Machine Learning 2022

# Milestone 2. Team: CS\_37

classification accuracy:0.82

total train time :13.8

total test time:0.06

**Hyperparameters**

Hyperparameter for SVC model is C with values (.1,0.01,0.001) and kernel('linear', 'rbf')

And it gives the best score with C': 0.01, 'kernel': 'linear'

Hyperparameter for LogisticRegression model is C with values (1,5,10,15)

And it gives the best score with C 10

Hyperparameter for MultinomialNB model is alpha with values (1,2,3,4) and fit\_prior with values (true ,false)

And it gives the best score with alpha 1 and fit\_prior true

Hyperparameter for DecisionTreeClassifier model is criterion : ("gini", "entropy"), splitter ("best", "random"),max\_depth(5,10,15,20) and max\_features:( "auto", "sqrt", "log2")

And it gives the best score with criterion:’gini’ ,max\_depth:15 , max\_features:”sqrt”,splitter:”best”

**Top features which have a correlation with the ‘PlayerLevel’ more than 0.0**

**our features are:**

['age','height\_cm','weight\_kgs','nationality','overall\_rating',

'potential', 'wage', 'preferred\_foot', 'international\_reputation(1-5)',

'weak\_foot(1-5)', 'skill\_moves(1-5)', 'body\_type','release\_clause\_euro', 'club\_team', 'club\_rating', 'club\_position','club\_jersey\_number', 'crosg', 'finishing', 'heading\_accuracy','short\_passing', 'volleys', 'dribbling', 'curve', 'freekick\_accuracy','long\_passing', 'ball\_control', 'acceleration,'sprint\_speed','agility', 'reactions', 'balance', 'shot\_power', 'jumping', 'stamina','strength', 'long\_shots', 'aggression', 'interceptions', 'positioning', 'vision', 'penalties', 'composure', 'marking', 'standing\_tackle',

'sliding\_tackle', 'GK\_diving', 'GK\_handling', 'GK\_kicking','GK\_positioni', 'GK\_reflexes', 'LS', 'ST', 'RS', 'LW', 'LF', 'CF','RF', 'RW', 'LAM', 'CAM', 'RAM', 'LM', 'LCM', 'CM', 'RCM', 'RM', 'LWB','LDM', 'CDM', 'RDM', 'RWB', 'LB', 'LCB', 'CB', 'RCB', 'RB','PlayerLevel', 'position\_first', 'position\_second', 'position\_third','upper\_work\_rate', 'down\_work\_rate', 'num\_of\_work\_rate']

**model\_params: list contains the model we used**

**(**LogisticRegression(),GaussianNB(),MultinomialNB(),

DecisionTreeClassifier()**)**

We used **GridSearchCV** to determine the best score and parameters for each

Model in model\_params list and the result

Graphical user interface, text, website

Description automatically generated

So **SVC** is the best model with **score 0.89**