International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 04 | Apr 2025 www.irjet.net p-ISSN: 2395-0072

Olympic Medal Prediction Using Python

Shivalingappa R. Tippa[1], Amogh S. khot[2], Soumya R. Lankal[3] , Vikas Timmanagoudar[4], P. K. Deshpande[5]

[1234]*Student, Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkote, India*
[5]*Assistant Professor, Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkote, India*

Abstract - *This project presents a machine learning model developed in Python to predict the probability of athletes winning medals in the Olympics. Using historical data and performance metrics such as age, nationality, prior achievements, and event specifics, the model aims to classify athletes into categories: gold, silver, bronze, or no medal. Various supervised learning algorithms, including logistic regression, decision trees, and random forests, were applied to train and evaluate the model. The system also incorporates data preprocessing techniques such as feature selection, normalization, and handling missing values to enhance prediction accuracy. The model's performance was assessed through metrics such as accuracy, precision, and recall, using cross-validation on the dataset. Additionally, a user-friendly web interface was developed to allow input of athlete data, providing real-time predictions. This system aims to serve as a useful tool for analyzing and forecasting athletic success in future Olympic Games, with potential applications in sports analytics and strategic planning.*

Key Words: Olympic Medal Prediction, Machine Learning, Sports Analytics, Historical Data, Performance Forecasting.

1. INTRODUCTION

In the 21st century, where data and technology drive decision-making, predicting Olympic medal outcomes has become increasingly valuable. The Olympic medal prediction system analyzes historical performances, socio-economic indicators, and athlete data to forecast a country's potential success, making it easier to manage large datasets.

Manual tracking of such data is time-consuming and error-prone, making traditional methods inefficient. By using machine learning algorithms, the Olympic medal prediction system simplifies this process, providing accurate forecasts.

This approach allows sports organizations and nations to make data-driven decisions, enhancing strategic planning and athlete development for future Olympic competitions.

2. OBJECTIVES

1. To develop a predictive model using machine learning that accurately forecasts Olympic medal counts for participating countries based on historical and socio-economic data.

2. To analyze key success factors, including GDP, population, sports funding, and previous performance, and evaluate their impact on a country's likelihood of winning medals.

3. To provide insights that assist sports organizations and policymakers in optimizing resource allocation and developing targeted training programs to enhance athletic performance.

4. To create a user-friendly interface that allows users to view predictions, interact with data, and explore different factors influencing Olympic success.

5. To encourage data-driven decision-making among sports organizations, coaches, and athletes by providing insights that guide training and strategic decisions for future Olympics.

3. MOTIVATION

The motivation behind Olympic medal prediction stems from the desire to harness data and technology to better understand athletic performance and outcomes. As the Olympics is a global stage for elite athletes, predicting medal results offers valuable insights for athletes, coaches, and analysts. By analyzing vast amounts of historical data and performance metrics, we can uncover patterns that influence success, helping athletes improve their strategies and training. Furthermore, these predictions foster a deeper connection between sports enthusiasts and the games by offering data-driven forecasts. As machine learning techniques evolve, the potential to refine and enhance prediction models grows, making it an exciting challenge to predict the unpredictable. The ultimate motivation is to blend sports with technology, providing a more analytical and objective approach to understanding Olympic success and empowering athletes and professionals to achieve peak performance.

4. LITERATURE SURVEY

1. A. Szmygin, M. Wojtowicz, Ż. Świderska-Chadaj and R. Roszczyk, "Prediction of athletes' performance results using machine learning algorithms," 2023 24th International Conference on Computational Problems of Electrical Engineering (CPEE), Grybów, Poland, 2023, pp. 1-5.

Key Findings:-

- A custom dataset with 17 features, including weather, competition locations, and athlete details, was created to predict 100-meter sprint performance.
- A multi-layer perceptron (MLP) machine learning model was used to predict sprint times based on the dataset.
- The MLP model achieved a 78% prediction accuracy with a tolerance of 0.13 seconds, showing promising results.

2. Z. Bo, Q. Chaoling, X. Xiaoli and Z. Fanbo, "GM (1,1) Model Gray Prediction for the Gold-Medal Result of Women's Put Shot in the 30th Olympic Games," 2011 International Conference on Future Computer Science and Education, Xi'an, China, 2011, pp. 334-337.

Key Findings:-

- The GM (1,1) gray prediction model was applied to forecast the gold medal result in the women's shot put event at 2012 London Olympics.
- The GM (1,1) model effectively predicts with limited and uncertain data, making it suitable for sports performance prediction the with sparse historical data.
- The authors demonstrated the GM (1,1) model's prediction accuracy, showing it provides a reliable forecast for the women's shot put event.

3. C. Thirumalai, S. Monica and A. Vijayalakshmi, "Heuristics prediction of Olympic medals using machine learning," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2017, pp. 594-597.

Key Findings:-

- Machine learning algorithms (decision trees, SVM) predict Olympic medal outcomes using historical data and key factors.
- Heuristic methods enhance predictions by considering economic conditions, population size, and past performances.

The combined approach of machine learning and heuristics ensures efficient and accurate forecasting of medal results.

4. G. R. LeTu, "A Machine Learning Framework for Predicting Sports Results Based on Multi-Frame Mining," 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2022, pp. 810-813.

Key Findings:-

- The paper introduces a machine learning framework that employs multi-frame mining techniques for predicting sports results.
- It demonstrates the effectiveness of combining historical performance data with multi-frame analysis to enhance prediction accuracy.
- The proposed framework shows promising results in forecasting sports outcomes, contributing to more reliable and efficient sports prediction models.

5. V. Asha, S. P. Sreeja, B. Saju, N. C S, P. G. N and A. Prasad, "Performance Analysis of Olympic Games using Data Analytics," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1436-1443.

Key Findings:-

- The paper uses data analytics to analyze performance trends in the Olympic Games, focusing on athlete performance, event popularity, and historical results.
- .It identifies key factors influencing Olympic success, including training patterns, competition strategies, and athlete preparation.
- The study offers valuable insights to optimize strategies for future Olympic Games and improve performance prediction models.

5. PROBLEM FORMULATION

Introduction

In the context of Olympic medal prediction, the problem formulation involves predicting the likelihood of an athlete or country winning a gold, silver, bronze, or no medal in various Olympic events based on available data. The key challenge lies in developing a reliable and accurate model that can process historical data, such as previous Olympic performances, world rankings, athlete attributes (age, experience, and physical characteristics), and external factors like weather conditions and competition specifics.

The problem can be approached as a classification task where the goal is to categorize the outcome of an athlete's participation into one of four classes: gold, silver, bronze, or no medal. The problem formulation requires addressing several factors, including how to handle missing data, the selection of relevant features, and the choice of appropriate machine learning algorithms to train the model.

Furthermore, the unpredictability of factors such as injuries, psychological states, and external influences adds complexity to the problem. Thus, the challenge is not only about using data-driven techniques to predict outcomes but also about overcoming the uncertainties inherent in competitive sports. The final model aims to offer valuable insights for athletes, coaches, and analysts to optimize performance and enhance predictions for future Olympic events.

Present System

The current systems for Olympic medal prediction use data analytics and machine learning to forecast the outcomes of athletes in the Olympic Games. These systems rely on historical data, athlete attributes, and event-specific factors to predict whether an athlete will win gold, silver, bronze, or no medal.

The foundation of these systems is built on comprehensive data, including past Olympic results, world rankings, athlete performance metrics, and external variables such as weather conditions, training history, and venue-specific factors. Athlete demographics, like age, experience, and injury history, are also considered.

Modern systems employ various machine learning algorithms, including classification models (e.g., decision trees, random forests, logistic regression) and more sophisticated methods like neural networks. These algorithms analyze patterns and relationships in the data to predict medal outcomes. The models are trained on historical performance data and are capable of classifying athletes into categories of gold, silver, bronze, or no medal.

Despite advancements, the current models face challenges in accuracy due to the unpredictable nature of sports. Factors like injuries, mental states, and other unforeseen events can heavily influence performance. Additionally, certain success determinants, such as mental resilience or psychological factors, cannot be accurately captured by data alone.

Nevertheless, these prediction systems provide valuable insights for athletes, coaches, analysts, and fans, offering data-driven forecasts to better understand and strategize for the Olympic Games.

Problems of Present System

Unpredictability of Sports Performance: Many factors influencing an athlete's performance are difficult to quantify, such as mental state, motivation, and psychological factors. These elements can significantly impact outcomes, making predictions less accurate.

Incomplete or Inaccurate Data: The quality and completeness of the data used for predictions are critical. Missing data or inaccuracies in historical performance records, athlete health status, and external factors like weather conditions can skew results and reduce the model's reliability.

Dynamic Nature of Athletes' Performance: Athletes' performance can fluctuate over time due to factors such as injuries, changing coaching strategies, or new competitors. Predictive models based on past performances may not fully account for these dynamic changes, leading to misleading forecasts.

External Influences: External factors such as unexpected injuries, psychological states, referee decisions, and weather conditions can have a substantial impact on Olympic events. These factors are difficult to model, adding uncertainty to predictions.

Over fitting and Generalization Issues: Some machine learning models may become too specific to the historical data they are trained on (overfitting), reducing their ability to generalize to new events and competitions.

Limited Data on Non-quantifiable Factors: Intangible elements like an athlete's mental resilience, their response to pressure, or team dynamics are difficult to capture in data, leaving the model incomplete.

Proposed System

The proposed system for Olympic medal prediction takes a data-driven approach, focusing on enhancing the accuracy and depth of predictions by leveraging machine learning algorithms that can process complex and multifaceted relationships. Unlike traditional systems that primarily focus on historical data, the proposed system incorporates a wide range of factors that influence an athlete's chances of winning a medal. Key elements of the system include historical performance data, which is crucial in identifying trends and patterns from previous Olympics. This is combined with economic indicators, such as national sports funding, which can influence athlete preparation and access to better facilities. The system also takes into account athlete-specific data—factors like age, training history, injury status, and psychological resilience, which are often overlooked in current models. Additionally, the system integrates recent competition trends, analysing results from other international events leading up to the Olympics. Through extensive feature engineering, the system creates a rich, comprehensive dataset that considers these interconnected variables. This allows the machine learning models to capture the intricate relationships between these factors and their impact on medal outcomes. By utilizing advanced algorithms such as deep learning and ensemble methods, the proposed system aims to provide more precise, reliable, and context-aware predictions, ultimately improving the overall accuracy and understanding of Olympic medal forecasting.

Problem Statement

System is to develop an accurate Olympic medal prediction system that integrates diverse data sources, such as historical performance, athlete profiles, economic indicators, and real-time competition trends. The system must overcome limitations of current models by incorporating complex, dynamic factors and improving prediction accuracy using advanced machine learning techniques.

6. SOFTWARE REQUIREMENT SPECIFICATION

FUNCTIONAL REQUIREMENT

These are specific features, functions, or tasks that a system must perform to satisfy the user needs. They define the behavior of the system and describe what it should do.

Functional Requirements of Olympic Medal Prediction using Python are:-

Data Collection: Gather historical data, athlete profiles, competition results, and external factors (weather, economic indicators).

Data Pre-processing: Clean and normalize data, extract relevant features for prediction.

Model Development: Use machine learning models (e.g., deep learning, ensemble learning) to predict medal outcomes.

User Interface: Create a web interface for displaying predictions, real-time updates, and interactive features.

Output: Generate medal predictions with confidence scores and provide insights or recommendations for athletes.

Performance Evaluation: Continuously evaluate model accuracy and adjust predictions based on new data.

NON-FUNCTIONAL REQUIREMENT

These are quality attributes or constraints of the system that define how the system performs its functions. They describe system properties like performance, scalability, usability, and reliability.

Non-Functional Requirements of Olympic Medal Prediction using Python are:-

Performance: The system should provide predictions in real-time, with minimal latency. It should handle a large volume of data efficiently, especially during live updates and event data processing.

Scalability: The system should be scalable to accommodate growing amounts of data and users, particularly during the Olympics when there is a surge in data volume and user traffic.

Availability: The system should be available 24/7, especially during live competitions, with minimal downtime to ensure predictions are accessible at all times

Reliability: The system must provide consistent and accurate predictions, with the ability to handle errors gracefully and provide fallback mechanisms in case of data inconsistencies or model failure.

Security: The system should implement robust security measures to protect sensitive data, including user information, athlete profiles, and competition data. User authentication and data encryption should be applied to ensure privacy.

Usability: The web interface should be user-friendly, with easy navigation and clear visualizations of predictions and insights. It should be accessible across devices (desktops, tablets, mobile phones).

Maintainability: The system should be easy to maintain, with clear code documentation and modular architecture to allow for future updates or changes in features.

Extensibility: The system should be designed to allow easy integration of new data sources or machine learning models in the future, without requiring major overhauls.

Compliance: The system should comply with relevant regulations and standards related to data protection, such as GDPR for user privacy and data handling.

1. Hardware and Software Requirements

Hardware Requirements

These specify the physical devices and components needed to run a system, application, or software effectively. They include servers, user devices, network equipment, storage devices, and other hardware resources.

Hardware requirements of Olympic Medal Prediction using Python are: -

Server: Multi-core CPU (Intel Xeon or AMD Ryzen), 16–32 GB RAM, 500 GB SSD storage, and a dedicated GPU (e.g., Nvidia Tesla) for machine learning tasks.

Networking: High-speed internet connection for real-time data processing and user access.

Backup: Cloud-based backup storage and UPS for data protection.

User Access Devices: Web servers for load balancing and client devices like desktops, laptops, and smart phones.

Database Server: Scalable DBMS (MySQL, PostgreSQL) for data storage.

Cloud Infrastructure (Optional): Cloud services (AWS, Google Cloud, Azure) for scalability and fault tolerance.

Software Requirements

These define the programs, operating systems, tools, and frameworks required to develop, deploy, and operate the system. They include the necessary software for coding, hosting, testing, and maintaining the application.

Software requirements of Olympic Medal Prediction using Python are:-

Operating System: Windows, Linux, or macOS.

Programming Languages: Python (for machine learning), JavaScript (React/Angular for UI), HTML/CSS.

Machine Learning Libraries: Scikit-learn, TensorFlow/Keras, PyTorch.

Web Development: Node.js, Bootstrap, Flask/Django.

Database: MySQL, PostgreSQL, or NoSQL.

Version Control: Git.

Cloud Services: AWS, Google Cloud, or Azure (optional).

Data Processing & Visualization: Pandas, NumPy, Matplotlib, Seaborn.

7. DESIGN

Architecture

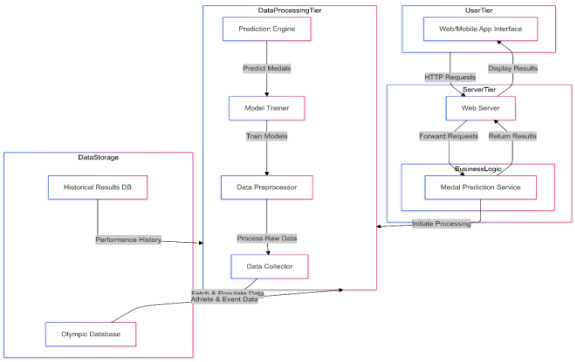

Fig. 7.1

The above figure represents the architecture of Olympics Medal prediction using Python

The architecture of Olympic Medal Prediction using Python provides a space for predicting the medal winners. It involves:

Data Storage: Holds athlete and event data in a Historical Results DB and Olympic Database.

Data Processing: Collects, pre-processes, and trains models to predict medals using the Prediction Engine.

User Interface & Server: A web or mobile app sends requests to a web server, which interacts with the prediction service to display results.

Data Flow Diagrams

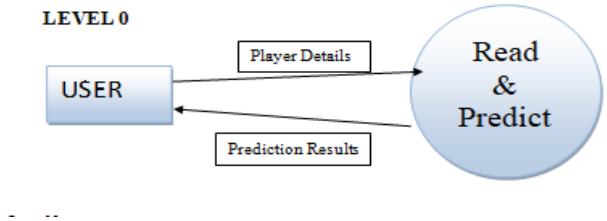

Fig. 7.2

Level0:
The context diagram shows the interaction between the user and the system Users provide player details for predicting the medal.

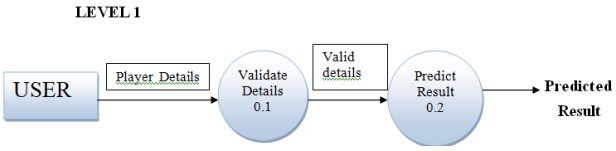

Fig. 7.3

Level1:
The system validates player details. After successful validation, the user can interact with the prediction module for predicting the medal.

Use Case Diagram

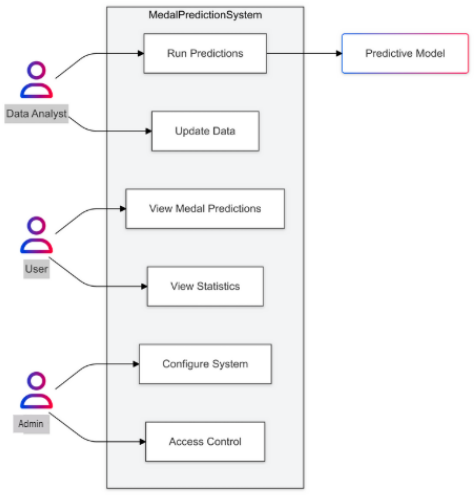

Fig. 7.4 Use Case Diagram

The above figure represents the use case of Olympics Medal prediction using Python

8. IMPLEMENTATION

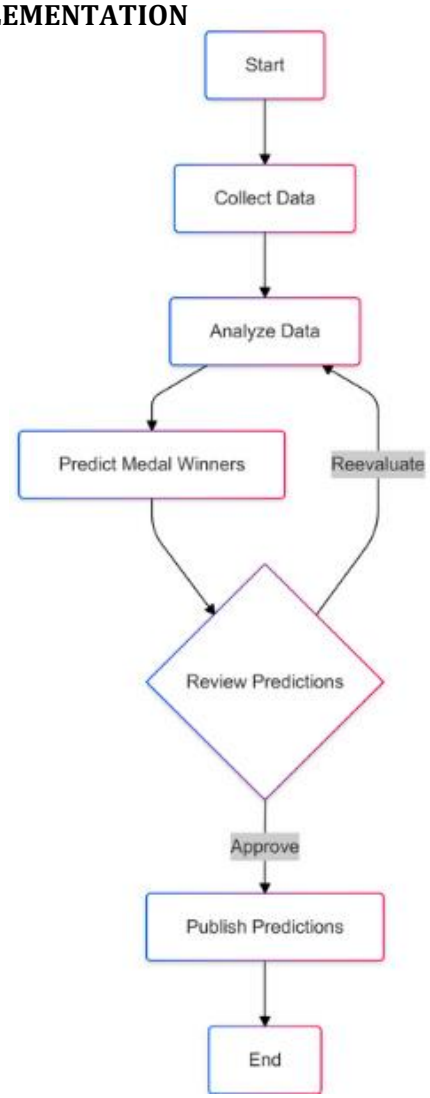

Fig. 8.1 Flowchart

The above figure represents the architecture of Olympic Medal Prediction using Python.

Tools

1. **Front-End Tools**

HTML: For structuring web pages and content.

CSS: For styling and designing the layout of the website.

Bootstrap: A CSS framework for responsive design and faster development.

JavaScript: For adding interactivity and dynamic behaviour to web pages.

React.js or Angular: JavaScript frameworks for building interactive user interfaces.

2. **Back-End Tool**

Programming Languages: Python (for machine learning and logic), Node.js (for real-time data handling).
Web Frameworks: Flask (lightweight), Django (feature-rich).
Database: MySQL/PostgreSQL (for structured data), MongoDB (NoSQL for unstructured data).
Machine Learning: Scikit-learn, TensorFlow/Keras, PyTorch.
Version Control: Gi

These tools collectively ensure the seamless functionality, interactivity, and reliability of the project, making it user-friendly and efficient for its purpose.

9. TESTING

Test Cases
Test Case 1: Invalid Age.

- **Input:** User enters the age below 18 and above 40.
- **Expected Result:** Enter the valid age.

Test Case 2: Invalid Sport.

- **Input:** User enters invalid sport name.
- **Expected Result:** Enter the valid sport name.

Test Case 3: Invalid Athlete name.

- **Input:** User enters invalid athlete name.
- **Expected Result:** Invalid athlete name

Test Case 3: Rank Metrics.

- **Input:** User enters the athlete's rank above 10.
- **Expected Result:** No medal is displayed.

10. RESULTS

- **Home Page**

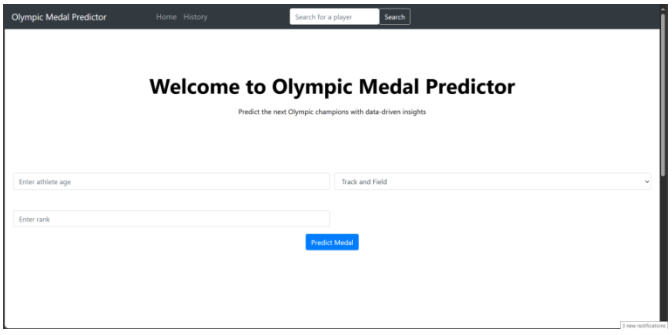

Fig 10.1.1 Home Page

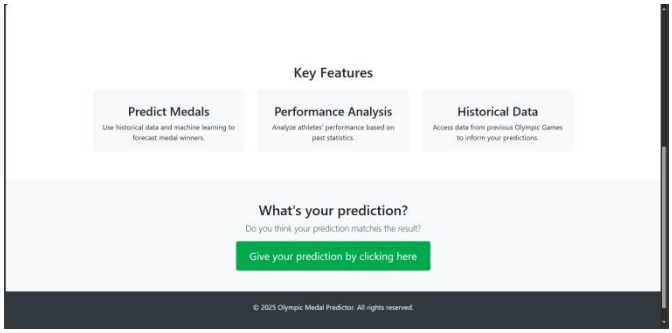

Fig 10.1.2 Home Page

The above figures represent the Home Page of the application.

- **Predicted Result Page**

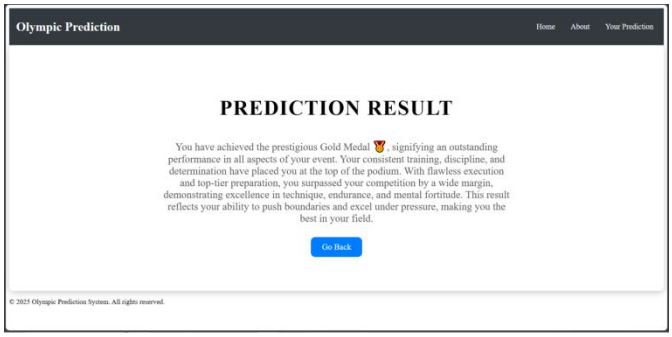

Fig 10.2 Predicted Result Page

The above figure represents the Predicted Result Page of the application.

- **Players Profile Page**

Fig 10.3 Players Profile Page

The above figure represents the Players Profile Page of the application.

- **User Prediction Page**

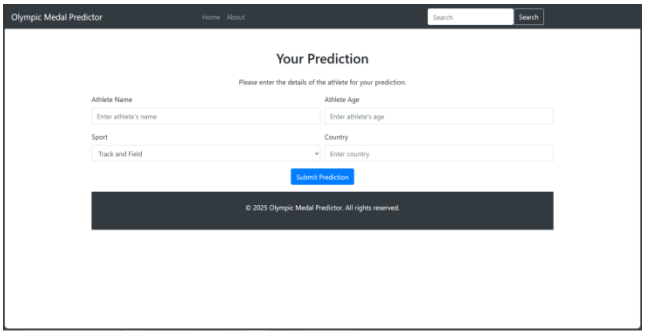

Fig 10.4.1 User Prediction Page

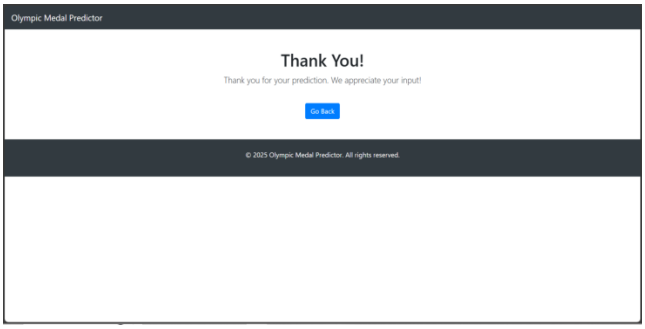

Fig 10.4.2 User Prediction Page

The above figures represent the User Prediction Page of the application.

- **Olympics History Page**

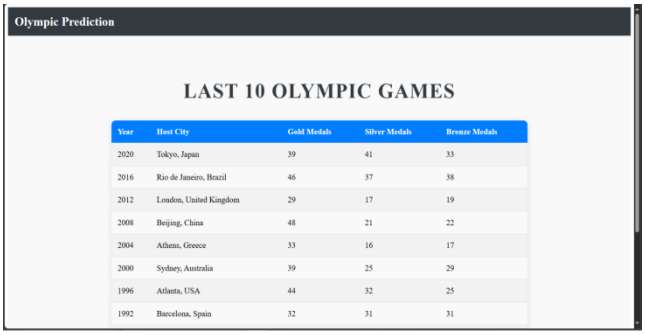

Fig 10.5 Olympic History Page

The above figure represents the Olympic History Page of the application.

11. CONCLUSIONS

The Olympic medal prediction system leverages advanced machine learning techniques and data-driven approaches to

provide more accurate and reliable medal forecasts based on historical performance and real-time data. By integrating diverse datasets, such as athlete performance, economic indicators, and competition trends, the system aims to enhance prediction precision compared to traditional methods. The proposed system focuses on scalability, high availability, and efficient data processing, ensuring seamless performance during large-scale events like the Olympics. With the implementation of secure authentication, interactive web interfaces, and robust back-end infrastructure, this system not only meets the demands of real-time predictions but also offers a user-friendly experience. Ultimately, the system has the potential to revolutionize how we analyze and predict outcomes in large sporting events, bringing deeper insights into the world of sports analytics. Users to make a positive impact on both their reading habits and the world.

REFERENCES

[1] A. Szmygin, M. Wojtowicz, Ż. Świderska-Chadaj and R. Roszczyk, "Prediction of athletes' performance results using machine learning algorithms," 2023 24th International Conference on Computational Problems of Electrical Engineering (CPEE), Grybów, Poland, 2023, pp. 1-5.

[2]Z. Bo, Q. Chaoling, X. Xiaoli and Z. Fanbo, "GM (1,1) Model Gray Prediction for the Gold-Medal Result of Women's Put Shot in the 30th Olympic Games," 2011 International Conference on Future Computer Science and Education, Xi'an, China, 2011, pp. 334-337.

[3]G. R. LeTu, "A Machine Learning Framework for Predicting Sports Results Based on Multi-Frame Mining," 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2022, pp. 810-813.

[4]C. Thirumalai, S. Monica and A. Vijayalakshmi, "Heuristics prediction of Olympic medals using machine learning," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2017, pp. 594-597.

[5] V. Asha, S. P. Sreeja, B. Saju, N. C S, P. G. N and A. Prasad, "Performance Analysis of Olympic Games using Data Analytics," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1436-1443.

[6] P. Badoni, P. Choudhary, C. P. Rudesh and N. T. Singh, "Predicting Medal Counts in Olympics Using Machine Learning Algorithms: A Comparative Analysis," 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech), Banur, India, 2023, pp. 116-121.

[7] Praveen Badoni, Priya Choudhary, Challa Parvathi Rudesh, NongmeikapamThoiba Singh, "Predicting Medal Counts in Olympics Using Machine Learning Algorithms: A Comparative Analysis", 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech).

[8] Chandra Segar Thirumalai, Monica Sankar, Vijayalakshmi A, "Heuristics Prediction of Olympic Medals using Machine Learning", IEEE - International Conference on Electronics, Communication and Aerospace Technology ICECA2017.