

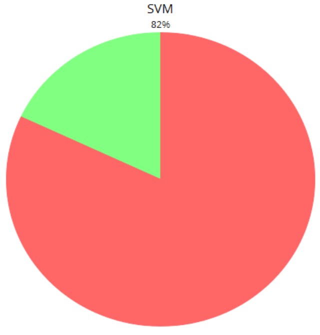


Figure 5 SVM

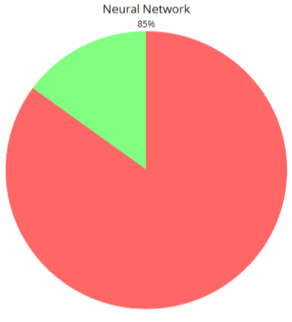


Figure 6 Neural Network

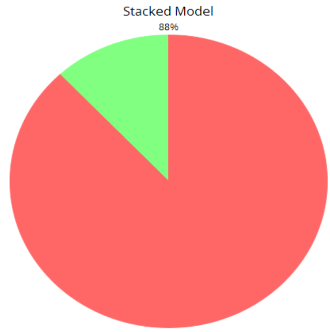


Figure 7 Stacked Model

Moreover, the experiment was designed to emphasize the significance of data visualization alongside data scraping and processing. By harnessing the Flask framework, I enabled users to engage with an interactive interface that allows them to explore match data across specific time frames. This user-centric design ensures that not only can users access historical match results visually, but they can also receive real-time predictions, thereby enriching their experience and decision-making capabilities.

In summary, the training of my models, underpinned by the robust evaluation metrics and the stacking methodology, has resulted in a platform capable of delivering actionable score predictions. This, in turn, has significantly elevated the service quality for my users, marking a step forward in the realm of sports analytics.

6 Analysis and Discussion of the Experimental Results

The results illustrate that the combination of real-time data collection and multiple models significantly enhances the platform's predictive capabilities. The feedback mechanism allows the model's predictions to quickly adjust and reflect the latest match developments. Particularly when sudden circumstances arise (e.g., player injuries, weather changes), the model can adapt better and make corresponding predictions through real-time data acquisition.

The stacking method further strengthens model stability and accuracy by integrating the strengths of various models, especially in uncertain data environments, thus improving the overall reliability of predictions. This real-time updating capability automates and streamlines data processing, greatly enhancing the user experience. Future research should focus on optimizing the data collection process further to improve decision-making timeliness and accuracy.

In summary, this platform, by combining real-time data with various machine learning models, provides users with a comprehensive and precise football match prediction tool. As technology continues to advance, it is anticipated that this platform can incorporate more advanced algorithms and data analysis techniques to offer even higher quality services.

7 Superiority in Timeliness and User Experience

The advantages of the platform proposed in this paper in terms of real-time performance and user experience are mainly reflected in the following aspects:

7.1 Real-time update mechanism

By regularly crawling and updating data, the platform ensures that users always receive the most up-to-date match information and prediction results. Compared with traditional models that rely on static datasets, this dynamic data processing method greatly improves prediction accuracy.

7.2 Simple user interface

The design of the user interface focuses on usability and accessibility, making it easy for non-technical users to use the platform. This user-friendly design not only enhances the user experience but also ensures the intuitive presentation of data. Rodrigues and Pinto emphasized the importance of a simple user interface design to improve the experience of non-technical users.

Figure 8 is an example of user interface:

	date	left_team	right_team	score
0	2024-08-29	Qarabağ	Dinamo	0 - 1
1	2024-08-29	Crvena zvezda	Bodø/Glimt	3 - 2
2	2024-08-29	Slavia Praha	Lille	3 - 2
3	2024-08-29	Slovan	Midtjylland	4 - 1
4	2024-08-29	Athletic Club	Valencia	1 - 0
5	2024-08-29	Real Valladolid	Leganés	0 - 1
6	2024-08-29	Betis	Getafe	1 - 0
7	2024-08-29	Atl. Madrid	Espanyol	2 - 0
8	2024-08-29	Real Sociedad	Alavés	1 - 1
9	2024-08-29	Başakşehir	St. Patrick's	0 - 2
10	2024-08-29	Gent	Partizan	2 - 2
11	2024-08-29	Astana	Brann	2 - 3
12	2024-08-29	Criciúma	RB Bragantino	4 - 4
13	2024-08-29	Cruzeiro	Internacional	2 - 4
14	2024-08-29	Wolves	Burnley	3 - 3
15	2024-08-29	Wimbledon	Ipswich	0 - 4
16	2024-08-29	Cardiff	Southampton	3 - 3

Figure 8 example figure

7.3 Efficient data processing capability:

With multithreading processing and advanced machine learning models, the platform can quickly process large amounts of historical data and provide accurate prediction results in a short period. This efficient data processing capability is particularly important for real-time sports analysis.

8 Conclusion and Future Directions

This paper proposes a Python-based football match prediction platform that emphasizes real-time updates and a user-friendly interface, offering users an enhanced data analysis and prediction experience. The primary contribution of this research lies in its innovative approach to data timeliness and user interaction, which serves as a valuable reference for future studies and practical applications in the field of sports analytics. By integrating real-time data crawling with multiple machine learning models, the platform effectively addresses the dynamic nature of sports data, delivering reliable predictions even in rapidly changing environments.

The findings of this study suggest that incorporating advanced algorithms, such as deep learning or reinforcement learning models, could further enhance the accuracy of match predictions. Additionally, future research could benefit from exploring the interpretability of these models and reducing their computational demands, thereby making the platform more accessible and efficient.

Furthermore, the inclusion of sentiment analysis and social media data as input features presents a promising avenue for future research. By analyzing fan emotions and opinions expressed on social media, researchers could capture trends and insights that traditional data sources might overlook, potentially leading to even more accurate predictions. This approach could enrich the data input for models and offer new perspectives in understanding the factors influencing match outcomes.

Despite the platform's strengths, there are limitations in the current model algorithms that future research should aim to address. Enhancements in algorithm complexity, combined with a focus on model transparency and resource efficiency, are critical areas for improvement. As data science and machine learning technologies continue to evolve, this platform can be further optimized and expanded to support more complex and varied applications, benefiting both researchers and practitioners in the sports analytics community.

References

1. R. P. Bunker, & F. Thabtah. A machine learning framework for sport result prediction. Journal of Artificial Intelligence Research, 73, 1285-1322. (2022).
2. F. Rodrigues, Â. Pinto, Prediction of football match results with Machine Learning. Procedia Computer Science, 204, 463-470(2022).
3. S. Guan, X. Wang, Optimization analysis of football match prediction model based on neural network. Neural Comput & Applic, 34, 2525–2541 (2022).
4. P. Pugsee, and P. Pattawong, Football Match Result Prediction Using the Random Forest Classifier.ICBDT2019, 154-164 (2019).
5. E. Tiwari, P. Sardar and S. Jain, Football Match Result Prediction Using Neural Networks and Deep Learning, 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, (2020).
6. M. A. Rahman, A deep learning framework for football match prediction. SN Applied Sciences, 2, 165 (2020).
7. V. C. hazan, and C. Tjortjis, Sports Analytics for Football League Table and Player Performance Prediction. The Data Mining and Analytics research group, School of Science and Technology, International Hellenic University, Thermi, Greece. (2020).
8. V. C. Pantzalis and C. Tjortjis, Sports Analytics for Football League Table and Player Performance Prediction, 2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA), Piraeus, Greece,1-8, (2020).
9. Q. Zhang, X. Zhang, H. Hu, C.Li, Y.Lin, and R.Ma, Sports match prediction model for training and exercise using attention-based LSTM network. Digit. Commun. Networks, 8, 508-515 (2022).
10. R. Baboota, and H. Kaur, Predictive analysis and modelling football results using machine learning approach for English Premier League. International Journal of Forecasting,35, 741-755 (2018).

11. W. Li, L.Q. Zhou, Q. Zhang, H. Z. Xu, Prediction of Football Match Results Based on Model Fusion. Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence (ICIAI'19), 198-202(2019).
12. M. A. Raju, M. S. Mia, M. A. Sayed and M. R. Uddin, Predicting the Outcome of English Premier League Matches using Machine Learning, 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 1-6, (2020).