In Summary, the steps that were taken to create this predictor were:

* Taking raw MMA data in the form of CSV records fighter information from 1994 – 2021 (source: Kaggle)
* Putting it as a data frame on Jupyter notebook
* Adding a ‘target’ column which brings the results of the fighter’s next fight back by 1 period to the one we are trying to predict for.
* Cleaning the data frame by importing the panda module (pd) and utilising it to filter/remove all data columns with null values
* Data is then scaled from 0 to 1 then ran through a sequential feature selector in which the machine chooses 10 factors/columns of data it deems the most optimal for prediction
* Then create a function ‘Backtest’ that splits data up into half years (typical active period of MMA fighters) and uses the first 2 6 months periods to predict the next 6 months. However, for it to work, we need 2 semi-annuals worth of data for it to work
  + Function takes in arguments:
    - Data: data frame being used
    - Model: model used for prediction: in our case will be ridge regression module we imported as RidgeClassifier
    - Start: Is the number of periods/ date we want to test the data for
    - Steps: How many number of periods we want to test for at a time
  + Categorises data into the training set and test set of data. The test set of data is specified by number (starting period) stated in ‘start’ argument this set is the value that is predicted, the testing set are all existing data previous to the starting period.
  + Results were then presented as 2 columns: actual and prediction.
    - Actual being the columns that is genuinely the future results of the specific periods (determined by ‘target’)
    - Predictions being the predictions made in previous function
* The accuracy score is then measured to test the validity of the predictor test
* Following that, we go back to the data frame:
  + Sort data into grouped by fighter names
  + Add information about next opponent and date of next fights
  + Following that, we merge the information of the fighter we are trying to predict for, with the information of the fighter’s next opponent utilising the similarity of date of their next fights.
    - Creating a ‘full’ data frame that merges all the info of fighter from both corners into a single row e.g, data such as losses, significant strikes.
* Finally, we take this next ‘full’ data frame and place it into the sequential feature selector and put it through the ‘backtest’ function again, the output of the backtest this time should be significantly improved.
* Once tested to an optimal accuracy, this model can be used to predict future MMA matches by taking the most recently updated UFC fight data and putting this data as the ‘testing’ data in the ‘backtest’ predictor function, thus, allowing it to test the results of future matches.