**Predicting the price value of FIFA players**

**Abstract**

The goal of this project is to analysis and clean list features of more than 18K pro-players that playing in FIFA (International Federation of Association Football), and to set a regression model to predict each player price base on the features that been given in the dataset. Based on the final result, we can estimate the price value of new players. However, only python code is encoded, hopefully next stage I will make a front-end page that can integrate with our python script.

**Design**

This project was design for Tuwaiq-bootcamp, in order to make a machine learning project. The acquiring data is very suitable for our project. From the data we use liner regression model to predict players values.

**DATA**

My dataset contains more than 18K cases with 56 features some of these features are [player\_age, height\_cm, weight\_kg, position, movement\_reactions, shooting, passing, preferred\_foot, power\_strength, movement\_sprint\_speed, value\_eur …]. Most columns are int except some which are objects. Using linear regression as my first algorithm with scaling to set the data patron in to suitable fit for the output prediction. However, the end result I had 85 features all in int form with about 16K rows.

Dataset Ref:( <https://www.kaggle.com/stefanoleone992/fifa-21-complete-player-dataset>)

**Algorithms**

The data cleaning results of more than 13K for rows, and 85 features. The training was 80% and testing was 20%. With linear regression model we had score about 78% of accuracy with Root-mean-square deviation: 2619163. However, I made a second predication using all features to predict players overall performance, the output result was 94% of accuracy with Root-mean-square deviation: 1.6858125893701004

**Tools**

• Data Processing: Pandas, Numpy.

• Modelling: SciKit-Learn, .

• Visualization: Matplotlib, Seaborn, plotly, hvplot

**Communication**

