**Proposal: Utilizing Neural Networks to Predict Men's 100m Rankings at the Paris Olympics**

**1. Introduction**

The objective of this project is to build a predictive model for the Paris Olympics 100-meter final rankings. By leveraging historical performance data, physiological metrics, and other relevant factors, we aim to create a robust and accurate predictive model.

**2. Background and Motivation**

Accurately predicting sports outcomes, especially for the 100-meter sprint, aids strategic planning and enhances competitive insights.Utilizing Web Mining to gather relevant data and constructing robust predictive models will help forecast rankings in the Paris Olympics 100-meter final, providing valuable insights for training and strategy.

**3. Objectives**

1. **Data Collection:** Gather historical performance data of male 100m sprinters, including past Olympic results, World Championships, and other major athletic events.
2. **Feature Engineering:** Identify and preprocess relevant features such as wind conditions, elevated conditions and the probability that an athlete may cause a false-start.
3. **Model Development:** Use traditional analytical models to confirm the accuracy of features in predicting final rankings, choose machine learning models such as neural networks to observe whether the prediction is better, and introduce NDCG (Normalized Discounted Cumulative Gain) methods to enhance accuracy.
4. **Evaluation:** Validate the model using various performance metrics and cross-validation techniques.

**4. Methodology**

1. **Data Collection and Preprocessing:**
2. **Target Variable (y)**: Predicted rankings for the Paris Olympics
3. **Potential Influencing Factors (X)**:

Physiological Indicators: Age

Historical Performances: Best performance of the year (net); Best performance international ranking of the year; Best net performance international ranking of the year; Championship Year (0/1); Average of top 10 performances of the year; International ranking of average top 10 performances of the year; Variance of top 10 performances of the year; Personal average of top 10 performances; Personal variance of top 10 performances; Number of times participated in championships; (Number of false starts)

**(3)Data preprocessing:**

* + Identify and fill in missing values (e.g., missing wind speed data) 
  + Quantify data through secondary processing (e.g., age, net performance)

1. **Feature Selection:**
   * Exploratory Data Analysis (EDA):Identify important features . And analyze the correlation between potential influencing factors and final rankings.
   * Utilize techniques such as Principal Component Analysis (PCA) to reduce dimensions if necessary.
2. **Model Architecture:**

**Traditional Models: **

* + Validate the predictive ability of selected features

**Machine Learning Models: **

* + Design a neural network architecture suitable for regression tasks, incorporating dense layers, dropout layers, and batch normalization layers 
  + Experiment with different activation functions, optimizers, and loss functions to improve model performance

**Model Optimization: **

* + Introduce the NDCG (Normalized Discounted Cumulative Gain) method for optimization

1. **Training and Validation:**
   * Split the data into training, validation, and test sets.
   * Train the model using the training set and tune hyperparameters using the validation set.
   * Evaluate the model on the test set using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared.

**5. Expected Outcomes**

* Construct a model that can predict the ranking of athletes in the Paris Olympics more accurately.
* Explore and verify the important factors affecting the ranking of athletes.
* Propose model optimization methods to improve the prediction accuracy.
* Predict the interval range of athletes' performance and evaluate the possibility of breaking the world record. (\*)

**6. Timeline**

**A screenshot of a graph

Description automatically generated**

**7. Conclusion**

This study aims to develop a robust predictive model for the 100-meter final rankings at the Paris Olympics by leveraging data analysis and machine learning techniques. With a well-defined methodology and a clear timeline, we are confident in our ability to deliver a robust predictive model that can provide valuable insights to athletes, coaches, and sports analysts.