Tournament Simulation Report

Project Title: AI-Powered Football Tournament Simulation
Developed By: Muhammad Kumail
Tools Used: Python, Pandas, XGBoost, NumPy

Project Overview

This project aims to **simulate a football tournament** using an **AI model trained with historical match data**. The model predicts match outcomes based on team performance metrics, and randomness is introduced for realistic variability.

Data Preparation

- **Datasets Used:**
  - matches.csv → Contained **30k historical matches**
  - teams.csv → Mapped teams to their confederations
  - qualified.csv → Contained the **16 teams** that advanced to the knockout stage
- **Feature Engineering:**
  - **Winning Form** (Last 10,20,30,40,50 Matches)
  - **Losing Form** (Last 10,20,30,40,50 Matches)
  - **Goal Difference Average**
  - **Form Against Confederation Teams**
  - **Home & Neutral Venue Effects**
- **Data Processing Steps:**
  Cleaned missing values
  Encoded team names for model training
  Merged team statistics with match data

AI Model Training

- **Model Used:** XGBoost (XGBClassifier)
- **Target Variable:**
  - 1 → Team 1 Wins
  - 2 → Team 2 Wins
  - 0 → Draw
- **Performance Metrics:**
  - **Training Accuracy:** 74.8%
  - **Feature Importance:**

Feature importance
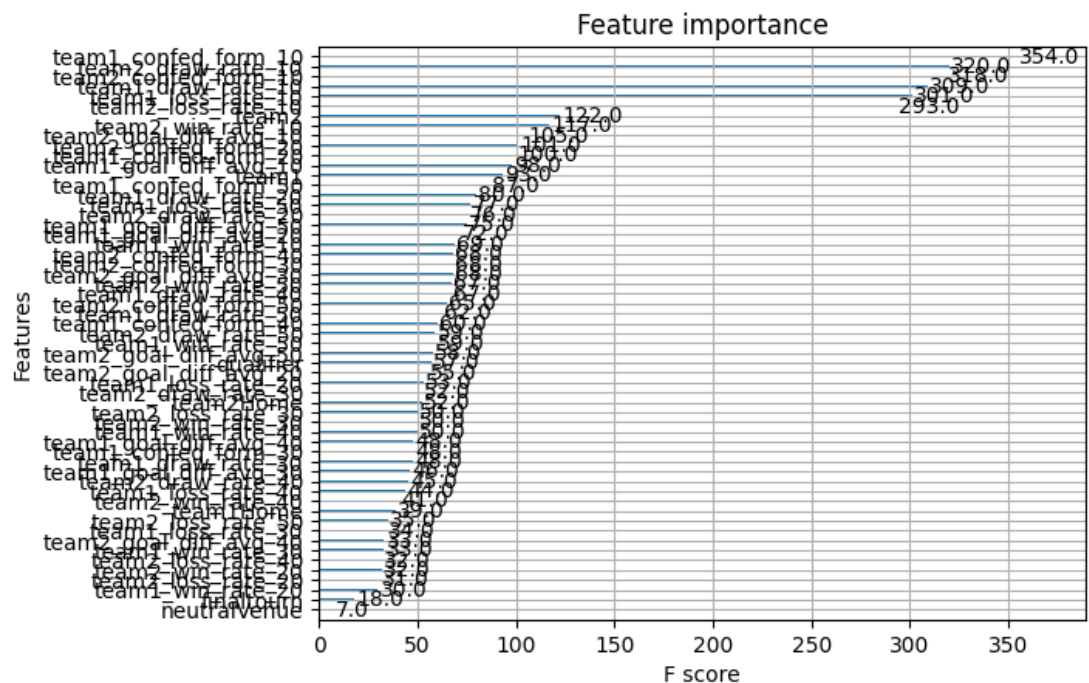
  -
- **Randomness Handling:**
  - If a **team wins convincingly (P > 50%)**, it wins outright.
  - If no clear winner, **randomly choose between a draw and the losing team**.

Tournament Structure & Simulation

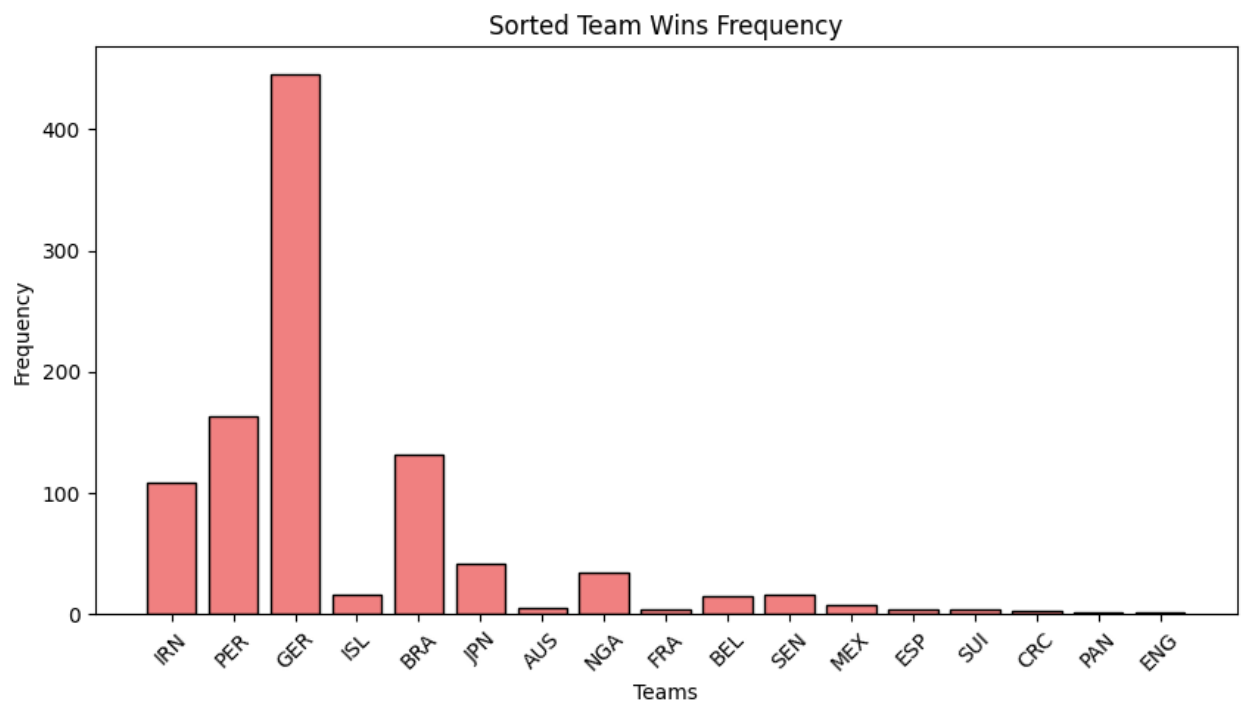### Group Stage

- Teams played **simultanious matches** within groups.
- **3 points** for a win, **1 point** for a draw, **0 points** for a loss.
- Top **2 teams from each group** advanced to the knockout stage.

Knockout Stage

- **Round of 16 → Quarterfinals → Semifinals → Final**
- **Single-elimination format**
- **Penalty Shootout Handling:** If a match was drawn, a **random winner was chosen** to simulate penalties.

Tournament Results & Analysis

Sorted Team Wins Frequency

Biggest Upsets: England being Eliminated in Group Stage 547/1000 iterations

- **Insights from the Simulation:**
  Teams with strong form and confederation strength performed well.
  Randomness allowed for occasional upsets, making results more realistic.
  Fine-tuning randomness threshold (0.5) balanced realism and predictability.

Future Improvements & Next Steps

More Factors for Realism: Injuries, fatigue, travel effects

Better Upset Handling: Adjust randomness for penalty shootouts

Live Match Commentary: Generate match events dynamically

Predicting Goal Differentials to perfectly simulate realistic results

Conclusion

This project successfully **simulated a football tournament using AI-powered predictions** while introducing controlled randomness for realism. **The methodology balances deterministic AI logic with real-world variability**, making it a practical tool for match simulations.