**Description of Web Application**

The application is will have three focuses:

1) Allows users to select fighters and get back a tables/views of their fighter traits like age, height, weight, etc. and stats on events they’ve fought in and matches they’ve fought.

2) Allows users to view match results by events and by fighters where users choose the event and get back a list of all the match ups and results from the fight.

3) The highlight of application will allow users to select two fighters within the same weight-class and predict who would win a match up as of today with a confidence interval (using some statistical model in the backend). The will be uniquely challenging and complex given predicting fight outcomes tends to involved particularly difficulties unlike other sports. For instance, one wrong move or slip up can lead to a knockout or injury and a lost match where as in other sports multiple players and longer games make it difficult for single mishap to directly result in an outcome of a match. Furthermore, more power statistical techniques such as nueral nets require large datasets and feature sets in order to train predictive models. Some research has been able to get to 68% accuracy on hundreds of thousands of data points. Despite this it will be interesting to see if we can formulate a juvenile predictive model from our database.

**Overview of Entities/Relationships**

We will have 6-8 entities in the application for example: players, fights, teams, weight classes, injury status, etc. The relationships will also be similar in number. For instance, “players to teams” are many-to-one, while players “belongs\_to” weigh-class is one-to-many.

**Data Sourcing**

We will use public available data from ufc.com and sherdog.com – two primary sources for mixed martial arts data. There are also available XML datasets online by weight class. If data is missing or incomplete, ufc.com has data that we can scrap directly from.

**Expanded-Design Option**

The expanded-design option is essentially part 3 of our original idea. We would expanded the number of fighters and resulting data set by ~10 fighters to say 2000 and try to increase the data sets to 50,000+ points which will help increase the accuracy of our predictive model. Given how complex the task is of developing an greater than random accuracy statistical predictive model, this design option is very complex.