Predicting The Method of Victory in UFC Fights

Navin Gopaul, navingopaul2018@u.northwestern.edu
EECS 349: Machine Learning, Northwestern University
[insert repo link]

Introduction

The Ultimate Fighting Championship (UFC) is a promotion company that hosts contests in mixed-martial arts (MMA). These contests occur in an arena called "the octagon" with 5 rounds at 5 minutes each for championship fights and 3 rounds at 5 minutes each for non-title fights.

My task is to determine the method of victory in a UFC fight (by knockout, submission, technical knockout, decision, or draw) based on attributes of each bout including age difference between the two fighters, weight difference, height difference, fighting class, the referee, and when the round ended. Being able to predict the outcome of a fight based on the attributes of each fighter could shift the odds against or in favor of different fighters prior to a fight and shift fan votes and expert predictions. This could have implications on the cost of betting for or against certain fighters or cost of betting on how a fight will end. Additionally, knowing what factors affect the outcome could influence how a fighter trains. For example, if a certain weight class tends to end by submission, competitors in that class would know that they have to focus on their ground game.

Dataset

I began with two sets of data in .xlsx format from www.sherdog.com, a website that does data analysis on UFC fights. The first dataset contained the data from every UFC fight from November of 1993 to February 2016 for a total of 3,569 fights. This attributes of each fight include the names of the two fighters, the event date, the event name, the event organizer, who won the fight, the referee's name, the method of victory (KO, TKO, submission, decision), and specifics about that method (ie Rear-Naked Choke, Armlock, etc.), and the round/time the fight ended. The second dataset contains data from every UFC fighter who has ever competed for a total of 1,561 fighters. This dataset includes attributes such as the fighter's name, nickname, birth date, height, weight, class, locality (where they're from), and the fighter's ID.

I decided to refine this data to one single .csv that I could analyze using Weka. I used a Python library for the purpose of .xlsx editing called openpyxl to write a script to create this refined dataset. I traversed the dataset containing all of the fight data and matched the two fighter's IDs to their IDs in the fighters data to create a new fights dataset with more interesting/meaningful attributes. I computed the fighters' age difference, weight difference, and height difference by subtracting the smaller of the two from the greater (ie older – younger, heavier – lighter, taller – shorter). I also copied over the other attributes from the fight dataset including the referee, the round the fight ended, the weight class, and method of victory (the output I'm examining).

The dataset was split into training and test using Weka's built in RemovePercentage filter. After removing percentage I saved that file as the training set (1784 instances) and used invertSelection to get the portion not removed and used that as the validation set (1785 instances).

Analysis and Implementation

Each model I used was trained on all UFC fights dating back to 1993 in the refined dataset that I computed, and then validated. I used 10-fold cross-validation for the data. The following table lists the

validation accuracy for each method.

Method	Accuracy
ZeroR	35.5182%
J48	62.6891%
BayesNet	60.014%
IBk	49.2152%

I began by using ZeroR to determine baseline performance as a benchmark for later classification methods. ZeroR ended up returning an accuracy of 35.5182%. After running J48, I received an accuracy of 62.6891%. BayesNet yielded an accuracy of 60.014%, and IBk had an accuracy of 49.2152%.

Classified as	Submission	TKO	KO	Draw	Decision
Submission	49	129	55	0	63
TKO	89	147	63	0	75
KO	60	102	52	0	19
Draw	0	0	0	7	0
Decision	0	0	0	0	836

Above is the confusion matrix yielded from running the J48 algorithm on our data. The model correctly predicts fights that end in a draw or end up going to a decision. The model classifies fights that end in Submission/TKO/KO with moderate inaccuracy. This may be because the number of rounds that a fight goes to gives us a clear distinction between these two types of outcomes – for example, fights that go one for longer rounds overwhelmingly end in a either a draw or a decision while shorter rounds end in TKO/KO/Submission so it is harder to differentiate between these three outcomes.

Based on the visual representation of the decision tree, we can see that the most important attribute in determining the outcome of a fight is the number of rounds that the fight goes on for. This intuitively makes sense since fights are either up to 3 or 5 rounds long, so in the tree we can see that if the round lasts longer than 2 rounds, the fight is very likely to end up going to the judge's decision. The second most important feature seems to be the referee – but this is likely misleading since different referees worked during different eras of UFC. For example, earlier UFC fights in the 90s were a lot less regulated so the referees during that time period had a tendency to ref fights that ended in KOs, TKOs, and submissions, whereas more fights today tend to end up going to a decision. The next most important feature seems to be weight class and weight difference between the two fighters. Heavyweight fights seem to have a tendency to end in KO or TKO in the earlier rounds whereas lighter weight fights (bantamweight, lightweight, flyweight, etc.) seem to end in submission more than heavier fights. Lastly, age difference and height difference don't seem to have a meaningful impact on the outcome of these bouts.

For future experiments related to the outcome of UFC fights, it would be interesting to include some other attributes to each fight, such as the different in reach of each fighter (arm span, which dictates how far their punches reach), the punching power of each fighter, and their win/loss record at the time

of that fight. Additionally, this experiment focused on the outcome of each fight rather than attempting to predict who will win – which is another area ripe for exploration.