Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

13

Sports Analysis Using Machine Learning

Shreya Ghogare,

master's in computer application, MIT WORLD PEACE UNIVERSITY Pune,

shreyaghogare10@gmail.com

Shounak Gandurkar,

master's in computer application, MIT WORLD PEACE UNIVERSITY Pune,

shounakgandurkarcode@gmail.com

Vishwajit Narkhede,

master's in computer application, MIT WORLD PEACE UNIVERSITY Pune,

jeetnarkhede.99@gmail.com

Dr. Irfan sayyed

syed.irfan@mitwpu.edu.in

Abstract:

With the advent of big data and machine learning approaches, sports analysis has experienced a substantial revolution. Large-scale data analysis has been transformed by machine learning, which has also revealed insights that were difficult or impossible to discover with more conventional qualitative methods in the past. This study analyzes machine learning based sports analysis, emphasizing the methodologies and approaches used as well as the applications and difficulties.

Data from numerous sources, such as player monitoring, social media, and video feeds, are analyzed using machine learning models. like neural networks and decision trees. Decisions on player choice, game strategy, injury prevention, and fan engagement may be made using the knowledge gained from these assessments. The need for more thorough and data, the

blending of many data and the moral application of data are the difficulties. Despite these difficulties, machine learning is set to become more widely used in sports analysis, offering insightful data and raising the level of competition.

KEYWORDS: Sports analytics, Data-driven sports performance, Machine learning algorithms, Athlete tracking, Game strategy development, Predictive modeling, Data quality, Interpretability, Performance prediction, Injury prediction, Sports technology, Wearable devices, big data, Sports science, Performance optimization.

INTRODUCTION

Modern sports performance has grown significantly impacted by sports analysis, which offers players, coaches, and analysts' insightful knowledge and data. Machine learning algorithms may be used to assess the massive volumes of data produced by the growing usage of sensors, cameras, and wearable technology. In order to handle and evaluate huge data sets, machine learning has emerged as a potent technology that can be utilized to provide a greater knowledge of sports performance and assist find areas for growth.

The goal of this study is to find out a comprehensive review of machine learning's current state-of-the-art in sports analysis. We'll look at the numerous approaches and procedures used in the industry, including feature selection, data pre-processing, and model construction.

We will also go through the uses of machine learning in sports analysis, such as player monitoring, injury prediction, and game plan formulation. We will also discuss a few difficulties with machine learning in sports analysis, including data quality, interpretability, and model generalization.

Ultimately, our goal is to provide a thorough overview of machine learning's application to sports analysis, emphasizing how it has the potential to revolutionize how sports are practiced, taught, and studied. Our results should stimulate further investigation and advancement in this fascinating area.

LITERATURE REVIEW

Traditional methods of sports analysis have depended on qualitative evaluations and subjective observations, which are vulnerable to prejudice and inconsistent results. The

emergence of big data and machine learning, however, has completely changed how sports analysis is done by making it feasible to analyze vast volumes of data and draw conclusions that were previously difficult or impossible to draw.

Athlete tracking is one of the key uses of machine learning in sports analysis. Sensors and cameras are used by athlete tracking systems to gather information about an athlete's performance and movement during practice and competition. The data may then be analysed using machine learning techniques to provide information on things like technique, speed, and endurance. For instance, it has been shown that using machine learning in athlete monitoring would increase the precision of performance forecasts and lower the risk of injury.

Game strategy creation is another area in which machine learning has been used in sports analysis. In order to find trends and forecast the results of upcoming games, machine learning algorithms may be trained on past game data. This may aid in the development of better game plans and the ability of coaches and analysts to make better game-related judgements. For instance, models that can accurately predict the results of soccer matches have been created using machine learning.

Unfortunately, there are a number of difficulties in applying machine learning to sports analysis. The quality of the data is one of the major issues. Machine learning models' accuracy may be impacted by sports data that is noisy, lacking, and inconsistent. Interpretability presents another difficulty. Coaches and analysts may find it difficult to comprehend the insights produced by machine learning models since these models might be tricky to interpret.

In conclusion, the field of sports analysis has been transformed by the availability of large amounts of data and machine learning. The way sports are played, taught, and studied might all be revolutionized by the machine learning in analysis. To fully use the capabilities of this technology, however, a number of problems are also presented by the machine learning in sports analysis.

IMPLEMENTATION

This model uses a Random Forest algorithm to make predictions by creating decision trees based on factors such as toss winners, players, venue, and DL methodology. It also uses SVM

to distinguish between different algorithms. Essentially, the model makes accurate predictions by analyzing data and using algorithms to identify patterns and make decisions.

1) **Random forest algorithm:** Random Forest is a type of machine learning that creates multiple decision trees using different sets of data. Each decision tree has the same starting point but different results, and these results are combined to find an average solution. The final decision is made by the random forest, which chooses the decision that is supported by the majority of the trees. In simpler terms, Random Forest is a method of combining the results of multiple decision trees to make more accurate predictions.

2) **SVM algorithm:** Support Vector Machine (SVM) is a type of machine learning algorithm used for classification or regression problems, although it's mainly used for classification. The algorithm plots each data point as a point in an n-dimensional space, where n is the number of features, and assigns each point a value. This allows the algorithm to predict not only the winner of a match, but also the expected number of runs to be scored by both teams.

RESULTS

Some important discoveries emerged from our thorough evaluation of the literature on sports analysis using machine learning. Secondly, we discovered that artificial intelligence has been used in a number of sports, including baseball, basketball, tennis, and soccer. The most commonly used machine learning method was supervised learning, which involved training models to predict outcomes or classify data.

Using unsupervised learning, dimensionality reduction and grouping were also accomplished. The creation of intelligent agents capable of making the best judgements in challenging circumstances made use of reinforcement learning. Natural language processing and picture identification were both accomplished using deep learning.

We also identified several applications of machine learning in sports analysis, including performance prediction and optimization, game strategy development, and injury prediction. Athlete monitoring data analysis and the creation of individualized training plans also utilized machine learning. Additionally, automatic highlight generating and real-time analytics have

been used in sports broadcasting using machine learning. We discovered that the two biggest problems in machine learning sports analysis were data quality and interpretability. Missing or incorrect data may have an impact on how well machine learning models perform, and it can be challenging to integrate and standardize data from many sources and formats. Moreover, machine learning models may be tricky to comprehend, making it difficult to obtain knowledge and base judgements on the findings.

Overall, our review highlights the potential of machine learning in sports analysis, but also emphasizes the need for careful consideration of data quality and interpretability issues.

DISCUSSION

The application of machine learning in sports analysis has the potential to revolutionize the way sports are played, coached, and analyzed. The ability to process and analyze vast amounts of data from various sources, such as player monitoring, social media, and video feeds, can provide valuable insights into player performance, injury prevention, game strategy, and fan engagement. The integration of machine learning in sports analysis has enormous potential to improve sports performance and enhance the overall fan experience. However, it is important to address the challenges associated with data quality and interpretability to ensure the effective and ethical use of machine learning in sports analysis. As technology continues to advance, it will be exciting to see how machine learning will further transform the world of sports.

CONCLUSION

Choosing the right team is important for winning a match. Our goal is to analyze IPL cricket data and predict the outcome of a match. We used three classification algorithms and compared them to find the best one. We used Anaconda Navigator and Jupyter for implementation. The Random Forest algorithm was found to be the most accurate with an 85.582% prediction rate. By using this prediction, we can form the best team.

REFERENCES

[1] M. Ks, "Applications of Artificial Intelligence in the Game of Football: The Global," - International Refereed Social Sciences Journal, p. 13, 2020.

[2] M. W. Farooq, "Critical Review on Test Match Format of Cricket," Global Social Sciences Review (GSSR), p. 9, 2021.

[3] Q. L. a. Q. Liu, "Prediction of Volleyball Competition UsingMachine Learning and," Woosuk University, Wanju, Republic of Korea, p. 8, 2021.

[4] A. Malhotra, "A Statistical Analysis of Bowling Performance in Cricket," Indian Institute of Technology, Bombay, India, p. 10, 2014.

[5] A. Sinha, "Application of Machine Learning in Cricket and Predictive Analytics of IPL 2020," Amity University Jharkhand Ranchi, p. 27, 2022.

[6] M. Daniyal, "Analysis of Batting Performance in Cricket using Individual and Moving Range," Department of statistics, the university of Sargodha, p. 9, 2014.

[7] A. Sahu, "Predictive Analysis of Cricket," Turkish Journal of Computer and Mathematics Education p. 14, 2021.

[8] J. D. Ulf Brefeld1, "Machine Learning in Sports," Report from Dagstuhl Seminar, p. 19, 2011.