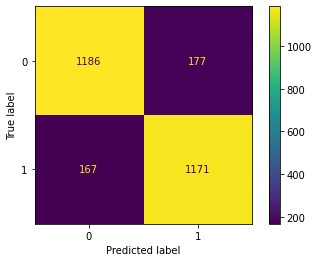
**Introduction**

In this project, I created a logistic regression model using UFC fight stats from every fight. The purpose of this model is to determine whether statistics can accurately predict a fight winner, in addition to determining which are the most important stats to look to when determining a winner. Much of this project was finding a good data source and creating a dataframe that could be used for this model. The data was originally arranged by fight round, but I simply summed the stat totals of each fight to make sure all fights have equal representation, regardless of rounds fought. After this, I added columns for opponent stats, allowing the model to factor in both fighters’ performance. In addition, I reduced the dataset to fights that ended with a winner, to allow myself to run this binary classification model.

To measure the accuracy of my model, I used a confusion matrix and accuracy score. While the accuracy score gives a good idea of the model’s strength, the confusion matrix will tell us a lot of what needs to be determined about the model. For instance, the model must predict a similar amount of winners and losers, as the real total amounts are equal.

**Discussion**

The model was able to perform quite accurately. It was able to correctly predict roughly 87% of the test data fights currently, a number which held consistent for the training data. From the confusion matrix I was able to find that the model does predict winners and losers near evenly, with 1348 predicted winners and 1353 predicted losers. In addition, the 177 false positives is very similar to the 167 false negatives.



After determining that the model is sufficiently accurate and predicts both sides evenly, I shifted gears to determining the strongest predictors. The strongest positive predictors turned out to be knock downs, strikes landed to the head, and strikes landed on the ground. An increase in either by one standard deviation in any of the three goes hand in hand with multiplying one’s odds of winning by over 2.25. Naturally the strongest predictors to losing a fight are one’s opponent’s knockdowns, strikes landed to the head, and strikes landed on the ground.

I predicted significant strikes landed to be the next strongest predictor after knockdowns. I was not correct, but not far off, as a standard deviation increase in significant strikes landed relates to odds being multiplied by 1.95. This does make sense, as strikes to the head tend to be more significant than those to the body or legs, that head strikes would have a higher odds coefficient than the whole of significant strikes. In addition, ground strikes having such a high odds coefficient also makes much sense. Ground strikes are usually the results of a takedown, knockdown, submission attempt, or reversal. All of these are strong predictors in favor of a fighter’s chances of winning.

**Conclusion**

In this project, I was able to successfully build a model which can use fight statistics to accurately determine a winner. In addition, I learned which statistics have the largest impact on whether or not someone will win a fight. In the future, I plan to continue to use this dataset to answer more of my questions about UFC fights and statistics.