FANTASY FOOTSCORES

Optimizing FPL Success with Predictive Analytics

By: Mu-izz Gbadamosi

Table Of Content

03 Problem Statement

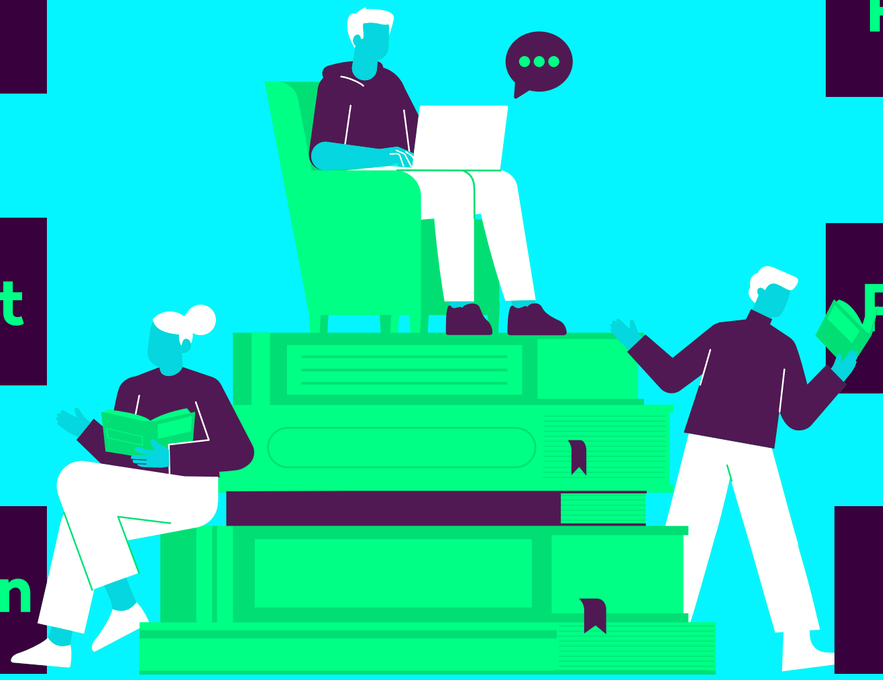
04 Proposed Vision

05 Potential Impact

06 Preliminary EDA

07 Model Evaluation

08 Next Steps

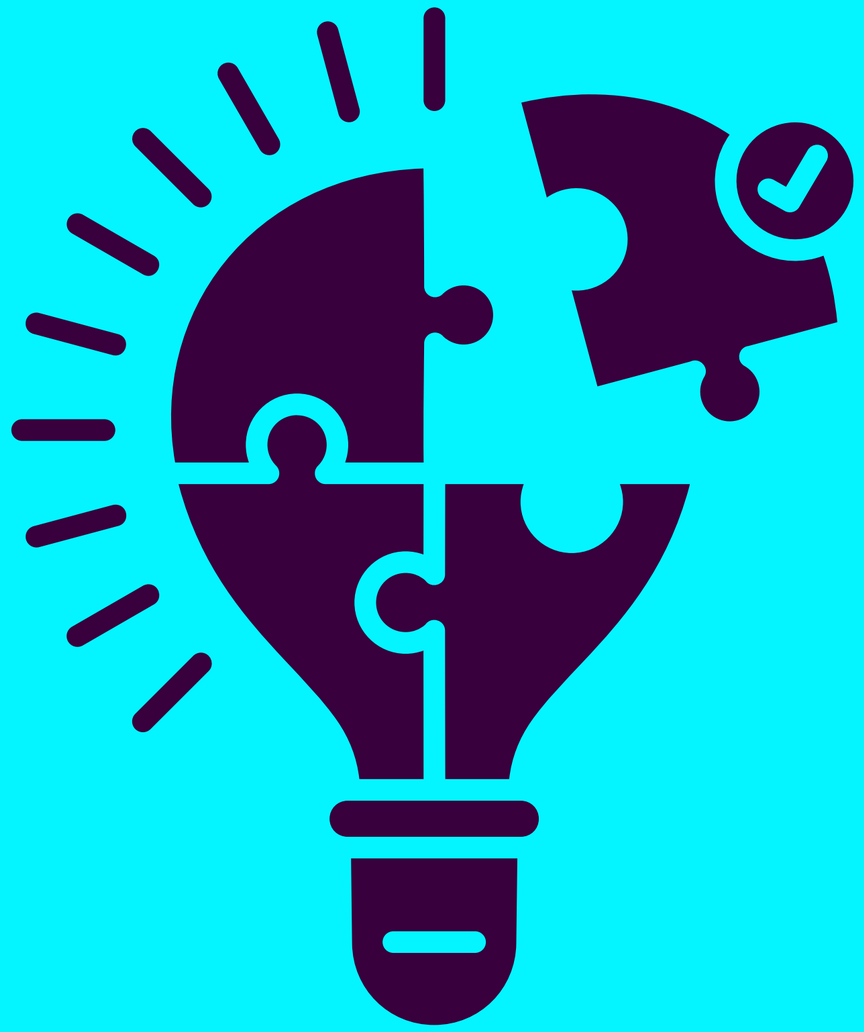


PROBLEM STATEMENT

Managing a successful team can be challenging due to uncertainty in player performance and strategic decision making.

KEY CHALLENGES

- Large Historical Data
- Trend Analysis



PROPOSED SOLUTION

Data Science Approaches

Time Series Forecasting

Enhanced User Experience

Potential Impact

01 Strategic Advantage for FPL Managers

02 Competitive Spirit

03 Platform Growth and Engagement

04 Improve Betting Companies Algorithm

DATASET

- Anand, Vaastav. (2023). FPL Historical Dataset.https://github.com/vaastav/Fantasy-Premier-League/
- 2022/2023 English Premier League Season
- 26505 rows and 41 columns

QUALITY CONCERNS

- Missing gameweek data
- No incremental data for each gameweek
- Rows with no information
- Player Transfers during the season

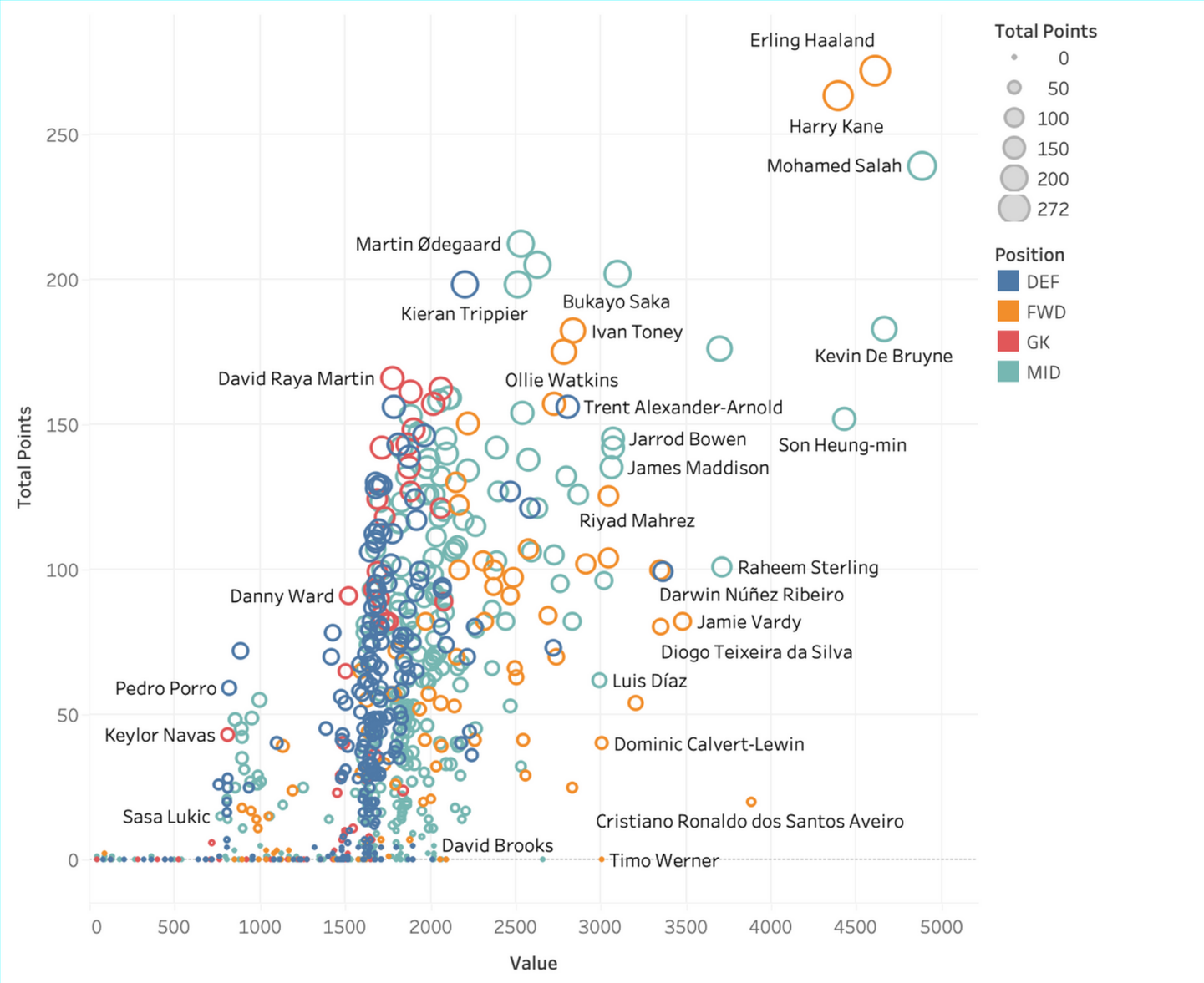
INDEPENDENT VARIABLES

- name
- position
- team
- xP
- assists
- bonus
- bps
- clean_sheets
- creativity
- element
- expected_assists
- expected_goal_involvements
- expected_goals
- yellow_cards
- expected_goals_conceded
- fixture
- goals_conceded
- goals_scored
- ict_index
- influence
- kickoff_time
- minutes
- opponent_team
- own_goals
- penalties_missed
- penalties_saved
- red_cards
- GW
- round
- saves
- selected
- starts
- team_a_score
- team_h_score
- threat
- transfers_balance
- transfers_in
- transfers_out
- value
- was_home

TARGET VARIABLE

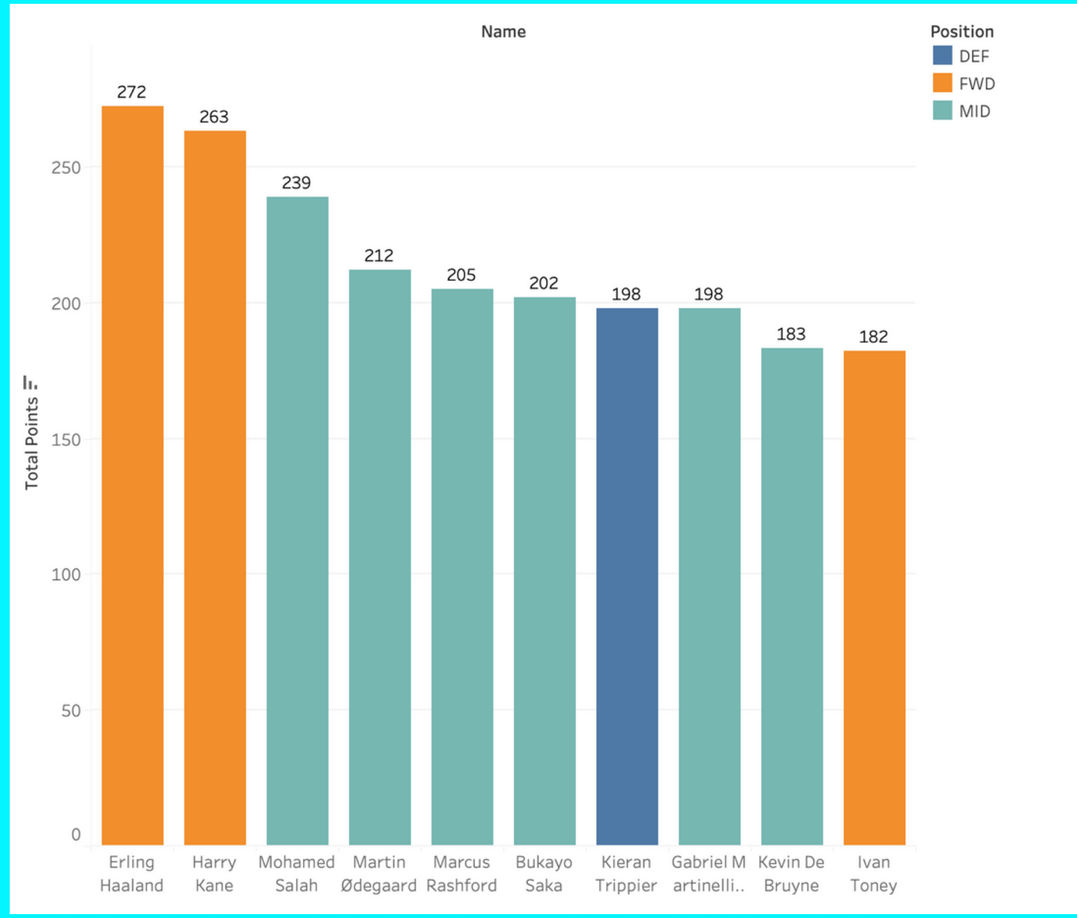
- total_points

PRELIMINARY EDA



TOTAL POINTS VS VALUE

TOP PLAYERS BY POINTS



MODELLING APPROACH

GRADIENT BOOSTING

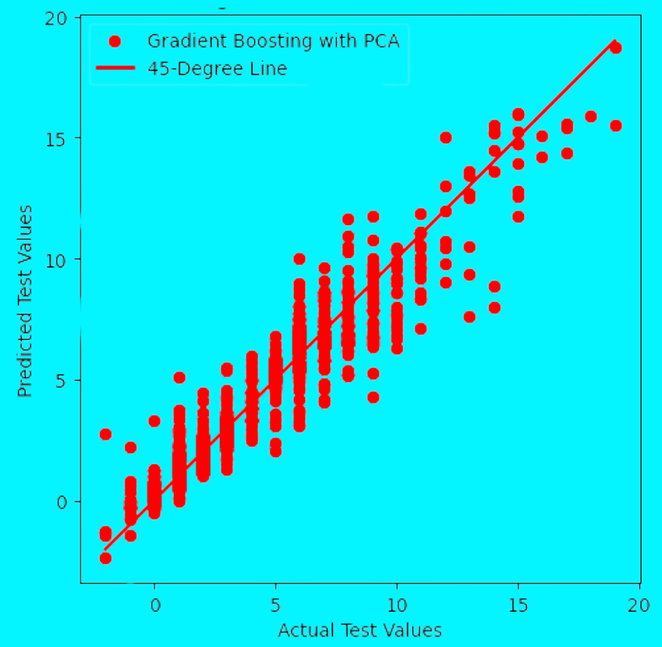
LINEAR REGRESSION

DECISION TREE

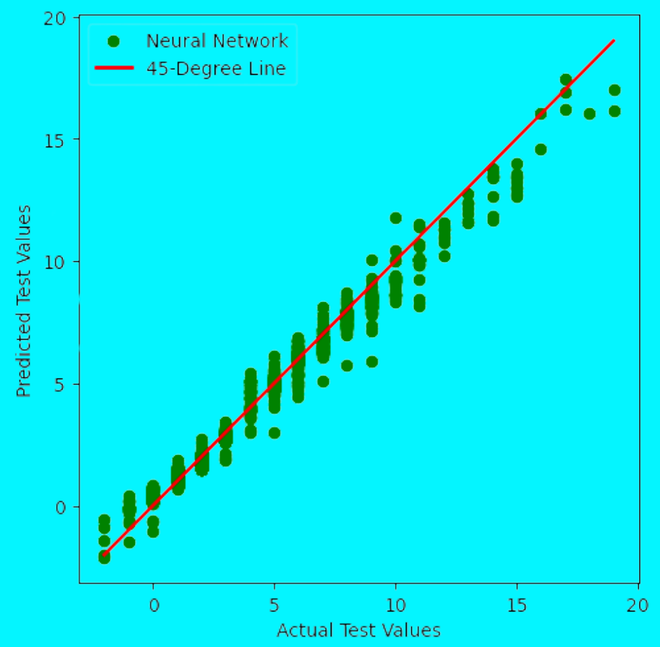
NEURAL NETWORK

ACTUAL VS PREDICTED VALUE

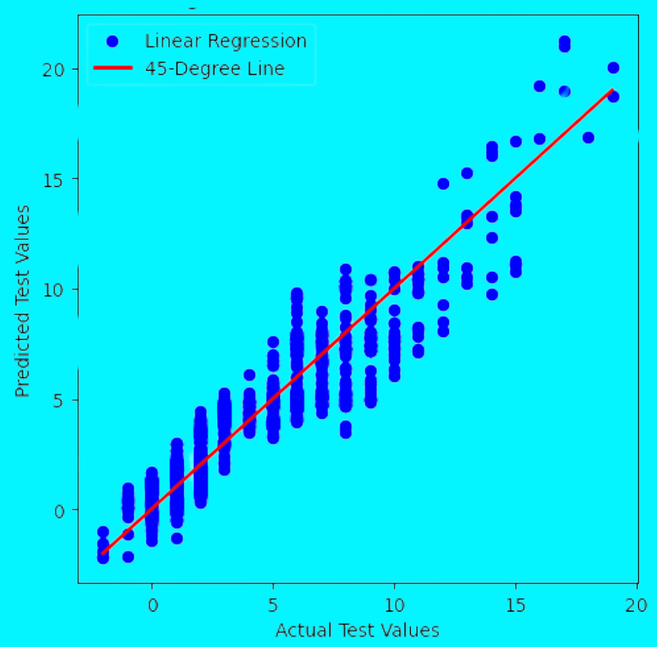
GRADIENT BOOST WITH PCA



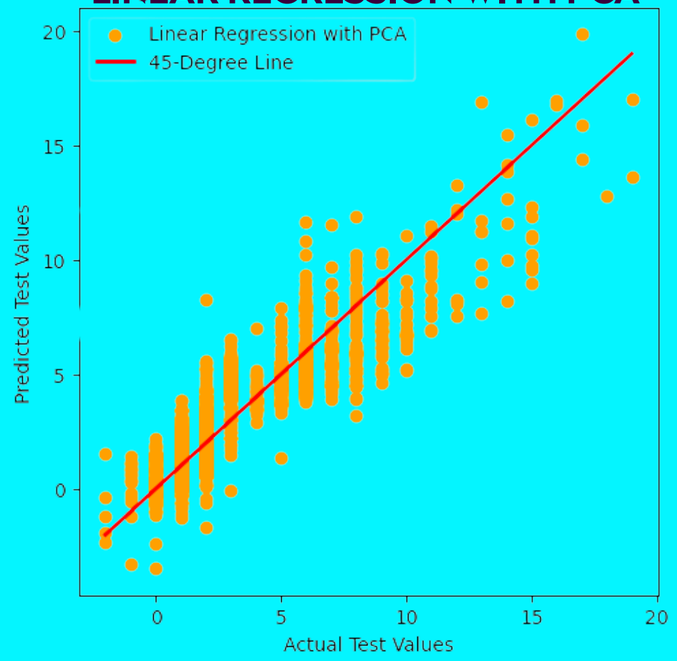
NEURAL NETWORK



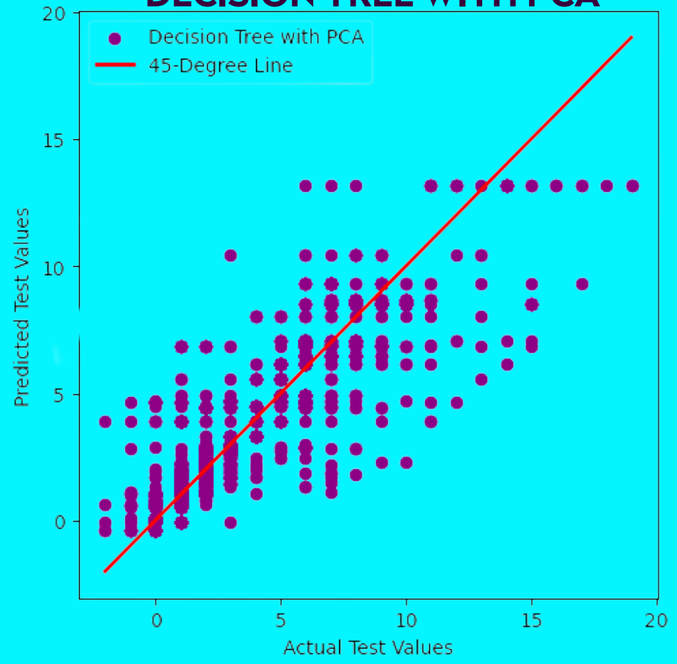
LINEAR REGRESSION



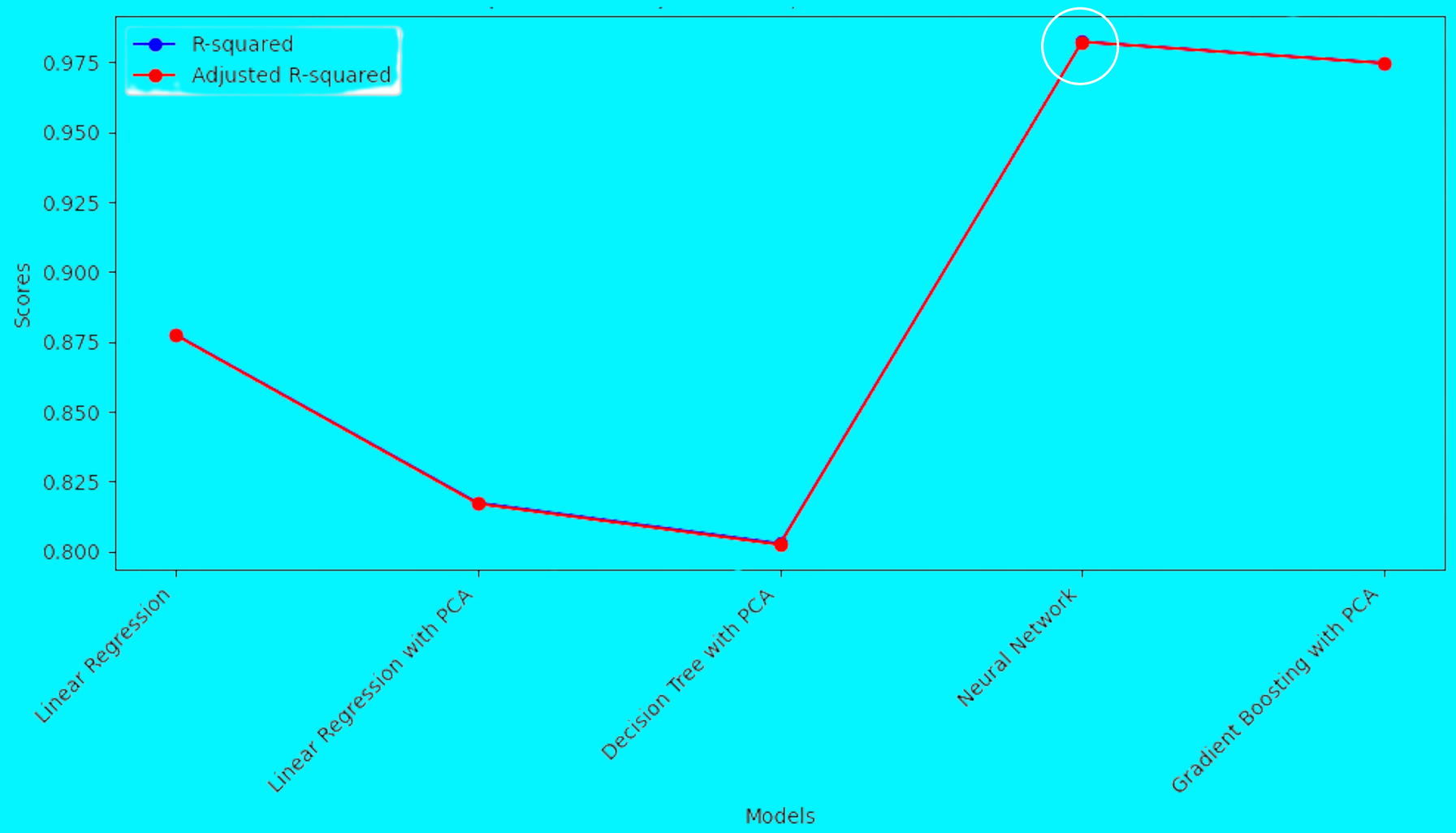
LINEAR REGRESSION WITH PCA



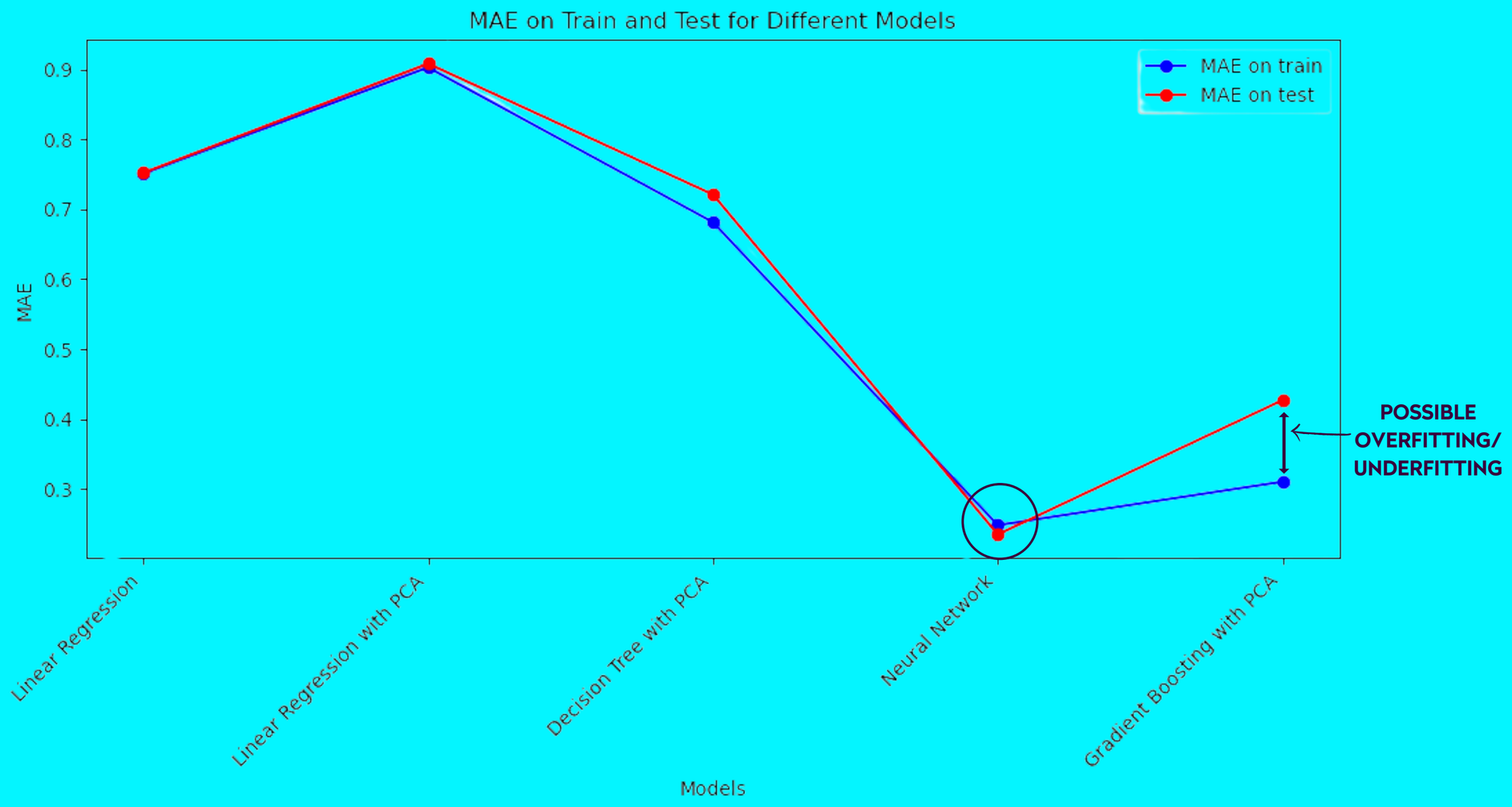
DECISION TREE WITH PCA



MODEL COMPARISON OF R-SQUARED AND ADJUSTED R-SQUARED



MODEL COMPARISON OF MAE FOR TEST AND TRAIN SET

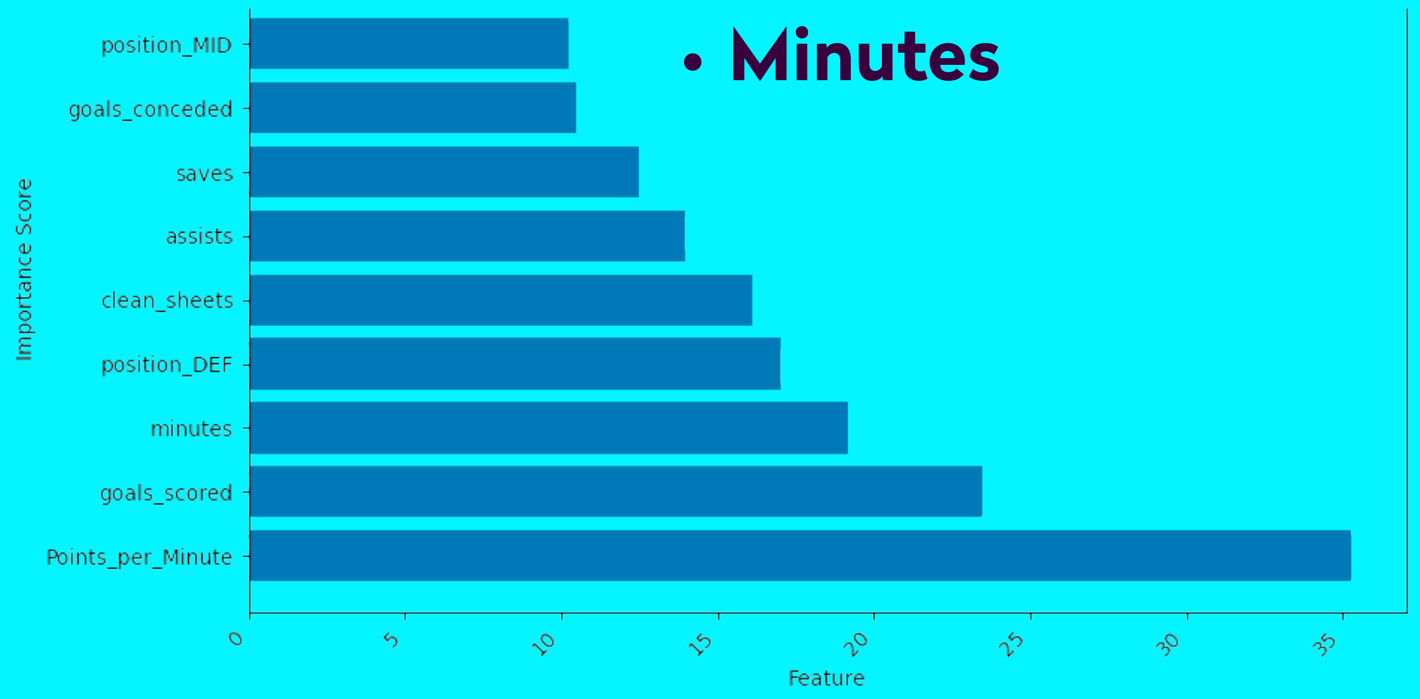


CHOSEN MODEL : NEURAL NETWORK

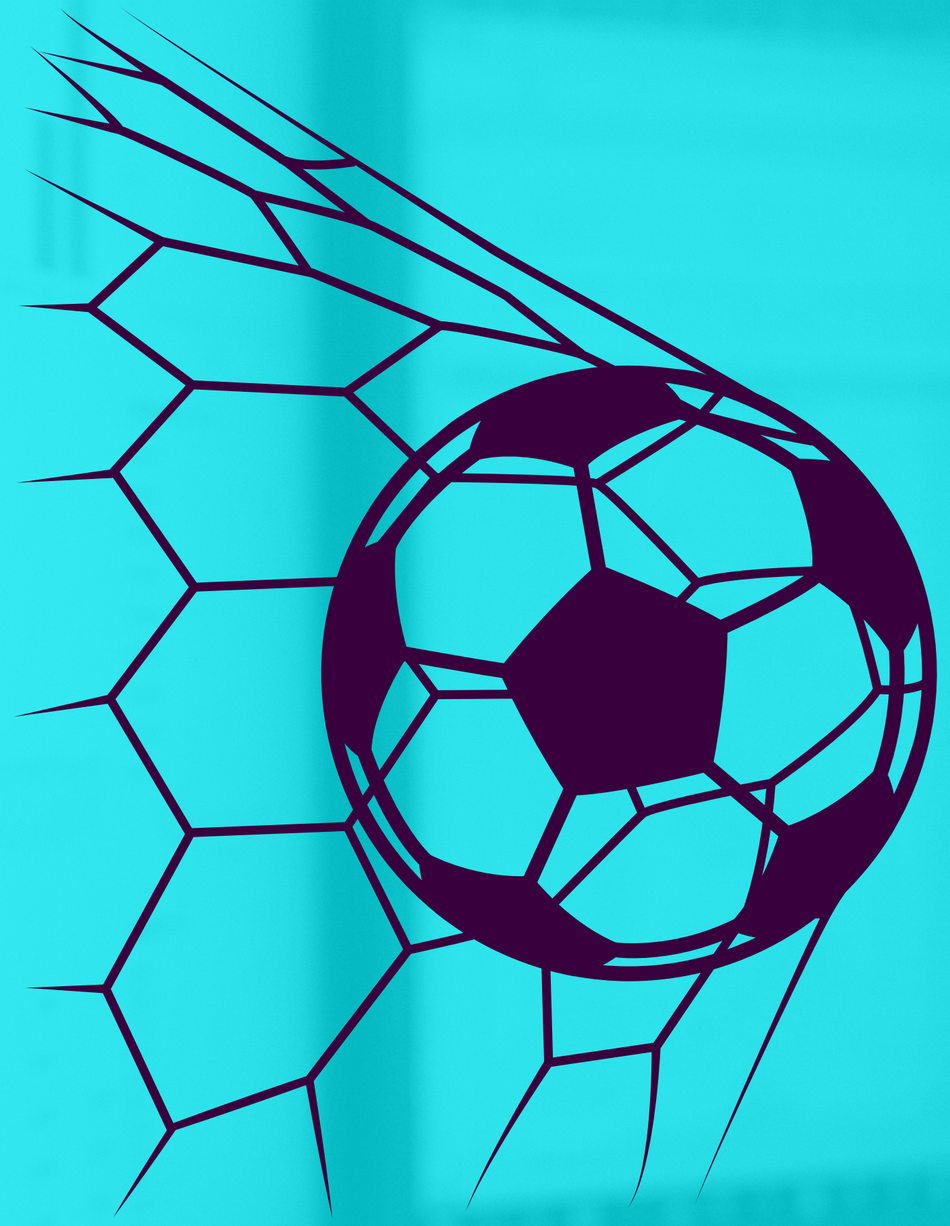
RSQUARED	98.2%
ADJ R_SQUARED	98.2%
MAE ON TRAIN	0.2486
MAE ON TEST	0.2352

TOP FEATURES

- Points_per_minute
- Goals_scored
- Minutes



NEXT STEPS

- Hyperparameter Tuning
- Try other models
- Apply Time Series

THANKS