Cricket Code Champions Hackathon

Aniket B. Barphe

Manager-Data Science, India First Life Insurance Company

Player Performance Prediction for World Cup Matches

https://www.linkedin.com/in/aniiketbarphe/

Outline

Introduction

1) Team Name :- *aniiket_barphe*

2) Member Name:- Aniket B. Barphe

3) About Member:- Dedicated and results-driven Data Scientist with a strong background in both the Insurance and Meteorology sectors. As a member of the Data Science team at IndiaFirst Life Insurance, I contribute my expertise to an organization backed by esteemed entities like Bank of Baroda, Union Bank of India, and Carmel Point Investments India Private Limited.

4) Theme:- Player performance prediction for World Cup matches.

Problem Statement

❖ The goal of hackathon is to create an advanced machine learning solution that leverages data science models and techniques. This solution aims to provide accurate predictions for individual player performances, specifically predicting the runs scored and wickets taken.

❖ This prediction will be based on a comprehensive historical dataset that covers both player statistics and team performance. The ultimate objective is to forecast how each player will perform in the highly anticipated ICC World Cup 2023 match scheduled for October 8, 2023.

Image Credit:- Bing, DALL E3

Solution

❖ The problem-solving process comprises two main phases: Data Preprocessing and Modeling.

❖ Data preprocessing involves optimizing the dataset for machine learning by conducting statistical analyses, examining data shape, addressing missing values, and assessing unique value counts.

❖ In the modeling phase, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model is employed, particularly effective for time series data. This model ensures accurate predictions for individual player performances, enhancing the overall effectiveness of the approach.

❖ The solution provides accurate predictions for player performances, offering strategic insights to teams. Evaluation metrics like prediction accuracy and correctness will assess its impact on ICC World Cup 2023 matches.
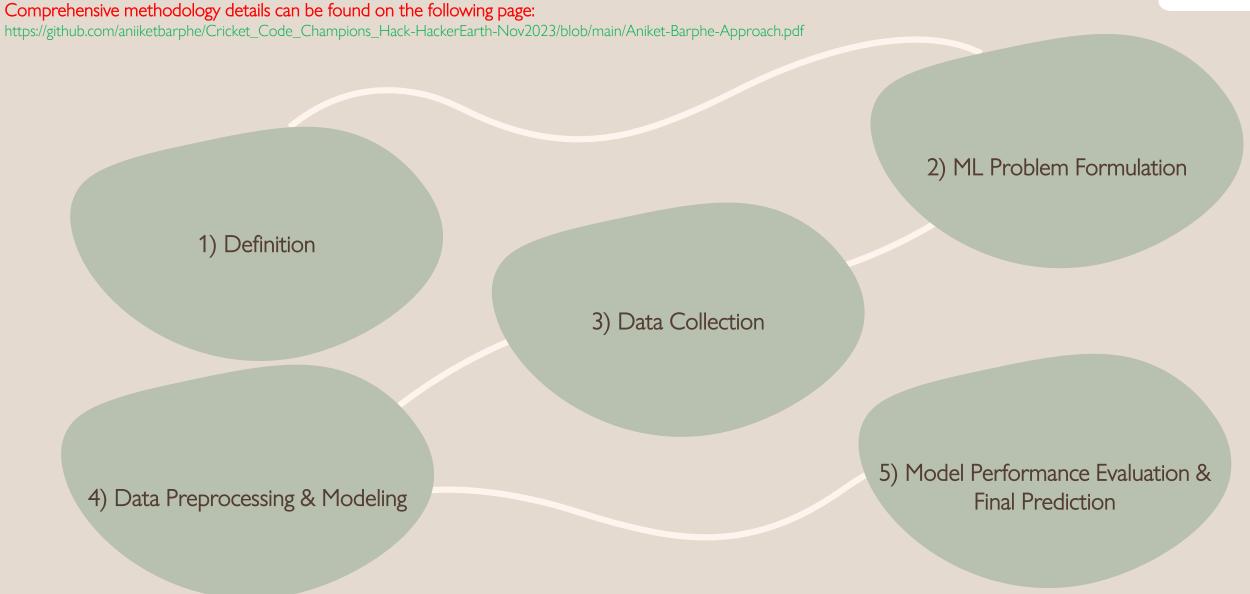
Cont.

❖ Python, along with data science and machine learning libraries like Pandas, NumPy, Scikit-learn, and Matplotlib, will be employed for an effective and efficient solution implementation.

❖ Key considerations for the project involve the quality and availability of historical datasets, prediction accuracy, and the teams' capacity to execute the recommended strategies. Assumptions, constraints, and decision points play crucial roles in shaping the overall effectiveness of the proposed solution.

❖ The proposed solution, utilizing mentioned frameworks and tools, is anticipated to deliver accurate player performance predictions, aiding teams in strategic decisions. With scalability, it can be applied to various similar applications, ensuring effectiveness and adaptability for future use.

Prototype Submission Phase-Aniket Barphe

Methodology

2) ML Problem Formulation

1) Definition

3) Data Collection

4) Data Preprocessing & Modeling

5) Model Performance Evaluation & Final Prediction

Prototype Submission Phase-Aniket Barphe

Working Prototype

Explore the detailed Python script for making predictions with the Machine Learning Model located at the following path:

https://github.com/aniiketbarphe/Cricket_Code_Champions_Hack-HackerEarth-Nov2023/blob/main/Player_Performance_Prediction_Fr_World-Cup_Matches.ipynb

Thank You