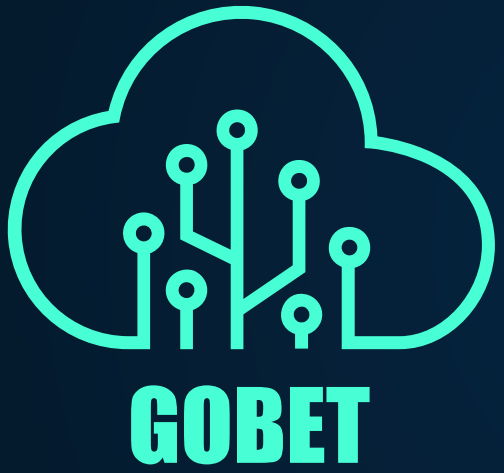
UFC CLASSIFICATION PROJECT

by Luigi Fiori



OUR COMPANY

GOBET is a bet site company mainly focused on football.

The Management is looking for opportunities to expand their Market.

Lately Combat Sports are raising in terms of popularity mainly because of the UFC boom.

As consultants, we have been asked to create a model capable of predicting wins and losses for UFC fights.

OUR GOALS

USABILITY

We want to build a Model that can be used on new random fights

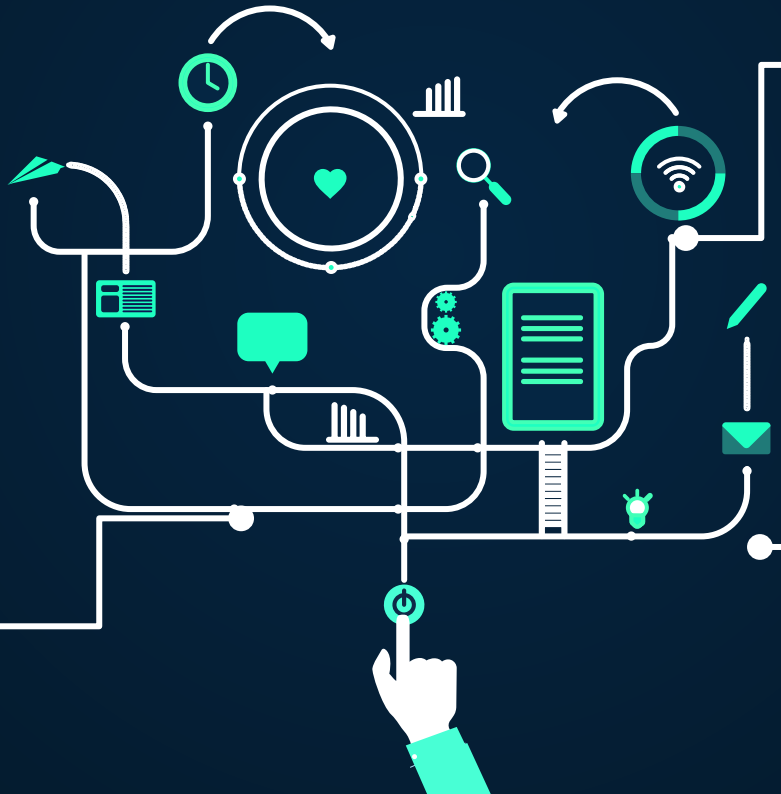
INTERPRETABILITY

We are looking to create a Model easy to interpret and understand

EXPANSION

We want to extend the bet options available in our website in order to cover an higher slice of Market

PROJECT STAGES



STEP 1

Clean Data

STEP 2

Check most important features for predicting a fight

STEP 3

Build a Machine Learning Model and Evaluate it

DATA PROCESSING

OBTAINING THE DATA	Dataset found on Kaggle.com. 5144 fights with 145 columns
DEAL WITH NAN VALUES	We dropped 20% of the rows because of missing values
FIXING MISTAKES	We dealt with bad data such wrong weight difference etc.
REDUCING COLUMNS	We reduced the columns to get a more manageable Dataset
CREATING NEW FEATURES	We created new features to fit our model

TOP 6 FEATURES

Takedowns
Result in taking the opponent on the floor

01

Ground Strikes
Strikes landed on the ground

02

Total Losses
Number of losses for the fighter

03

04

Head Strikes
Total number of strikes landed to the head

05

Total Seconds Fought
Total number of seconds spent by the fighter in official fights

06

Age
Fighter's age

TOP MODEL SCORES

LOGISTIC REGRESSION

62%

DECISION TREES

62%

RANDOM FOREST

62%

We tried different Machine Learning Models and as we can see, all the Models performed 12% better that random guessing, an overall decent result.

It’s interesting to notice that there is no actual difference between them, in fact the score is always 62%

RECOMMENDATIONS

We recommend to consider these features when deciding the fight's odds:

1. Total Number of Losses of the Fighter

2. Age of the Fighter

3. Takedown Capabilities(Related Fighter Background Techniques)

4. Total Number of Rounds Fought by the Fighter

FUTURE WORK

Having more time to spend on the project it would have been interesting to analyse more in depth the features in the dataset.

In particular spending some time on creating new features based on the one we have could lead to some interesting results.

Finally to have a more accurate prediction we could use a more resource intensive models such as Support Vector Machines or XBoost.

Thank For Watching!

Does anyone have any question?