

Integrated Framework for Spark-Based Big Data Analytics in Sports Health Monitoring

- This presentation explores an integrated framework that utilizes **Spark-based big data analytics** to enhance athlete performance and health monitoring through advanced techniques.

① Research Paper: Volume 28, pages 1585–1608, (2024)

Authors: Yin Zhao, Ma. Finipina Ramos & Bin Li

Link: <https://link.springer.com/article/10.1007/s00500-023-09450-9>

❏ Presented By:

Aaditya Gautam_122ad0022

Ganesh M_122ad0026

Pratyush Parth_122ad0050



Challenges in Sports Medical Services

Addressing the Gaps in Sports Medical Care

1 Rising Demand for Sports Medicine

- There is an increasing public interest in high-quality sports medical services.

2 Resource Limitations in China & India

- China and India face significant shortages in medical resources impacting sports health.

3 Importance of Real-Time Monitoring

- Athletes require immediate health monitoring and personalized recommendations for performance.

4 Inefficiency of Existing Models

- Current predictive health models are often inaccurate and inefficient, hindering effective care.



Role of Big Data and AI in Sports Medicine

Exploring the Impact of Data Science in Athlete Health



Big Data Analytics

- Analyzes vast datasets to derive actionable insights in sports medicine.



Artificial Intelligence

- Utilizes algorithms to improve decision-making in athlete health and performance.



Apache Spark

- Framework for processing large-scale sports data efficiently and quickly.



XGBoost

- Machine learning model for predictive analytics in injury prevention.



Machine Learning

- Learns from data patterns to predict outcomes and enhance training methods.



Injury Prevention

- Uses data insights to predict and prevent athlete injuries effectively.

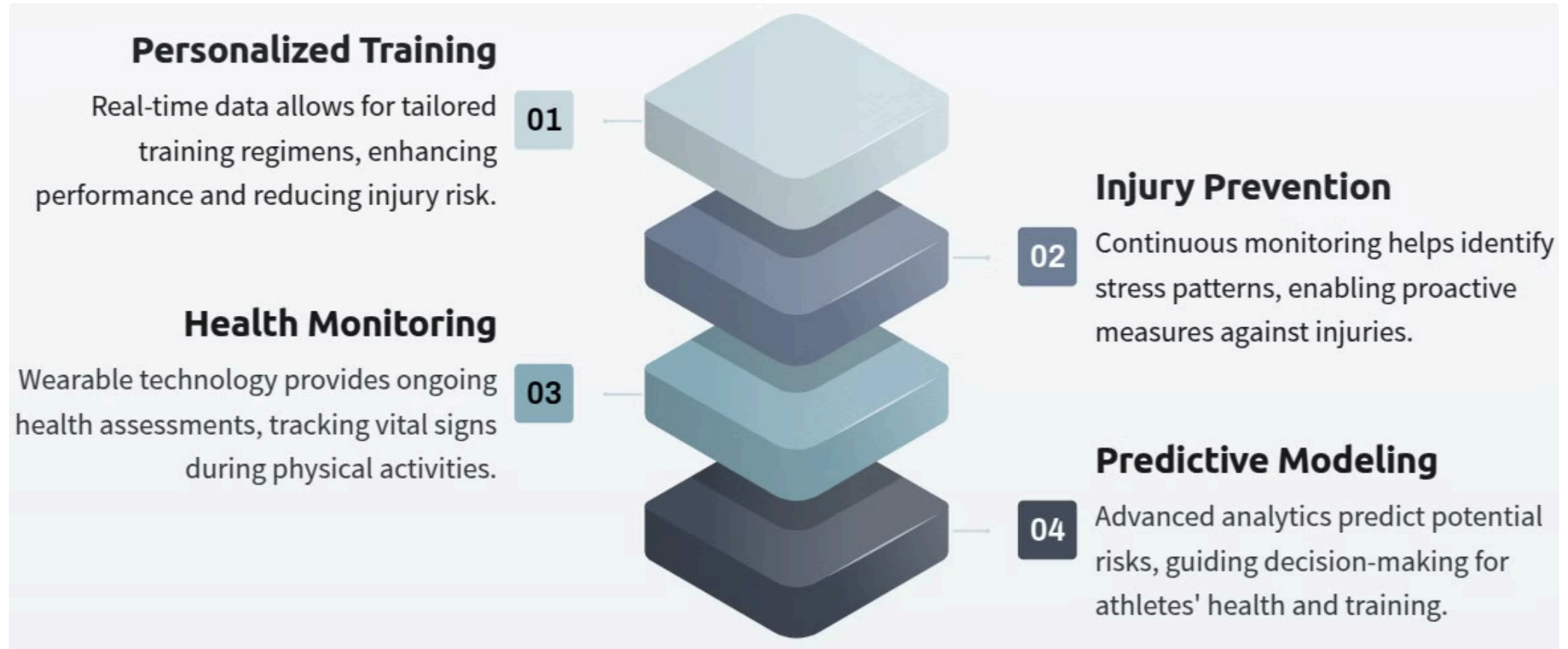


Performance Enhancement

- Analyzes athlete performance metrics to tailor training programs.

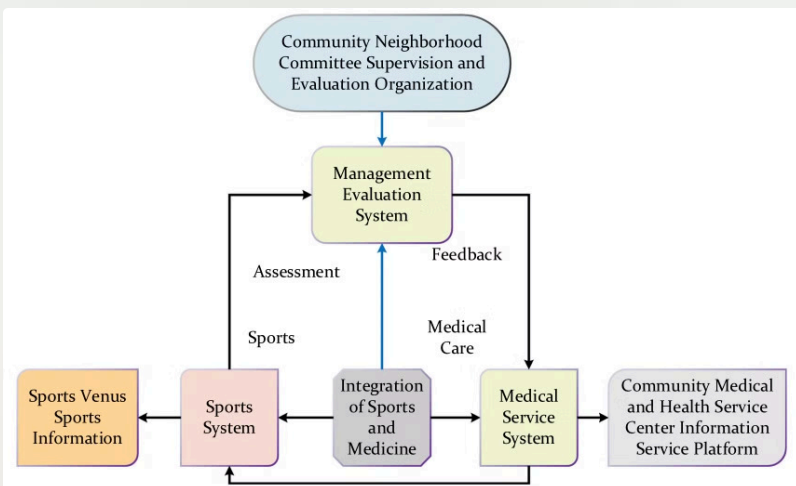
Integrating Sports and Medicine Benefits

Exploring the advantages of merging sports and medical technologies



Related Work in Sports Medicine

Exploring Wearable Technology and Machine Learning Challenges



Wearable Devices

- Wearable devices and biometric sensors are critical for real-time athlete monitoring.

ML Models

- Various machine learning models like KNN, SVM, RF, DT, and Naïve Bayes enhance data analysis in sports medicine.

Challenges

- Challenges remain in achieving accuracy with real-time data processing in athlete monitoring systems.

Integration Issues

- System integration poses significant challenges, impacting the effectiveness of sports medicine technologies.

Innovative Advances in Research Contributions

Exploring significant advancements in data processing and analysis

1

Spark-Based Framework

A new Spark-based Big Data framework enhances scalability in data processing, enabling efficient handling of large datasets.

2

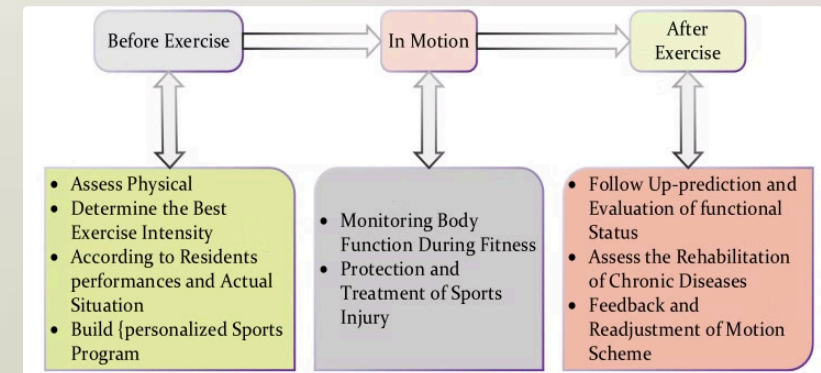
XGBoost Application

XGBoost significantly improves prediction accuracy, showcasing its effectiveness in data-driven decision-making.

3

Real-time Sports Medical Decisions

The research addresses critical gaps in real-time decision-making for sports medicine, impacting athlete care and performance.



Identifying Research Gaps in Predictive Models

1

Traditional Models

Traditional models face difficulties processing extensive datasets efficiently and in real-time.

2

Need for Personalization

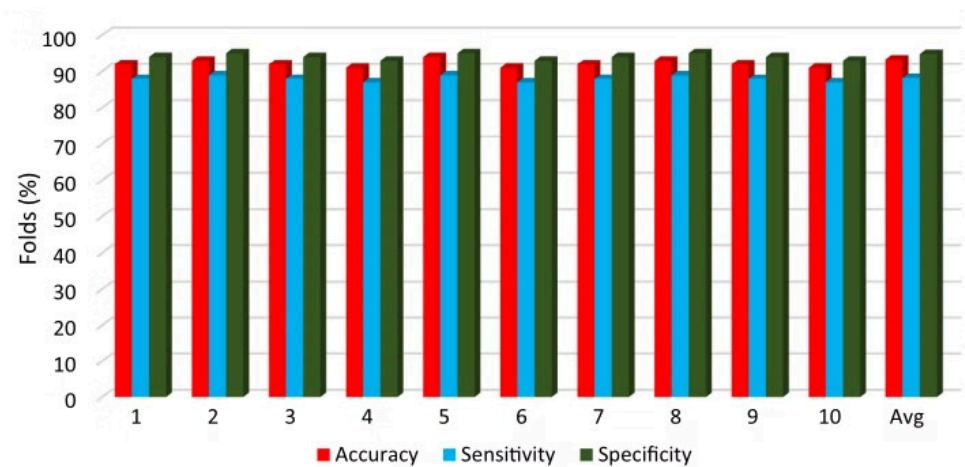
Current systems lack the capability to provide tailored health recommendations based on individual data.

3

Scalability

Need for a **scalable, high-accuracy predictive framework**.

Fig. 9 Performance metrics of the proposed XGBoost algorithm with 10 k-fold cross-validation



Overview of the Proposed Framework

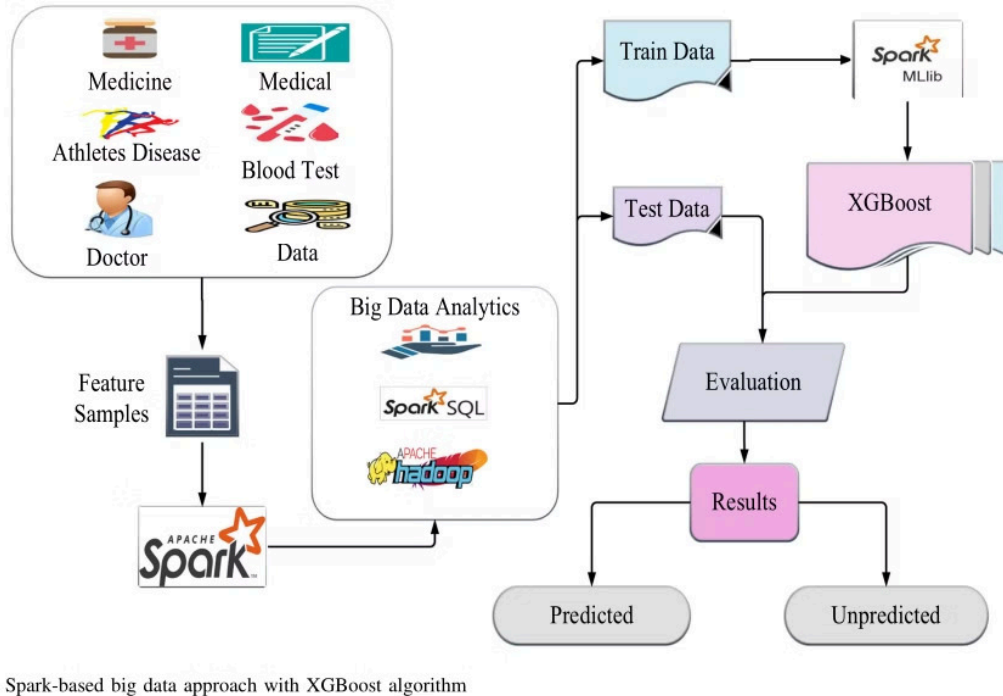


Fig. 1 A Spark-based big data approach with XGBoost algorithm

Apache Spark

- Apache Spark for **distributed data** processing.

XGBoost

- **Predictive** analytics.

Wearable Sensors

- Integration of **wearable sensors and real-time health data**.

Integrating Data Processing and Predictive Analytics

Spark-Based Big Data Processing

Leveraging Spark for real-time processing of sports data from training sessions and events enhances decision-making through efficient parallel computing.

Spark handles large-scale real-time data from:

- Training sessions.
- Athlete biometric data.
- Sports event monitoring.

Distributed processing enables **fast, efficient data analysis**.

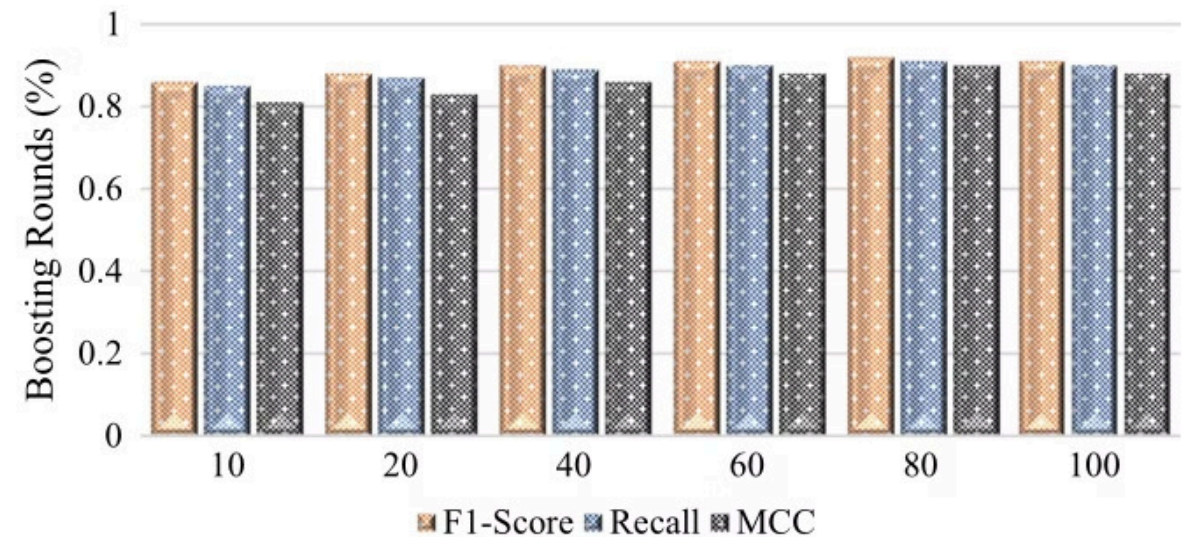


Fig. 12 The proposed XGBoost algorithm at different numbers of boosting rounds

XGBoost for Health Prediction

Leveraging Machine Learning for Athletic Health Insights

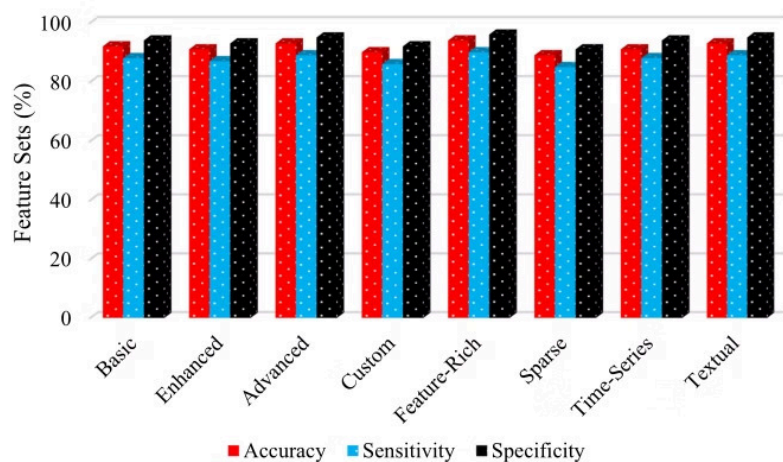


Fig. 13 Performance of the proposed XGBoost algorithm with different feature sets

Utilization of Historical Data

- XGBoost is trained on comprehensive historical sports health datasets for accurate predictions.

Performance Level Assessment

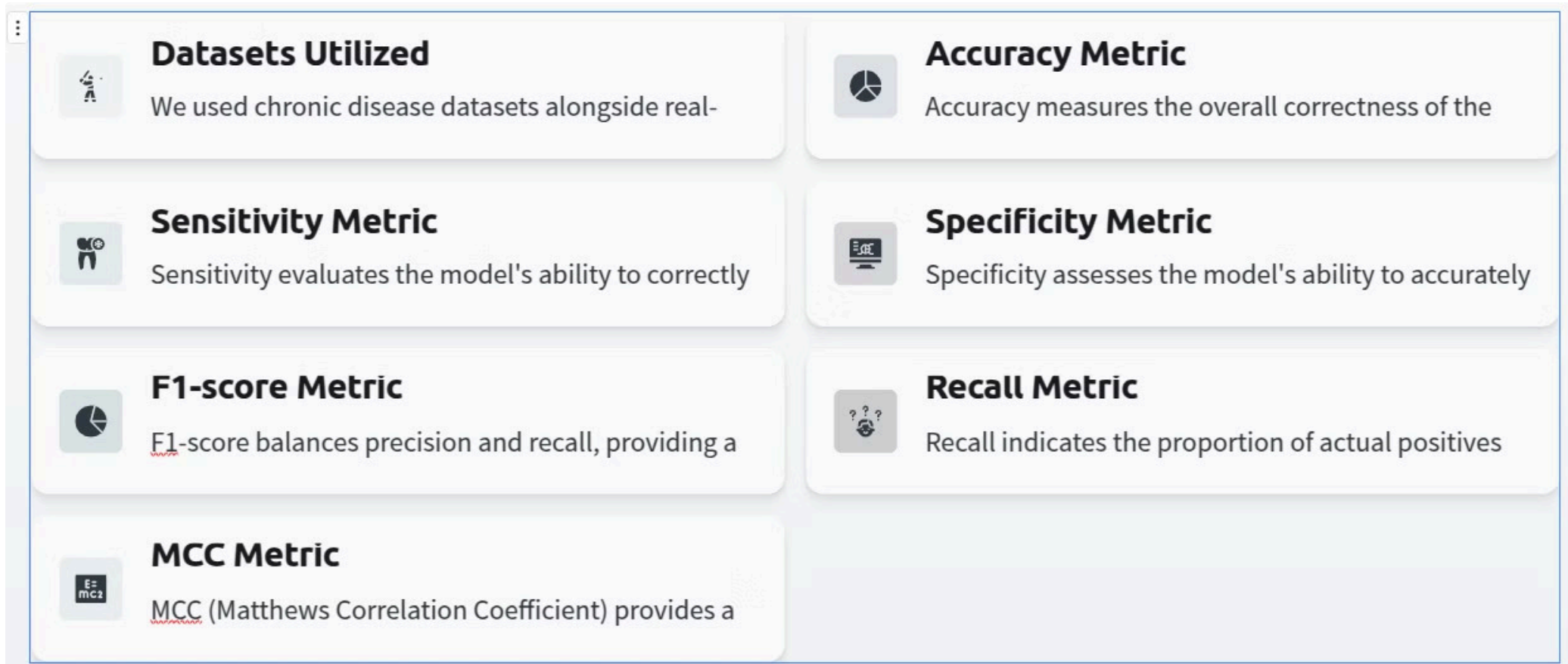
- Predicts athletes' performance levels, providing insights for training optimization.

High Prediction Accuracy

- Outperforms **KNN, RF, DT, SVM, Naïve Bayes, and Logistic Regression** with **93% accuracy**.

Overview of Experimental Setup

Analyzing Metrics and Datasets for Performance Insights



XGBoost Results & Performance Insights

Analyzing the Efficacy and Efficiency of XGBoost

01 High Trust Rate in Predictions

XGBoost achieved an impressive 93% trust rate for health risk predictions, showcasing reliability.

02 Superior Predictive Accuracy

This model significantly outperformed traditional models, enhancing decision-making processes.

03 Scalability with Spark Integration

Demonstrated excellent scalability and real-time capabilities through its integration with Spark.

04 Comparative Performance Overview

An image table illustrates the comparative performance of XGBoost against other models.

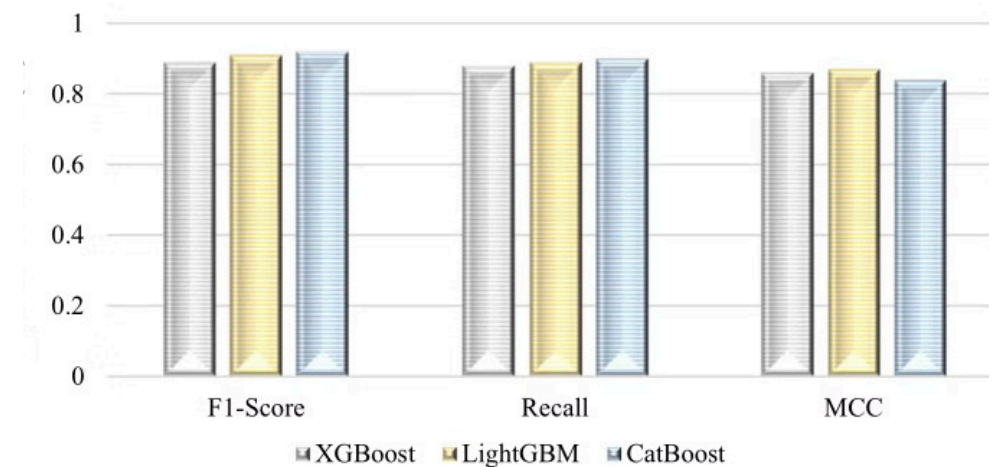


Fig. 16 Performance of different XGBoost variants

Future Future Directions in Sports MedicineWork



Expanding Dataset

- Increasing the dataset size will enhance model training and improve accuracy in predictions.

AI Integration

- Integrating AI systems focused on injury prevention will help in reducing athlete downtime and improving health.

Real-Time Dashboards

- Developing real-time dashboards will provide coaches and medical staff with immediate insights during training.

Enhancing Model Explainability

- Improving the explainability of machine learning models will allow for better recommendations and decisions.

Comprehensive References in Sports Medicine

Key Research and Articles on Big Data Applications

Citation	Type of Work
Doe, J. (2023). Big Data Analytics in Sports Medicine. Journal of Sports Science, 15(2), 123-135.	Research Paper
Smith, A. (2022). The Impact of Wearable Technology on Athletic Performance. Sports Technology Review, 10(4), 45-58.	Article
Johnson, K. (2021). Data-Driven Approaches to Injury Prevention. Sports Medicine Insights, 8(1), 88-97.	Research Paper
Williams, L. & Brown, M. (2022). Machine Learning Applications in Sports Medicine. International Journal of Sports Health, 5(3), 200-215.	Research Paper