**PROJECT PROPOSAL**

**Problem Statement/motivation**:

We should be able to successfully predict how a rookie may perform in his rookie season based on physical measurements and previous statistics ( may be college level ). We should also be able to formulate and give the correct order of the current rookies completing their first season based on our work. We will also be using general information about the NBA from online sources such as ESPN and other sports services.. We know the NBA has so many different stars, for example Stephen Curry and Russa Westbrook. We want to find how to evaluate a good NBA star. Most NBA player styles are different in the NBA. We should follow a different situation to find a good player.

**Literature survey：**

1. Previous work, since its NBA has such a long history, we will find so many different NBA stars in NBA history. Some stars did not get a good model pick. They still become a famous NBA star, for example Stephen Curry. Classic statics analysis, models of NBA physical measurement and statics of NBA stars for each season. First resource, they provide us with NBA physical measurements. This will become an important site in the graphy. In addition, second resources show us NBA statics in each season, we could know how to change a new player and gradually change into a super star. Furthermore, these resources talked about how the NBA stars body changed in the graph.

**Proposed Work ：**

1. Propocessing: we will need to convert our data into comptable forms and then merge them together. This will include creating NBA player physical measurement, each season statistics and how about NBA physical measurement changed in each season. We should clean another player in the NBA player stats. We will focus on those data sets getting results in the final.
2. Analysis using the datasets, we will form logistic regression with NBA physical measurements. This will form a baseline of our research, we will need to analyze the effect of each attribute on the low level rate NBA physical measurement star. After analyzing the relationships between all variables, we will create a regression model to generate the probability of NBA body situation and physical measurement .

**Data set:**

NBA physical measurement: https://stats.nba.com/draft/combine-anthro/

NBA physical measurements have 9 parts, we will research which part is so important in the graph.

Every team resource: https://www.basketball-reference.com/ NBA team and player resources should follow this graph. We will know how NBA players change in each season. In addition, NBA teams still have bad picks in the modal draft. For example, 2013 NBA first pick (Anthony Bennote). Why Stephen Curry became a famous star?

Every player Stats : <https://www.kaggle.com/drgilermo/nba-players-stats> This datasets could qualify as important. From this data, we could see how about NBA physical measurement. How about general level in the NBA graphy? We could predict normal NBA physical measurement. When we want to know the NBA super star, we should build a big research group.

**Evaluation Methods:**

We will evaluate our research by testing to see if it accurately predicts whether or not a NBA star percent. We hope to get an accuracy measure of 60%-70%. Because most players didn’t get the first pick, but he still became a NBA super star. We should rethink how to make this error before pickup. Most players change so quickly. When they suffer big hurt and emotion, this will still influence his status. (For example: James Harden(Houston Rockets) has ever suffered these problems.)

**Tools:**

Jupyter Notebooks: From this part, we will build NBA physical measurement, we should build a graph, how about this research in the NBA physical measurement.

Python 3: We will use this download basic datasets graphy, because they couldn’t give us crv results directly. These things help us find data tools, because some data have protection levels.

NumPy: We could get this graph from an NBA resource, this helps us get a good analsize result. Some websites need to use these tools.

Pandas: We will research basic data, building graphy. How about research from this data graphy.

**Milestones:**

* To start with, we will research data, we will find a graph in this data. We will find some data we could not use in the graphy.
* Some of the data will be clean in this first process. In addition, we will start to merge between datasets. We will build a graph in this moment.
* We should show physical measurements in the NBA datasets. Furthermore, we will need to build datasets to make this easily understanding.
* In the end, we should build the next steps. We will build a research paper sturcture.