

PUBG Player Placement Prediction

Group No - P09

Rachit Shah (rshah25)

Sourabh Sandanshi (ssandan)

Udit Misra (umisra)



What is PUBG?



- Online Multiplayer Battle Royale game
- Max 100 players fight in either solo, duo or squads
- However, there are events and custom games which allow team sizes more than 4.
- The last person or the last team alive wins the match.
- During the game, players search buildings, ghost towns to find weapons, vehicles, armor, and other equipment.



Dataset

<https://www.kaggle.com/c/pubg-finish-placement-prediction/data>



- Training Set: 4.45 million
- Test Set: 1.93 million
- Features: 28



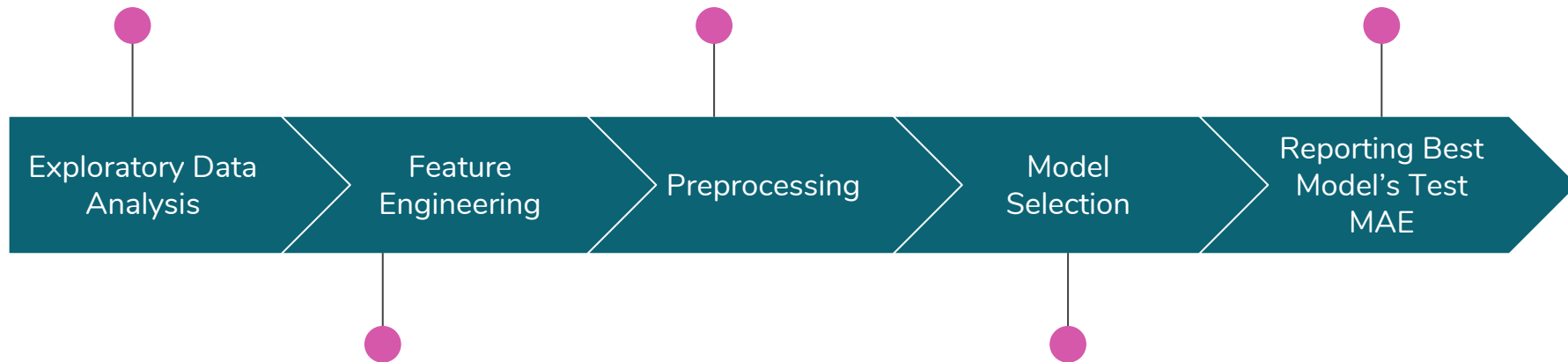
Procedure



Plotting different plots
and gaining insights
from the data

Normalising, removing
NA and Infinity values,
and handling
categorical data

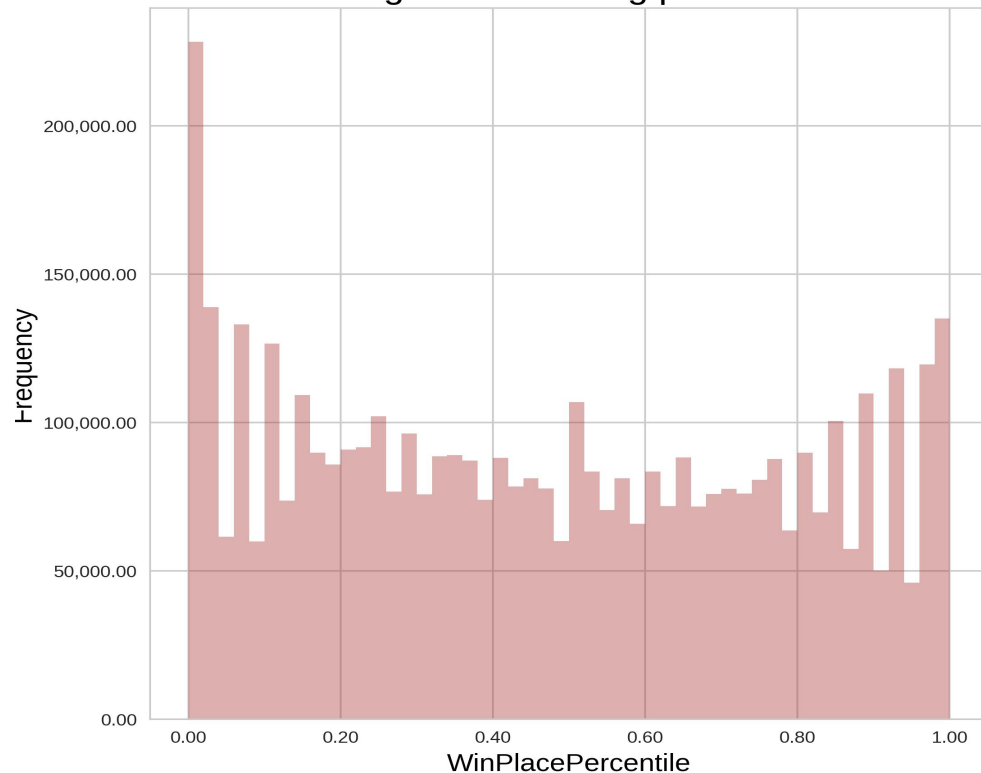
Fitting the test dataset
on the best model and
reporting Final MAE



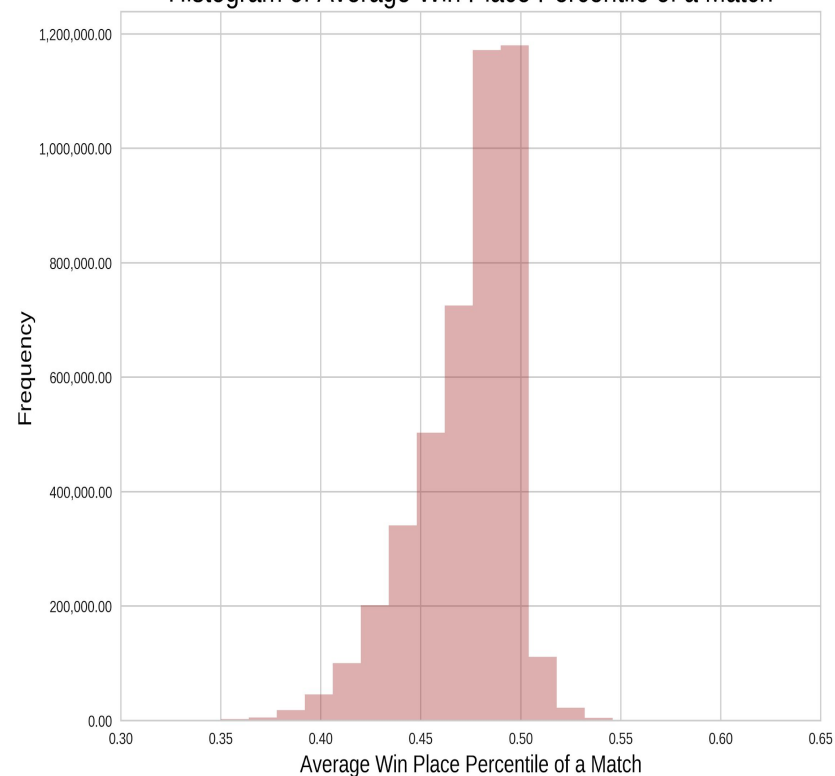
Adding relevant new
features to boost
model accuracy

Comparing different
Regression models
and calculating their
MAE

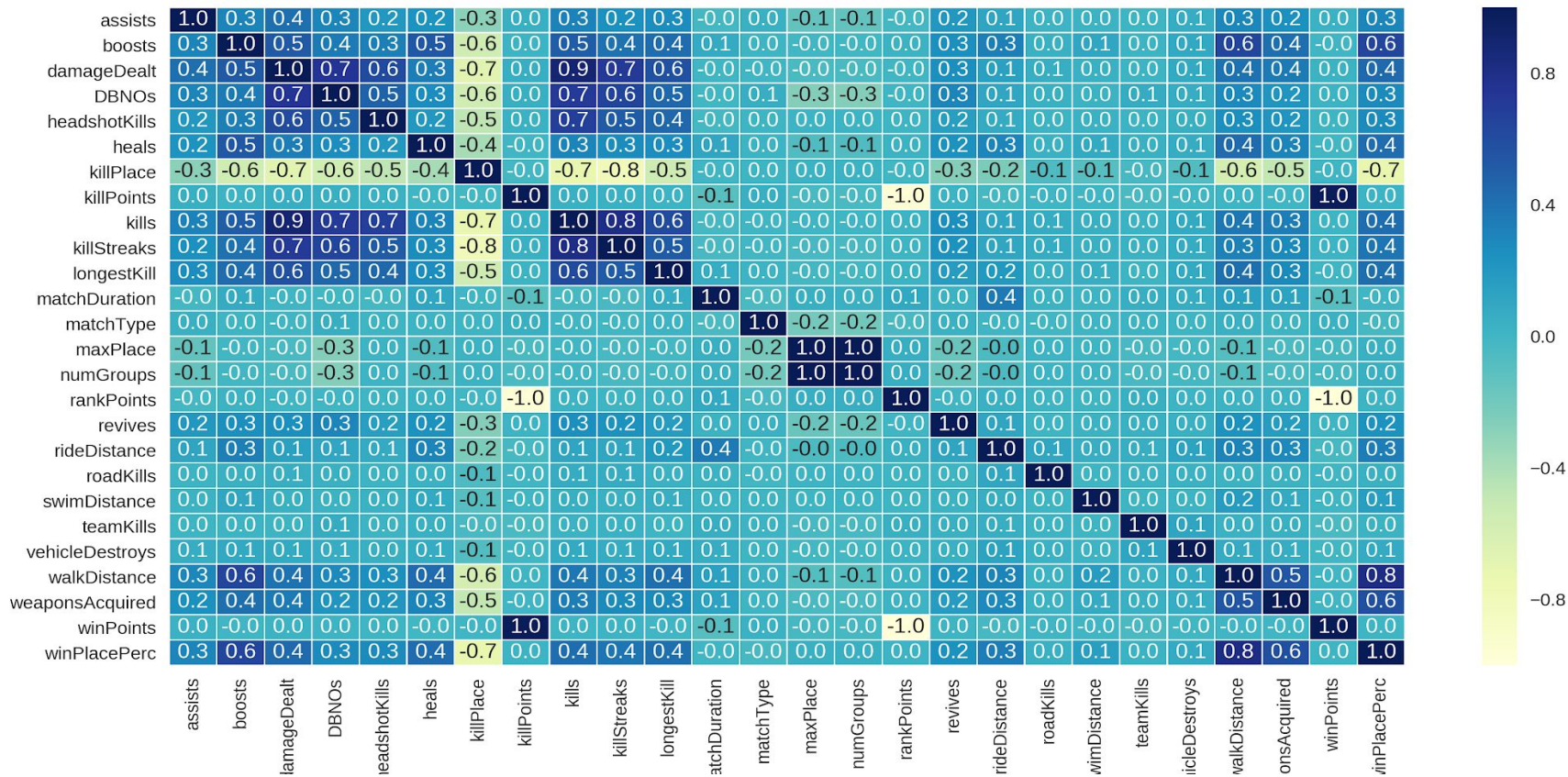
Histogram of winning percentiles



Histogram of Average Win Place Percentile of a Match

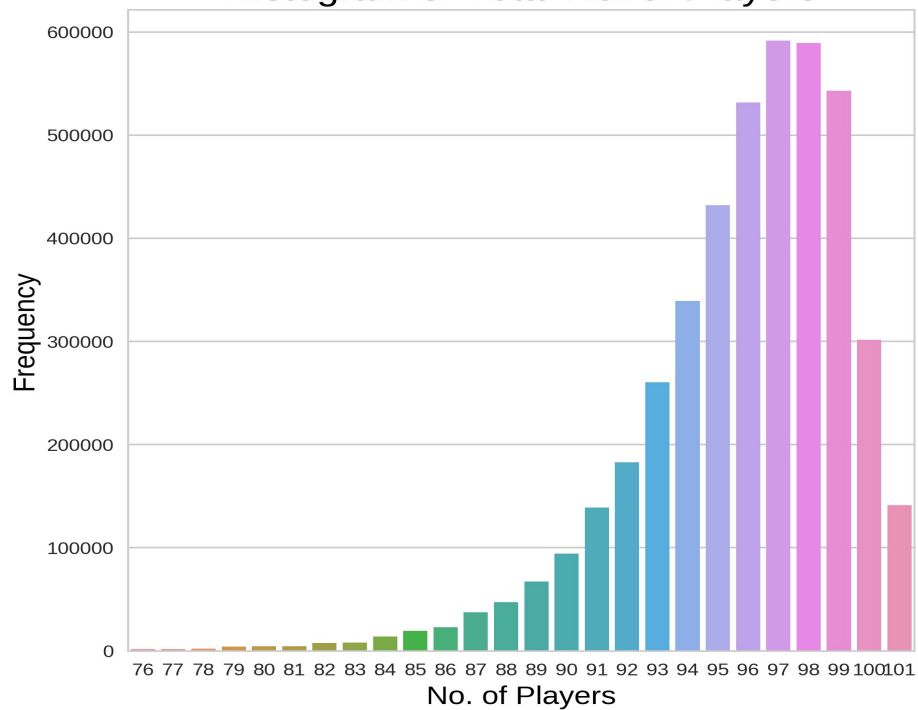


Distribution of Target Variable

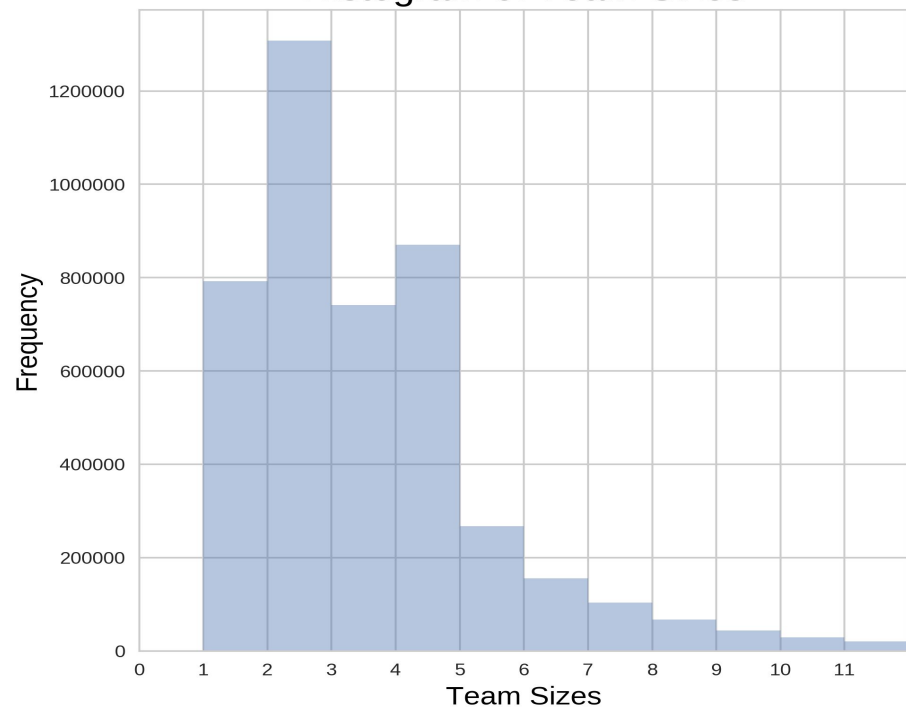


Correlation

Histogram of Total No. of Players

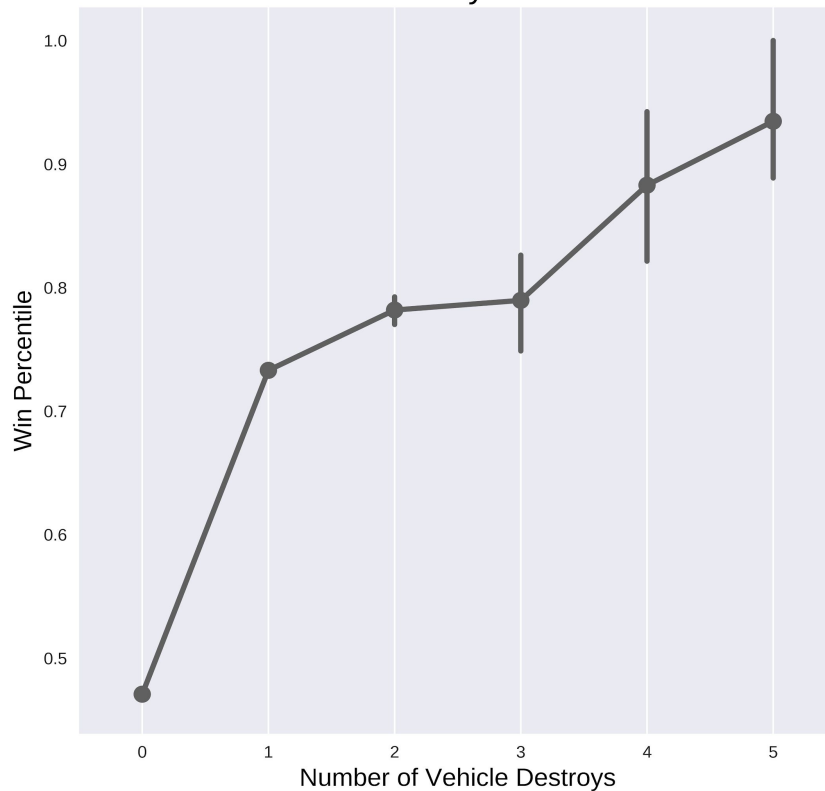


Histogram of Team Sizes

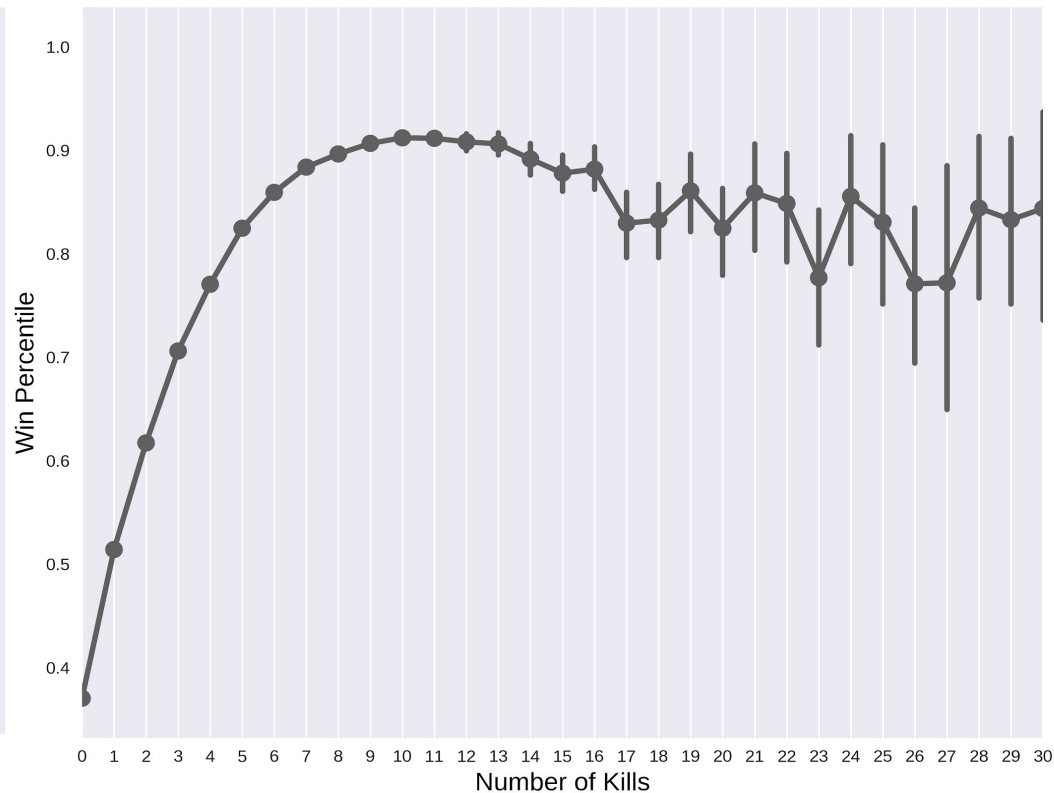


Total Players in the game

Vehicle Destroys/ Win Ratio



Kills/ Win Ratio



Skill Analysis



Feature Engineering



Features	Explanation
totalPlayers	Total Players in the match
teamSize	Total team member in a team
normMatchType	Match type like solo, duo,quad etc.
totalDistance	Swim+Ride+Walk Distance
maxPossibleKills	Total kills by team
itemsUsed	boost+heals+weapon



Model Selection

Model → Scores ↓	Linear	Ridge	LASSO	Elastic Net	AdaBoost (lr=0.8)	Gradient Boosting (lr=0.8)
MAE	0.08755	0.08755	0.12084	0.113	0.0974	0.05861
R2	84.83%	84.83%	74.42%	77.06%	82.79%	93.01%

Model → Scores ↓	Random Forest (n_est=10)	Decision Tree	LGBM	Simple MLP (adam) (lr=adaptive, e=23)	Deep Learning
MAE	0.05751	0.07711	0.05392	0.08630	0.06478
R2	93.00%	86.92%	93.70%	85.42%	90.71



Deep Learning

MAE=0.0623

Parameters:

Epochs: 20

Input: (3112875,43) , Output: (3112875,1)

No. of hidden layers: 4 with batch normalization and dropout (0.1)

Optimizer: Adam with learning rate=0.01, epsilon=1e-8, decay=1e-4,

Activation: ReLu for hidden and sigmoid for output

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 512)	22528
batch_normalization_1 (Batch Normalization)	(None, 512)	2048
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 256)	131328
batch_normalization_2 (Batch Normalization)	(None, 256)	1024
dropout_2 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 128)	32896
batch_normalization_3 (Batch Normalization)	(None, 128)	512
dropout_3 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 1)	129

Total params: 190,465

Trainable params: 188,673

Non-trainable params: 1,792



Grid Search - LightGBM

Grid Parameters:

Learning rate: [0.05, 0.1, 0.3, 0.002]

N_estimators : [100, 250, 10, 50]

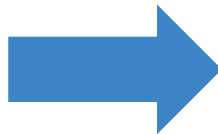
Num_leaves : [60, 200, 25, 125]

Boosting_type : [gdbt, dart, goss]

Objective : mae

Early Stop Rounds : [10, 50, 200]

Random State : 501



Best Parameters:

Learning rate: 0.3

N_estimators : 250

Num_leaves : 200

Boosting_type : gdbt

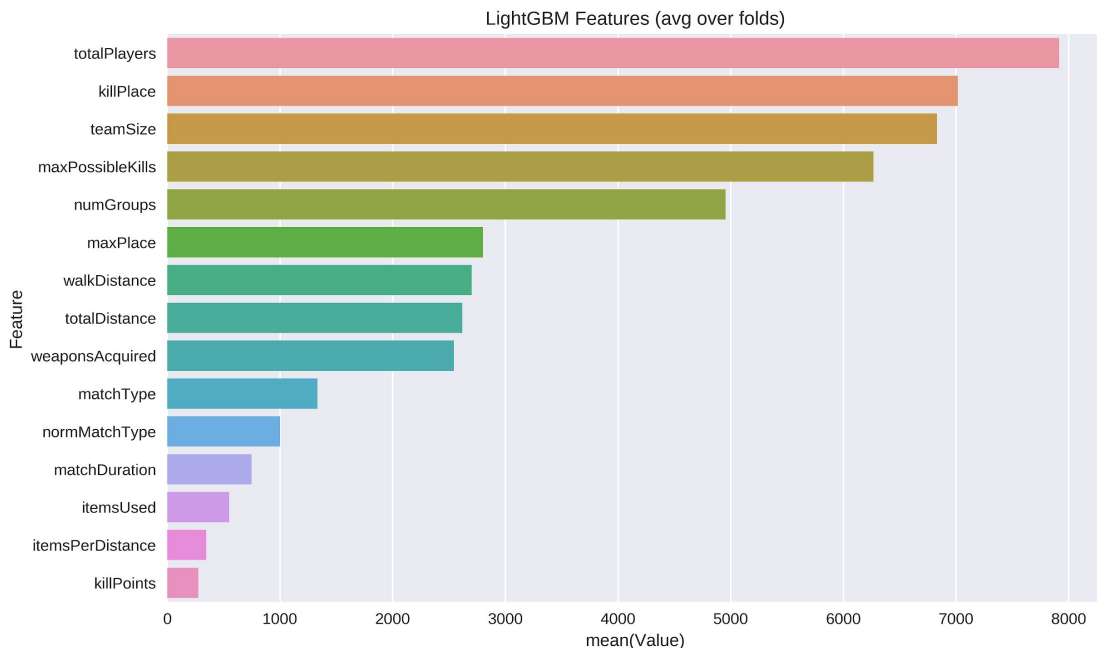
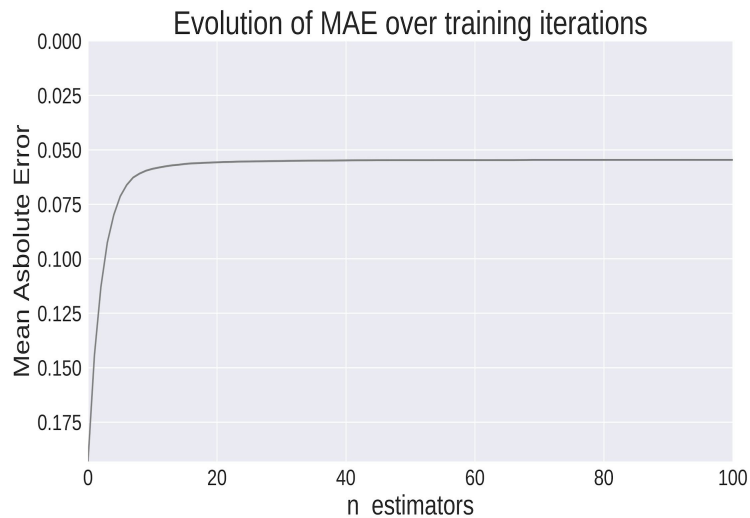
Objective : mae

Early Stop Rounds : 10

Random State : 501



LGBM (Light Gradient Boosting Machine)



Parameters: $n_est=250$, $lr=0.3$, $n_leaf=200$,
 $early_stop_rounds=10$

MAE: 0.607 → (After feature engineering) → 0.5392



Conclusion

- Applied Validation Dataset to 11 Models
- LGBM Gave highest MAE
- Got following result on Test Data by LGBM :

Name	Submitted	Wait time	Execution time	Score
sample_submission.csv	a day ago	2 seconds	14 seconds	0.0539

Complete