**Machine Learning Project**

**NBA Players Salary Prediction Using AdaBoost Algorithm  
  
Problem Statement:**The Problem statement is to do some data analysis about basketball nba players’ salary data. Through this analysis, we are predicting using linear regression model. Through this we are predicting the each person’s salary of NBA players using the collected Dataset.  
  
**ML Methodology:**Adaboost iteratively generates K weak classifiers. The final model is the addition of K weak classifiers. Each base classifier is multiplied by its own weight. For the two-class problem (-1,1), each weak classifier will calculate a prediction result of -1 or 1. These prediction results are multiplied by the weighting coefficients of the respective weak classifiers and added. If the final result is greater than 0, it is predicted to be positive (1); otherwise, it is predicted to be negative (-1).

**Data set description:**

The dataset was scraped from Basketball Reference. This dataset is so simple that only includes information about player names, teams and salary amount.   
1. Player: Player name

2. Tm: Team name each player belongs to at the beginning of the 2017- 2018 season

3. seson17\_18: Salary price for 2017-2018 season (USD)

**Pre-Processing:**

Pre-Processing which is to process the data’s before using them in algorithms.  
**Example:** from sklearn.preprocessing import Normalizer  
Most variables are skew, thus we should rescale them.

**Building ML model:**

ML Should be built step by step with each process according to the algorithm. Collect the data set and wrangle the data then find Average Points Per Game Ranking and essential details to visualize the data’s. Now preprocess the data and apply the algorithm to predict the outcome result of the problem statement.

**Training and Evaluation of all ML model:**

For each ML models the data has been trained and tested and it is evaluated before proceeding with the Algorithm models. It is to differentiate the values from original data and the trained data.  
1. Define a function to evaluation regression algorithms, model is fitted algorithms 2. Predict is for if display comparison of prediction and true value of test data

**Details of code repository in GitHub:**

<https://github.com/Clamencia/ML-Project/blob/master/nba-players-salary-prediction%20(1).ipynb>

**Conclusion:**

The results of the AdaBoost algorithm show that the average playing time and efficiency rate have the most important impact on salary, while the average score is only ranked third. This is quite different from the decision tree results.