PROCESSES INVOLVED IN THE ZS CHALLENGE

DATASET IS LOADED

DATA PREPROCESSING

DELETING UNRELEVANT COLUMNS LIKE Match id etc

MERGING COLUMNS LIKE Type Of shots

DEALING WITH NA’S WHICH ARE FILLED BY BACK Pass and forward pass

DIVIDING THE DATA INTO TRAIN/TEST

Then Selecting important features for the model

Normalisation

Model generated by algorithms like

KNN

SVM

XGBOOST

DECISION TREE

ADABOOST

Then model is used to calculate the probability of goal prediction

PROPER ANALYSIS CAN BE FOUND IN THE PYHTON CODE