

ANALYZING DETERMINANTS OF TEAM SUCCESS IN THE ENGLISH PREMIER LEAGUE

Reproducible Research project

Turgud Valiyev
Akshma Atreja
Marco Ayuob
Muhammad Ibrahim Hanif

Why We Chose This Topic

Football is globally loved and highly data-rich.



The EPL is one of the most competitive leagues.

We are passionate about combining data science with sports.

There's a practical demand from analysts, bettors, and clubs.

Research Objective & Problem Statement Objective:

To analyze which factors significantly impact EPL match outcomes and team success.

Key Research Questions:

What metrics most influence success in the EPL?

How well can we predict outcomes using machine learning models?

Do these metrics change in importance over time?



Data Overview



25 years of match data from the English Premier League
Covers seasons from 2000 to 2025
Sourced from: football-data.co.uk

Key Variables:

Match outcomes (win, draw, loss)

Goals (Full-time, Half-time)

Home/Away team stats (shots, corners, fouls, cards)

Disciplinary data (yellow/red cards)

Market odds and betting lines

Asian handicaps & over/under odds

Tools & Modeling Approach

Tools Used

Python – for data preprocessing, modeling, and evaluation

Jupyter Notebooks – for reproducible and collaborative work

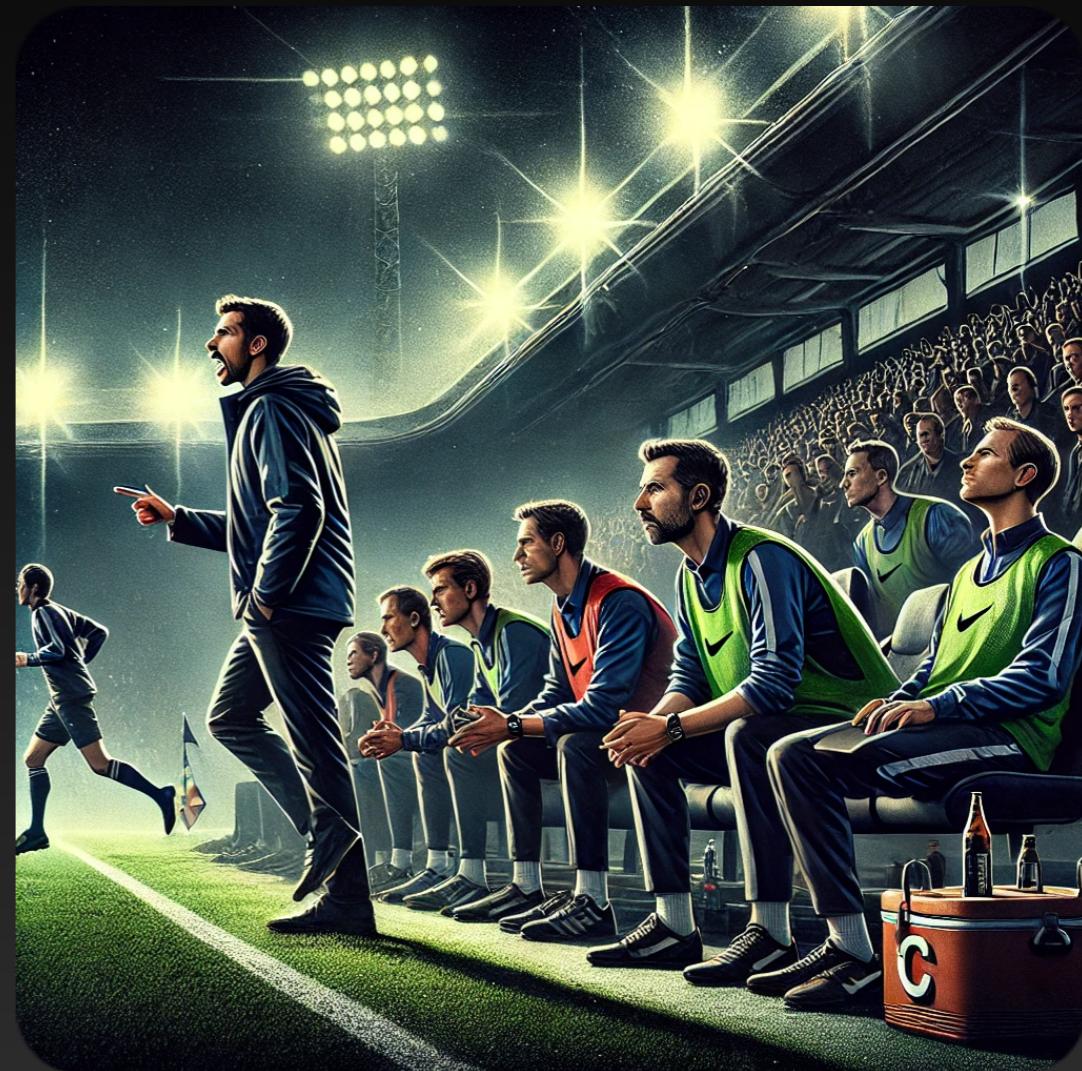
Explainable AI (XAI) – to understand model decisions

Models Applied

Linear Models: Logistic Regression

Non-Linear Models: Random Forest, XGBoost

SHAP (SHapley Additive exPlanations): for interpreting feature influence



Team Responsibilities

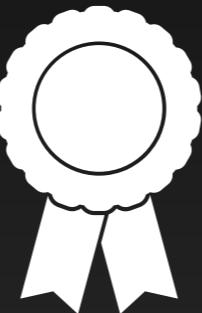
Turgud Valiyev

Team lead, machine learning modeling, presentation delivery



Akshma Atreja

Data collection & visualization, match odds analysis



Muhammad Ibrahim Hanif

Results interpretation, fairness check, and documentation

Marco Ayuob

Feature engineering, model tuning, evaluation metrics

Impact & Relevance

Why This Matters

Enhances strategic decision-making in football

Provides insights for clubs, analysts, and bettors

Demonstrates how machine learning supports sports analytics

Brings transparency via explainable AI



A wide-angle shot of a large, modern stadium at night. The stadium is brightly lit from within, with numerous spotlights illuminating the dark sky above. The field is a vibrant green with white boundary lines. In the foreground, the silhouettes of spectators are visible against the bright lights of the stadium.

Thank you for your
attention