IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

XAI-POWERED SPORTS ANALYTICS SUITE FOR PLAYER PERFORMANCE, INJURY RISK, AND MATCH OUTCOME

[1]Taufiq R. Nagori, [2]Anil D. Marathe, [3]Gopal V. Thakare, [4]Saurav A. Band, [5]Ketaki A. Kadu
[1]Student, [2]Student, [3]Student, [4]Student, [5]Professor
Department of Computer Science & Engineering
Prof. Ram Meghe College of Engineering & Management, Amravati,
India

Abstract: In modern sports, data and artificial intelligence are widely used to improve player performance, predict injuries, and support strategic planning. However, many AI models are difficult to interpret because they do not clearly explain how their predictions are generated, which reduces trust among coaches and analysts. This research proposes an Explainable Artificial Intelligence (XAI)-based sports analytics system that integrates player performance analysis, injury risk prediction, and match outcome prediction within a unified framework. The system analyzes sports data, including player statistics, match history, and injury records, using machine learning techniques, while XAI methods are applied to clearly explain the reasoning behind each prediction. The results demonstrate that the proposed system provides accurate predictions along with meaningful explanations, thereby enhancing decision-making and increasing confidence in AI-based systems. Overall, the study highlights the importance of XAI in making sports analytics more transparent, reliable, and suitable for real-world applications.

Keywords: Explainable AI, Sports Analytics, Machine Learning, Injury Risk Prediction, Player Performance Analysis, Match Outcome, SHAP, Data Science

I. INTRODUCTION

Sports are more than just games; they play a vital role in promoting physical fitness, mental well-being, and social interaction. They also help develop important life skills such as teamwork, discipline, leadership, and perseverance. Over time, sports have evolved from simple recreational activities into highly competitive and well-organized systems where performance, outcomes, and strategic planning are equally important. In the modern era, sports are no longer driven solely by talent and experience but are increasingly influenced by technology and data-driven approaches.

Today, sports generate a vast amount of data from various sources, including player movements, match events, training sessions, fitness levels, and injury records. The process of collecting and analyzing this data to support decision-making is known as sports analytics. By examining factors such as player speed, stamina, positioning, and historical performance trends, coaches and analysts can identify strengths and weaknesses, enhance training programs, design effective match strategies, and reduce injury risks. Data-driven decision-making enables teams to perform more consistently and maintain player fitness throughout long and demanding seasons.

The integration of Artificial Intelligence (AI) has significantly enhanced the capabilities of sports analytics. AI systems can process large datasets quickly and accurately, enabling teams to predict match outcomes, evaluate player performance, analyze opponent strategies, and recommend optimal team formations. AI is also widely applied in injury prediction, fitness monitoring, talent identification, and tactical planning. By combining AI with traditional analytical methods, teams gain deeper insights and make evidence-based decisions rather than relying solely on intuition or experience.

One notable area of research focuses on improving football performance evaluation using the expected goals (xG) model. Instead of relying only on final match scores, which may be influenced by luck or random events, the xG model assesses performance based on the quality of scoring opportunities. Researchers have analyzed data from more than 315,000 shots taken in top European football leagues and applied machine learning techniques to estimate the probability of scoring. To enhance interpretability, Explainable Artificial Intelligence (XAI) methods were used to illustrate how shot-related factors, such as position and angle, influence scoring probability. This approach allows coaches and analysts not only to assess performance outcomes but also to understand the underlying reasons behind them, making the model both accurate and trustworthy.

Despite these advancements, a major limitation of many AI systems is that they function as black-box models. Although such models may provide highly accurate predictions, they often fail to explain how or why specific decisions are made. This lack of transparency creates trust concerns, particularly in critical areas such as injury prediction and player selection. Coaches and medical staff may hesitate to rely on AI recommendations if the reasoning behind them is unclear.

To address this issue, Explainable Artificial Intelligence (XAI) has emerged as an effective solution. XAI aims to make AI models transparent and interpretable by highlighting the key factors and patterns influencing predictions. In sports analytics, XAI can clarify why a player's performance is improving or declining, why a particular outcome is predicted, or which variables contribute most significantly to injury risk. By providing understandable explanations, XAI enhances trust and enables coaches, players, and analysts to confidently apply AI-driven insights in practical decision-making.

The primary objective of this research is to develop an explainable AI-based sports analytics system that is both accurate and interpretable. Unlike conventional black-box models, the proposed system not only generates predictions but also clearly explains the reasoning behind them. This integrated approach supports improved player performance, injury prevention, optimized team strategies, and informed decision-making.

In conclusion, the integration of sports analytics, artificial intelligence, and explainable AI represents the future of modern sports management. Sports analytics provides structured and meaningful data, AI delivers predictive capabilities, and XAI ensures transparency and trust. Together, these technologies create a smarter, safer, and more strategic sports environment. This research aims to bridge the gap between advanced AI methodologies and practical sports applications by delivering a reliable and interpretable explainable AI framework.

II. LITERATURE REVIEW

Today, sports generate a vast amount of data from matches, training sessions, and player monitoring systems. This data includes information such as player positions, passes, shots, speed, heart rate, and fitness levels. Teams use this data to enhance player performance, develop match strategies, and reduce injury risks. Artificial Intelligence (AI) and Machine Learning (ML) are widely applied to analyze such data because they can efficiently process large volumes of information and identify patterns that may not be easily recognized by humans. For example, AI systems can determine which players perform best in specific situations, analyze how team formations influence results, and detect signs of player fatigue. This enables coaches and analysts to make informed decisions and improve both individual and team performance over time.

Although many AI models provide accurate predictions, they often fail to explain how those predictions are generated. For instance, a model may indicate that a player has a high probability of scoring but may not clarify the contributing factors. This lack of explanation reduces trust among coaches and medical staff. Explainable Artificial Intelligence (XAI) addresses this limitation by revealing the reasoning behind each prediction. For example, XAI methods may highlight that a player's position, recent form, or shooting angle significantly influences scoring probability. This transparency enhances trust and facilitates the practical application of AI in football analytics [6].

Certain AI models emphasize the quality of scoring opportunities rather than relying solely on final match results. These models analyze data from thousands of shots, passes, and game events to estimate goal probability. Explainable AI techniques can identify key factors—such as distance from goal, shooting angle, defensive pressure, and player ability—that most strongly influence scoring outcomes. This information assists coaches in understanding why goals are scored or missed and helps in designing targeted training sessions to improve decision-making and finishing skills [9].

Machine learning techniques are also used to evaluate individual players based on fitness, technical skills, and historical performance data. These models can recommend suitable playing positions and identify areas requiring improvement. For instance, a player with high stamina and strong passing ability may be better suited for a midfield role, whereas another player with strong defensive skills may perform better in defense. Such analysis supports personalized training programs and data-driven decision-making. However, many traditional models focus solely on prediction accuracy without offering explanations, which may reduce confidence in their recommendations. Therefore, interpretable and explainable models are essential for effective player development [8].

Various approaches exist for predicting football match outcomes. Some models forecast categorical results, such as win, loss, or draw, while others estimate the number of goals scored. Both approaches provide valuable baseline insights for football analytics and help teams identify the factors influencing match performance. These models support strategic comparisons, team selection optimization, and opponent analysis. Integrating explainable AI into match outcome prediction further enhances understanding by clarifying the reasons behind predicted results [11].

Injury prediction is another critical application of machine learning in sports analytics. Accurately predicting injuries is challenging due to limited data, imbalanced injury records, and complex influencing factors. Explainable AI assists by identifying risk factors such as excessive workload, inadequate recovery, or abnormal movement patterns. With these insights, coaches and medical staff can implement preventive measures, adjust training intensity, and minimize injury occurrences, thereby maintaining team competitiveness [12].

Several studies have proposed general frameworks for implementing explainable AI in sports analytics. These frameworks aim to highlight the most influential features while maintaining model simplicity and interpretability. Although some frameworks were initially developed for other sports, they can be effectively adapted to football analytics. The adoption of such frameworks enhances understanding of both individual and team performance, supports strategic planning, and strengthens trust in AI-driven systems. Overall, explainable AI significantly improves the practical applicability and reliability of sports analytics [10].

III. PROBLEM STATEMENT

In modern football, a large volume of data is collected from matches, training sessions, and player tracking systems. This data includes player movements, fitness metrics, match events, and injury records. Although such data is highly valuable for improving performance and preventing injuries, manual analysis is time-consuming and inefficient. Identifying meaningful patterns becomes difficult when the dataset is large and complex.

To manage this data effectively, Artificial Intelligence (AI) and Machine Learning (ML) techniques are widely used. These technologies assist in predicting player performance, match outcomes, and injury risks. However, most traditional AI models provide only final predictions, such as performance scores or risk levels, without explaining how those predictions are generated. As a result, coaches and medical staff may hesitate to rely fully on these systems.

Another limitation of existing approaches is that they often focus on a single task, such as goal prediction or injury forecasting, rather than integrating multiple analytical functions within one system. Additionally, in the absence of interpretability, it becomes difficult to determine which factors—such as fitness level, workload, or player role—significantly influence the predictions. This lack of transparency restricts the practical application of AI in football decision-making.

To address these challenges, this project proposes an Explainable Artificial Intelligence (XAI)-based football analytics system. The system applies machine learning models to predict player performance, injury risk, and match outcomes within a unified framework. In addition to generating predictions, it provides clear and interpretable explanations for each result.

Explainable AI techniques such as SHAP and LIME are utilized to illustrate how different features influence the predictions. These methods identify important factors such as fitness condition, recent performance, workload, playing position, and match context. This enhances transparency and enables users to understand the reasoning behind each prediction.

Overall, the proposed explainable football analytics system supports more informed and reliable decision-making. By combining predictive accuracy with interpretability, the system contributes to improved performance management, effective injury prevention, and strategic match planning in a practical and dependable manner.

IV. PROPOSED METHODOLOGY

The proposed methodology outlines the systematic procedure followed to design and develop the Explainable Artificial Intelligence (XAI)-based football analytics system. It explains how raw football data is collected, processed, analyzed, and transformed into predictions accompanied by clear explanations. The primary objective of this approach is to convert unstructured football data into meaningful, interpretable insights that assist stakeholders in making informed decisions.

The process begins with data collection. Football-related information is gathered from various reliable sources, including match statistics, training session records, player tracking systems, fitness assessments, and injury reports. The collected dataset contains key attributes such as player workload, playing position, performance indicators, match outcomes, recovery details, and medical history. To maintain uniformity and simplify processing, all data is stored in CSV format.

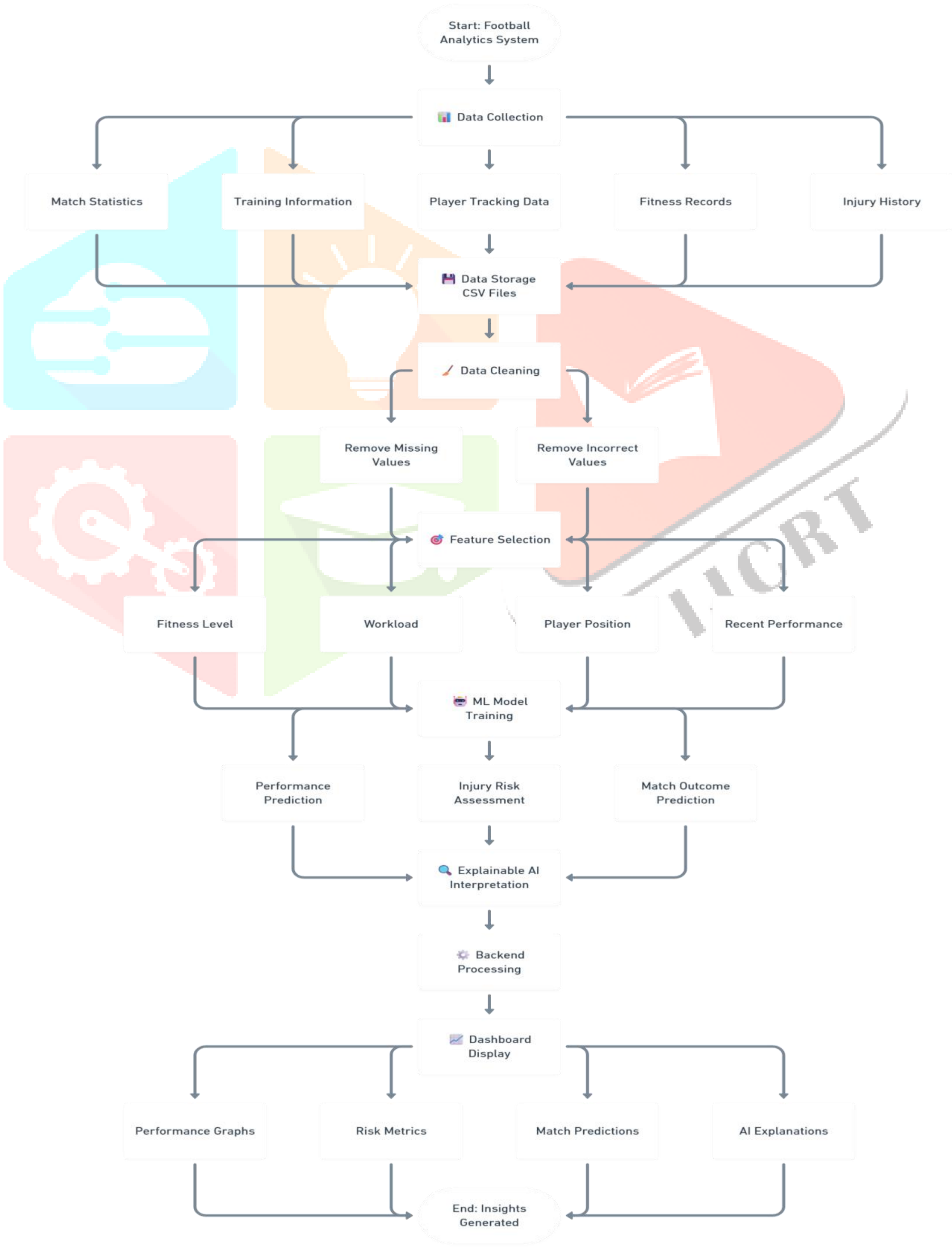
The next stage is data preprocessing. During this phase, data quality is enhanced by addressing missing values, eliminating duplicate entries, and correcting inconsistencies. The dataset is then transformed into a structured format suitable for machine learning algorithms. Feature selection techniques are applied to identify the most relevant factors influencing predictions, such as fitness level, workload intensity, recent performance trends, player position, and match conditions. This ensures that only meaningful and clean data is used for model training.
After preparing the dataset, machine learning models are developed and trained for specific analytical tasks. Separate models are designed for player performance prediction, injury risk estimation, and match outcome forecasting. These models learn patterns from historical data and apply that knowledge to generate predictions on new inputs. Through training, the system identifies relationships between different variables and their impact on performance, injuries, and match results.

To enhance transparency, Explainable Artificial Intelligence techniques such as SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are integrated into the system. These methods interpret model outputs by highlighting the contribution of each feature to a specific prediction. For instance, they can demonstrate how increased workload may elevate injury risk or how consistent match participation may improve overall performance. This interpretability component builds trust and improves the reliability of the system.

The system follows a frontend–backend architecture. The backend is developed using FastAPI, which manages data processing, model execution, and explanation generation. The frontend is implemented using Streamlit, providing an interactive dashboard where users can input data and visualize prediction results along with explanation graphs.

Finally, prediction results and explanations are presented through visualizations such as bar charts, line graphs, pie charts, and feature importance plots. These visual tools enable coaches, analysts, and medical professionals to easily interpret the outputs. By integrating machine learning, explainable AI techniques, and intuitive visualization, the methodology ensures that the system is practical, transparent, and applicable to real-world football analytics scenarios.

V. SYSTEM ARCHITECTURE



VI. RESULTS

In this project, player performance prediction plays a vital role in the overall system. The machine learning model analyzes historical football data, including player fitness levels, workload, playing position, and recent match statistics to evaluate how effectively a player is performing. Based on these factors, the system classifies each player's performance into three simple categories: Low, Medium, and High. This categorization makes the results easy to interpret for coaches, analysts, and other users.

The predicted performance outcomes are displayed using visual representations such as bar charts or line graphs. In these graphs, player names are presented along the x-axis, while their predicted performance levels are shown on the y-axis. These visual tools allow coaches and analysts to quickly compare players and make informed decisions. Players identified with High performance levels may be selected for crucial matches, whereas those categorized as Low performance may require additional training, rest, or performance improvement strategies. Therefore, performance prediction supports effective team selection and structured training planning. The player performance results are illustrated in Figure 1.

Additionally, the system provides a detailed explanation of each player's performance score. This explanation clarifies the factors that influenced the prediction. For example, a higher performance score may be assigned to a player who consistently participated throughout the season, regularly appeared in the starting lineup, and actively contributed in multiple matches. Such continuous involvement and experience increase the reliability of the model's prediction and justify the assigned performance level.

Figure 1 Player Performance Prediction Results

Figure 1.1 Player Performance Prediction Analysis

In addition to performance evaluation, the system also includes an injury risk prediction module, which plays a crucial role in maintaining player fitness and overall team stability. The injury prediction model analyzes several important factors, such as player workload, current fitness condition, training intensity, and previous injury records to estimate the probability of injury. Based on this analysis, the system categorizes injury risk into three levels: Low, Medium, and High, making the results simple and easy to interpret.

The predicted injury risk levels are presented using a bar graph, where player names are displayed along the x-axis and their corresponding injury risk levels are shown on the y-axis. This visualization enables coaches and medical staff to quickly identify players who may require special attention. Players categorized under High risk can be assigned rest periods or reduced training intensity, while those with Low risk can continue with their regular training schedule. By taking preventive measures based on these predictions, the system helps minimize injuries and maintain team fitness.

The injury risk prediction results are illustrated in Figure 2.

Figure 2 Injury Risk Prediction

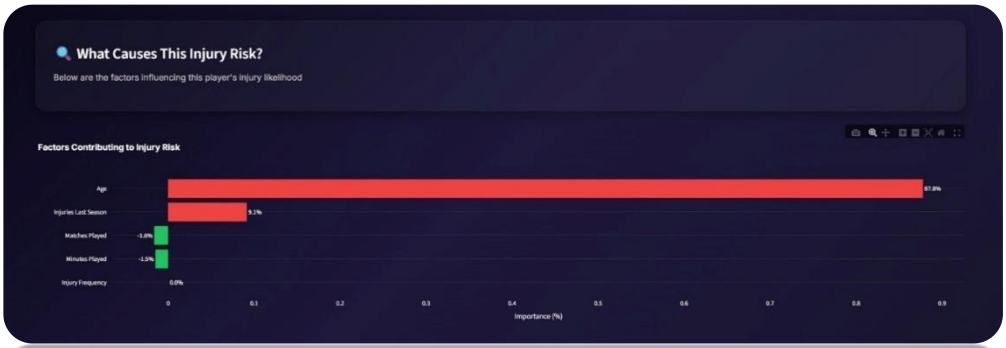


Figure 2. 1 Causes of this Injury risk

The system further supports match preparation by incorporating a match outcome prediction module. By analyzing factors such as overall team performance, current player form, and historical match data, the model estimates the probability of three possible outcomes: Win, Draw, or Loss. These predictions provide valuable insights that assist teams in developing effective pre-match strategies.

The predicted probabilities are presented through visualizations such as a pie chart or a probability-based bar chart. These graphical representations clearly display the likelihood of each possible result, making the information easy to interpret. Coaches and analysts can use these insights to adjust tactics, refine player selection, and improve overall match planning.

The match outcome prediction results are illustrated in Figure 3.

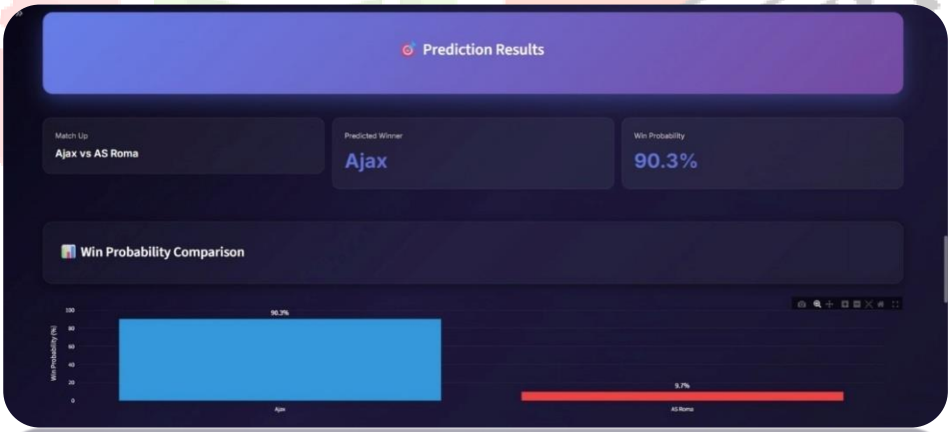


Figure 3 Match outcome

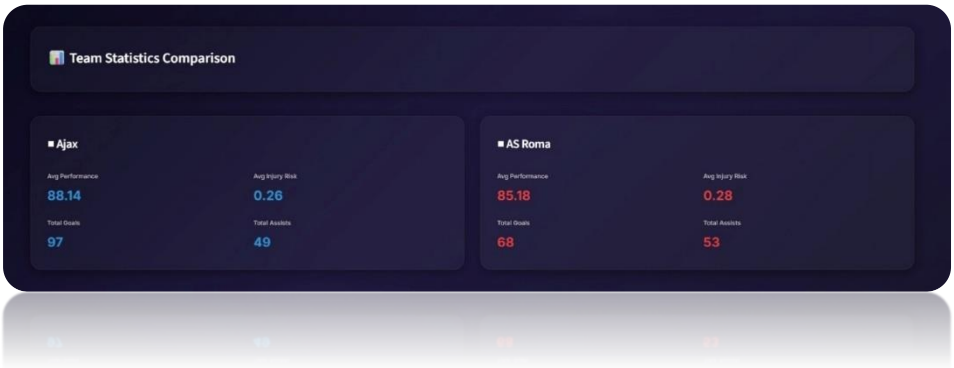


Figure 3.1 Team Statistics Comparison

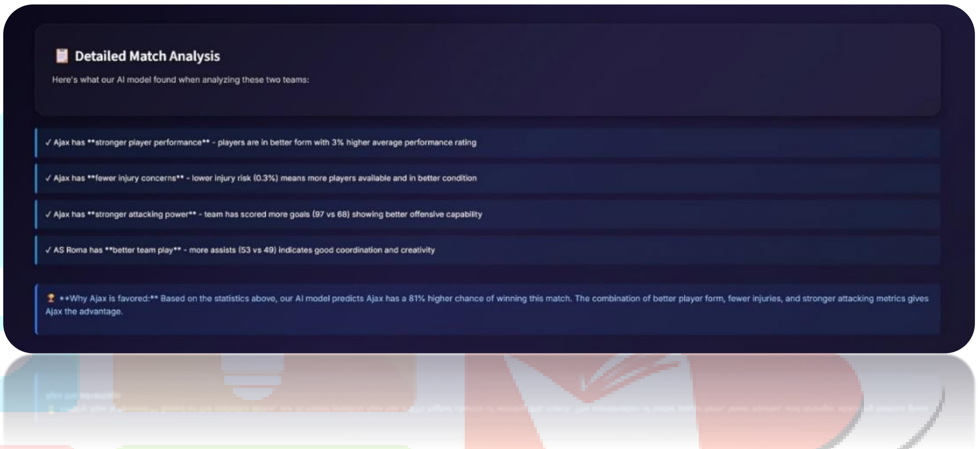


Figure 3.2 Match outcome Analysis

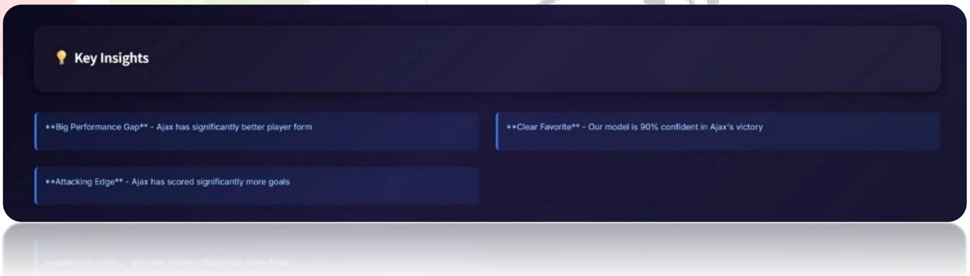


Figure 3.3 key insights of Match Outcome

CONCLUSION AND FUTURE SCOPE

This project develops an Explainable AI-based sports analytics system that helps analyze player performance, injury risk, and match results on a single platform. The system studies football data such as player fitness, workload, match performance, and injury history to provide useful insights. Instead of showing complex or difficult results, the system clearly explains why a particular prediction was made. This makes it easier for coaches and analysts to understand the information and take better decisions.

The project shows that when data is explained in a simple and clear way, it becomes more helpful and reliable. The system is built using FastAPI for handling the backend processes and Streamlit for creating a user-friendly and interactive interface. Overall, this project demonstrates that explainable analytics can support better performance planning, reduce injury risks, and improve match preparation in football.

In the future, the system can be improved by adding real-time data from player tracking devices and fitness sensors. Including data from more seasons and different leagues can make the predictions more accurate. The platform can also be expanded to support other sports and provide automatic suggestions to improve performance and match results.

REFERENCE

[1] R. Takamido, J. Ota, and H. Nakamoto, “PassAI: An explainable machine learning for predicting soccer pass outcomes using multimodal match data,” pp. 132884 132895, 2025,doi: 10.1109/ACCESS.2025.3589903

[2] S. Moustakidis, S. Plakias, C. Kokkotis, T. Tsatalas, and D. Tsaopoulos, “Predicting football team performance with explainable AI,” 2023, doi: 10.3390/fi15050174.

[3] C. Feng and L. Wang, “Football action recognition based on explainable artificial intelligence assisted by unmanned aerial vehicles,” 2025, doi: 10.1109/TCE.2025.3571434.

[4] D. O. Capanema, T. V. de Souza, J. C. Serrão, and A. C. M. Pereira, “Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports,” 2019, doi: 10.1186/s40798-019-0202-3.

[5] Z. Pu, Y. Pan, S. Wang, B. Liu, M. Chen, H. Ma, and Y. Cui, “Orientation and decision-making for soccer based on sports analytics and AI: A systematic review,” pp. 37–57, 2024, doi: 10.1109/JAS.2023.123807

[6] A. Procopiou and A. Piki, “The 12th player: Explainable artificial intelligence (XAI) in football: Conceptualisation, applications, challenges and future directions,” pp. 213–220, 2023, doi: 10.5220/0012233800003587.

[7] S. Kranzinger, C. Halmich, D. Hofer, and C. Kranzinger, “A scoping review of explainable artificial intelligence in sports science,” pp. 1–17, 2026, doi: 10.1007/s44163-025-00709-8.

[8] S. Sneha, B. S. Prithvi, M. Niranjanamurthy, H. K. Kiran Kumar, and P. Dayananda, “Machine learning based assessment of elite football players based on anthropometric and motor fitness parameters with regard to their playing positions,” pp. 1–15, 2024, doi: 10.1007/s42979-024-03261-x.

[9] M. Cavus and P. Biecek, “Explainable expected goal models for performance analysis in football analytics,” pp. 1–25, 2022.

[10] T. Zhao, J. Cabral, and G. Zhu, “A novel explainable artificial intelligence framework using knockoffs techniques with applications to sports analytics,” pp. 1–30, 2025, doi: 10.1007/s10479-025-06575-y.

[11] A. Tsokos, S. Narayanan, I. Kosmidis, G. Baio, M. Cucuringu, G. Whitaker, and F. Király, “Modeling outcomes of soccer matches,” pp. 77–95, 2019, doi: 10.1007/s10994-018-5741-1.

[12] A. Majumdar, R. Bakirov, D. Hodges, S. Scott, and T. Rees, “Machine learning for understanding and predicting injuries in football,” pp. 1–10, 2022, doi: 10.1186/s40798-022-00465-4.