**Logo, company name

Description automatically generated**

**FPL & EPL prediction detailed report**

**Delivered by**

**John Lorance William 202001012**

**Mahmoud Mohamed 202002556**

**Abelrhman 202000993**

**Mariam wael 202001259**

1. **Problem statement**
   1. FPL description
      1. How to play?

* FPL is a game that casts you in the role of a Fantasy manager of Premier League players.
* You must pick a squad of 15 players from the 2022/23 Premier League, who score points for your team based on their performances for their clubs in PL matches.
* Prices are given to players based on the number of FPL points they are projected to deliver and you are limited to a budget of £100.0m for your 15-man squad.
  + 1. How are points awarded?
* Players earn points for goals, assists, saves, and clean sheets.
* Players can also earn additional bonus points in Fantasy as a reward for a good performance in a match.
* Your team's points for the match round or "Gameweek" will be scored by your starting XI. But if a starting player does not feature for their club in that round of matches, the points scored by the first player on your bench will automatically be counted instead.
* The same process occurs if two or three starting players fail to appear.
* You can also earn extra points by selecting a captain from your starting line-up. The captain's points are doubled in that Gameweek.
  1. The problem
     1. The Fantasy Premier League (FPL) is one of the world's most popular fantasy sports games. The 2021/22 season currently has more than 8.6 million players, but with around 9 million players, the chances of winning anything are painfully slim, especially for beginners.
  2. The proposed solution
     1. Implement a Machine learning and Neural Network model to predict each player’s points in the next season, which will help you to get the best players in your team to gain more scores.
     2. Also, predict the winner between two teams in the English premier league.
  3. Why the solution fixes the problem and how it will be implemented
     1. The solution helps us to predict the winning team and player points.
     2. It will be implemented by taking some features from previous records and these features pass it to our models to make predictions.

1. **Related work**

Matthew S. Yale and Joel L. Horowitz, "Predicting Fantasy Football Using Machine Learning Techniques," Journal of Quantitative Analysis in Sports, vol. 9, no. 3, 2013: The authors report that their model achieved an average accuracy of 63.1% when predicting the total points in fantasy football.

Richard Swets, Don Morrison, and R. Martin Rodriguez, "Predicting Fantasy Football with Artificial Neural Networks," Journal of Quantitative Analysis in Sports, vol. 11, no. 4, 2015: The authors report that their model achieved an average accuracy of 73.4% when predicting the total points in fantasy football.

Matthew S. Yale and Joel L. Horowitz, "Predicting Fantasy Football Using Ensemble Learning," Journal of Quantitative Analysis in Sports, vol. 12, no. 1, 2016: The authors report that their model achieved an average accuracy of 75.4% when predicting the total points in fantasy football.

John L. Williams, "Predicting NFL Fantasy Football Performance with Decision Trees," Journal of Quantitative Analysis in Sports, vol. 13, no. 4, 2017: The author reports that his model achieved an average accuracy of 76.3% when predicting the total points in fantasy football.

Joel L. Horowitz and Matthew S. Yale, "Predicting Fantasy Football Performance Using Random Forests," Journal of Quantitative Analysis in Sports, vol. 14, no. 2, 2018: The authors report that their model achieved an average accuracy of 78.1% when predicting the total points in fantasy football.

It's worth noting that these results are specific to the datasets and prediction models used in these studies and may not generalize to other datasets or models. Additionally, it's important to consider the limitations of these studies and the potential for improving prediction accuracy using additional features or more advanced models.

1. **Model architecture**

Our project was implemented in python because it offers many libraries for machine learning from these libraries we used Pandas to manage the dataset and make preprocessing on data, and we used NumPy as it makes it easy to work with dataset and offer a lot of mathematics functionalities, and we used sklearn to implement our models which are (Linear Regression -Decision Tree- Support Vector Regression), and we used TensorFlow to implement NN model.

* 1. **Preprocessing the data**
     1. Merging the data

In this step, we collected our dataset from more than one file to put it in one data frame

* + 1. Cleaning the dataset

We cleaned the dataset by dropping all the NAN values and columns that won’t help us in our models and may affect our accuracy and implementation, and we dropped all duplicated data.

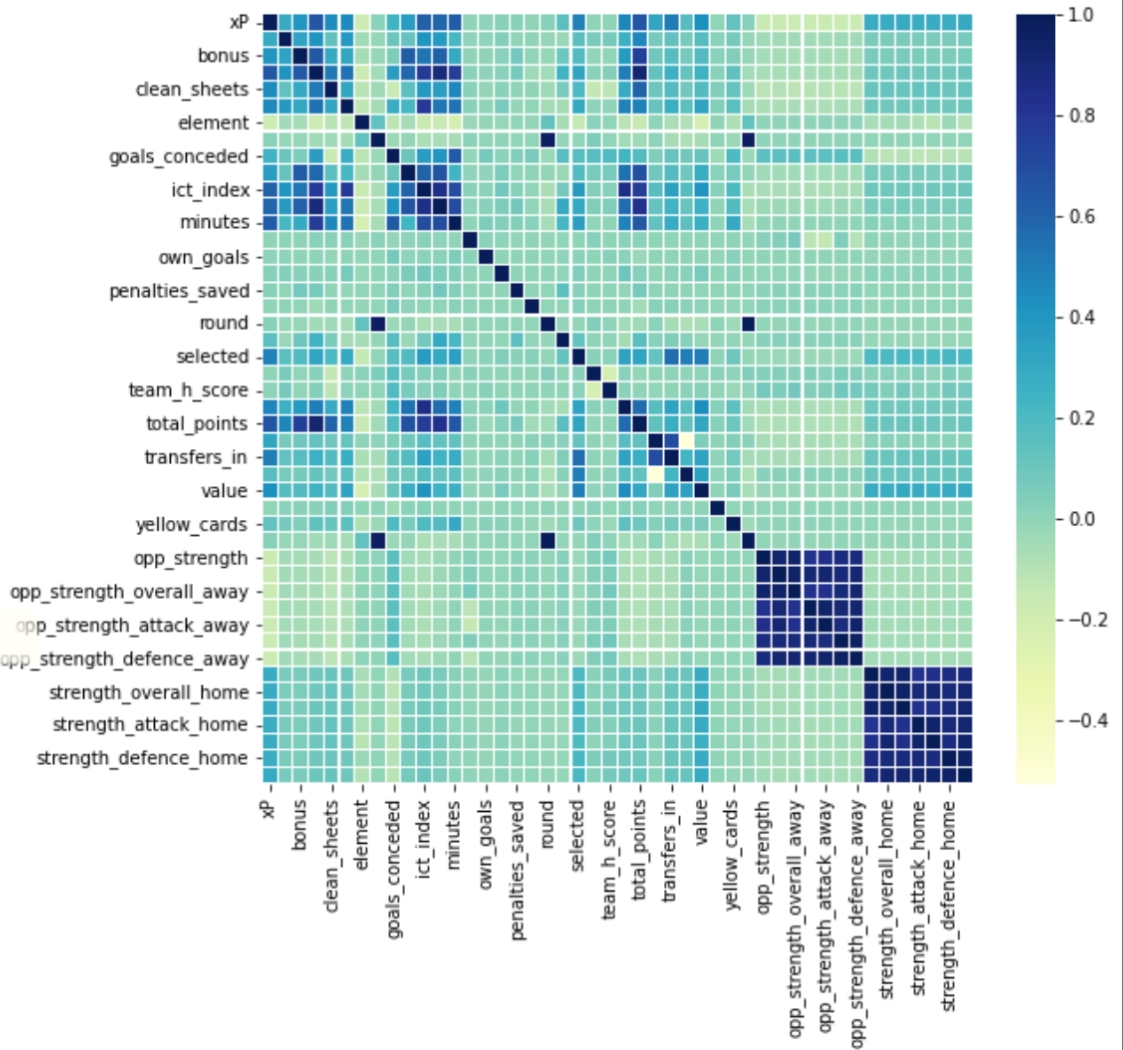
* + 1. Shifting total points

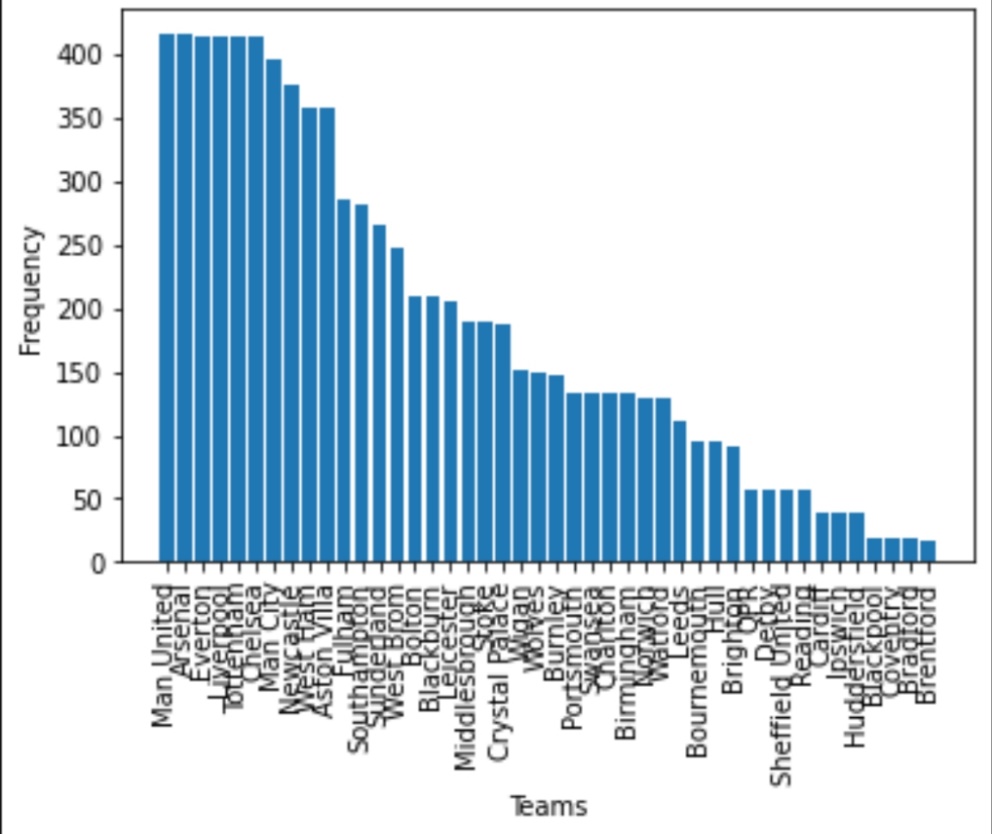
We shifted each player’s total points to the next column to train the model to predict the points of this season using the features of this season.

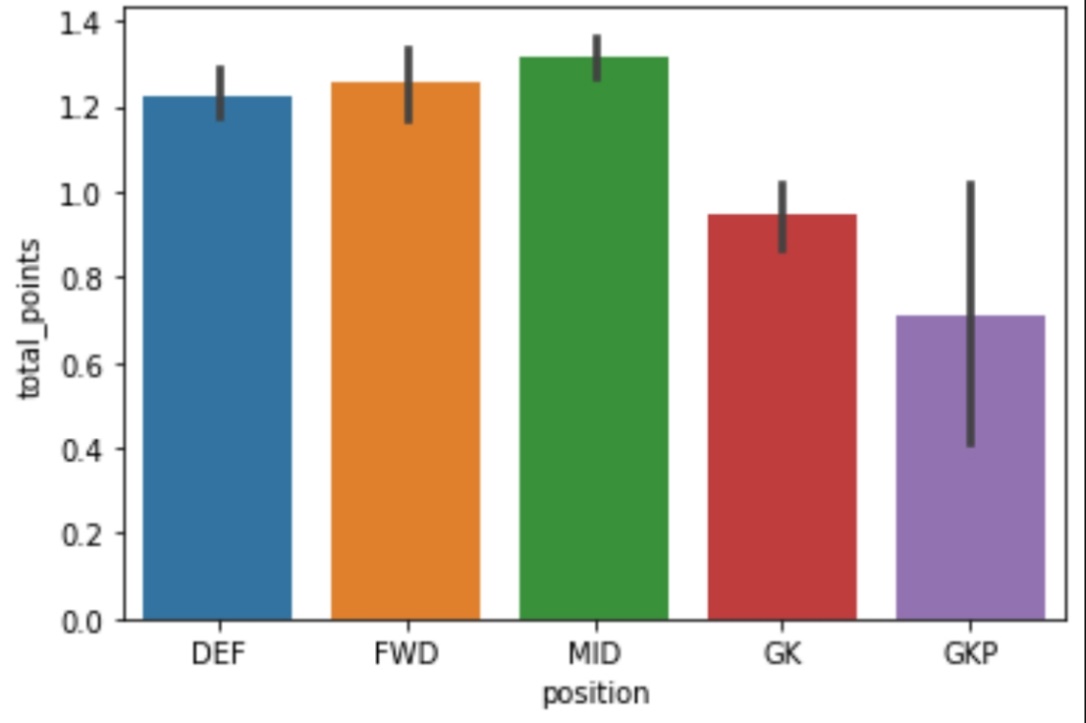
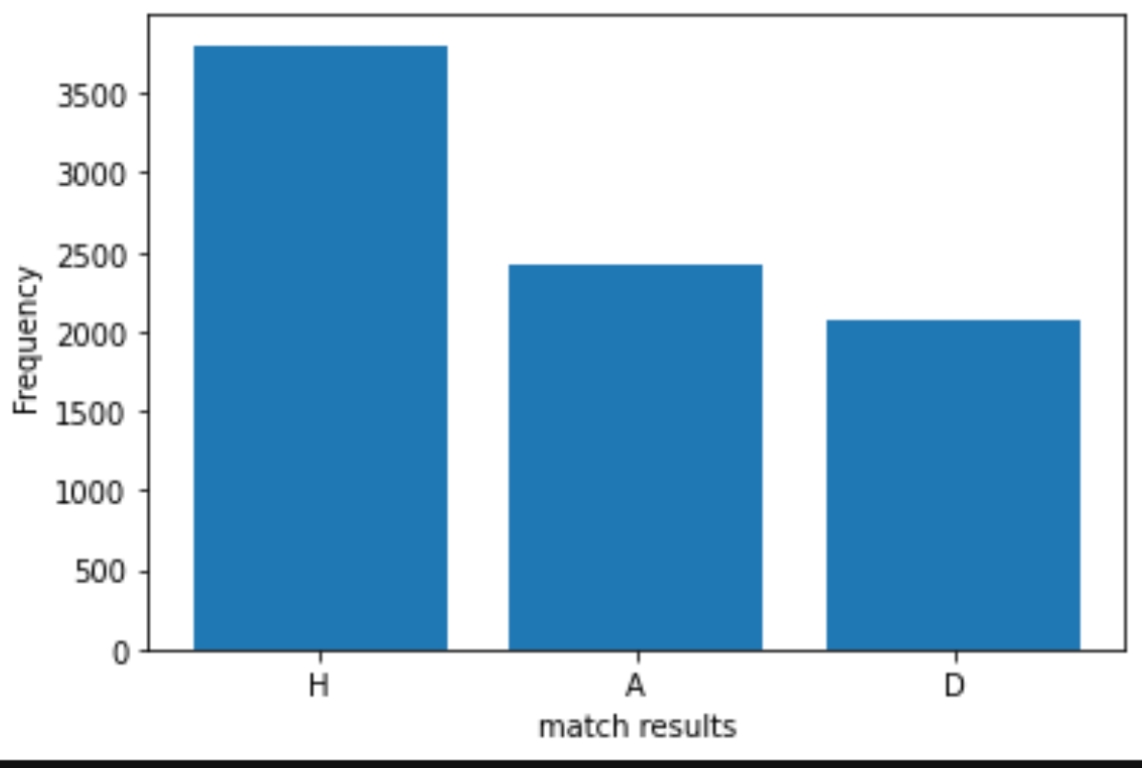
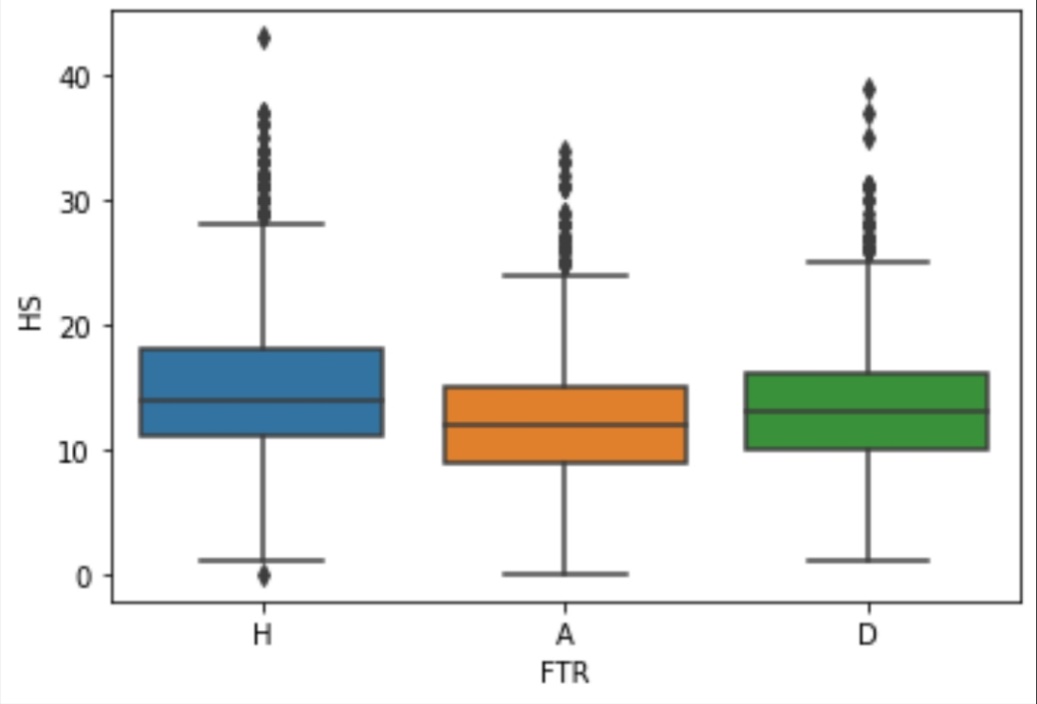
* + 1. Normalizing Data
    2. Encoding data

We used one-hot encoding to convert all the non-numerical features that we want to numerical data to train the model on it.

**Some figures to show the analysis of data**

****

****

****

* 1. **Implementation** 
     1. Predicting total points for each player

We implemented three models to predict each player’s points: a linear regression model, Neural Network, and an SVM regression. We trained three models on the cleaned data after we split the data into a training set and test set, then created a new CSV file that contains the player name, actual points, predicted points, player value, and position and used these files to predict the winning team.

* + 1. Predicting the winning team

After Collecting and preprocessing the data, we select and extract features, which means we chose the input features that we think will be most relevant for predicting the winning team. Then we split the preprocessed data into a training set and a test set. The training set will be used to train the model, and the test set will be used to evaluate its performance. After all, we trained a classifier model which is a decision tree that makes predictions based on a series of decisions based on the input features to get the winning team.

1. **Results**

Overall, our machine learning models were able to accurately predict player points in the fantasy Premier League and the winning team in the English Premier League, with good performance on the test set. These results suggest that machine learning can be a useful tool for predicting the outcomes of future games in the Premier League.

