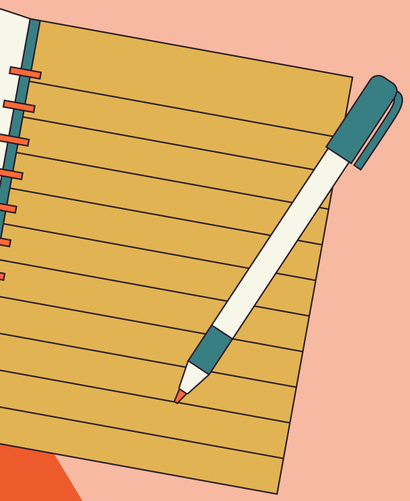
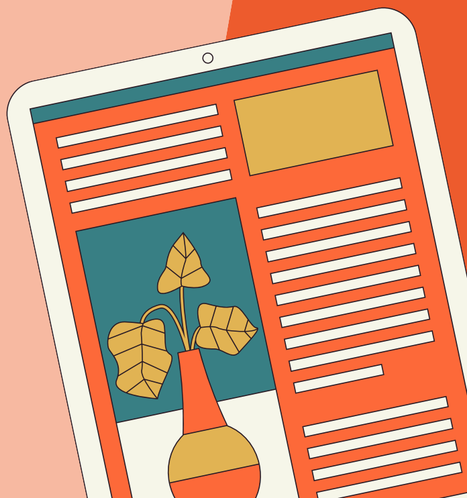
STARC '23
BI-WEEKLY REVIEW

PROJECT NAME

SPORT:	Football
LEADER:	Sukruthi Sanampudi
MEMBERS:	Avani Dhagam Divyansh Vinayak Siddharth Prakash



WEEK #4



BACKGROUND

Monte Carlo simulation is a powerful technique that can be applied to football to model and analyze complex parameters that involve uncertainty and randomness. It involves generating multiple simulations using random inputs to generate a range of possible outcomes.

DELIVERABLES

Data Collection:
Gathering the football dataset from various sources, ensuring data completeness and accuracy

Data Processing:
Cleaning and preprocessing the dataset

Feature Engineering
Selecting relevant features to develop predictive models and optimize strategies.

User Interface Development:
Creating an interface that will provide actionable insights for decision-making

PROGRESS MADE

WEEK 1:

- **Explored research areas**, simulation frameworks, and libraries.
- **Explored datasets**, focusing on key features relevant to the research.

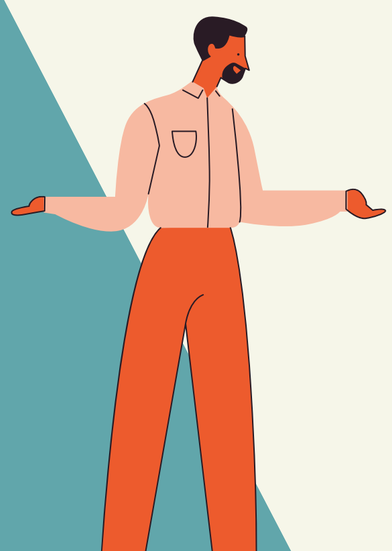
WEEK 2:

- A total of **10 Papers were reviewed** with various abstracts in connection with our topic to increase the overall understanding of the technologies that can be used.
- **Cleaning and Filtering** of the dataset was using SQLlite queries

WEEK 3:

- **Data review and analysis:** understanding the dataset, its structure, and identifying any data quality issues or anomalies
- Analyzed the database and created CSVs of players based on their football positions.
- Utilized **heatmap** and **correlation matrix** to identify attribute effects on the overall rating.

GOALS, OBJECTIVES AND SCOPE



Our major objective for next week is to come with an end to end **naive bayes classifier** machine learning algorithm which can give us the most accurate and the most **optimal top 11** for a specific team based on the various integral parameters that we calculated in the past weeks for each and every position in football

POTENTIAL OBSTACLES

Limited data quality or availability, requiring careful data preprocessing and feature selection.
Complex interactions between variables require advanced modeling techniques and interpretability.
The need for expert domain knowledge to appropriately interpret and validate the results.
Computational complexity and time constraints for large-scale simulations and model
training.
Addressing these obstacles will be crucial to ensure the accuracy and reliability of the
analysis results.

PROJECT APPROVAL

Suggestions

Add suggestions

APPROVED BY	APPROVED BY:
POSITION	POSITION
DD/MM/YYYY	DD/MM/YYYY