Last login: Tue Dec 1 11:52:06 on ttys000

valindavanam@Valindas-Air ~ % cd /Users/valindavanam/Documents/ML/Final

valindavanam@Valindas-Air Final % ls

Datasets model.py

Final Project Report.docx newDataset.xlsx

\_\_pycache\_\_ requirements.txt

featureDataset.xlsx teamData.xlsx

main.py ~$nal Project Report.docx

valindavanam@Valindas-Air Final % python3 main.py

Displaying top-5 rows of Countries data

id name

0 1 Belgium

1 1729 England

2 4769 France

3 7809 Germany

4 10257 Italy

Displaying top-5 rows of Match data

id country\_id league\_id season stage ... GBD GBA BSH BSD BSA

0 1 1 1 2008/2009 1 ... 3.25 4.00 1.73 3.40 4.20

1 2 1 1 2008/2009 1 ... 3.25 3.75 1.91 3.25 3.60

2 3 1 1 2008/2009 1 ... 3.20 2.50 2.30 3.20 2.75

3 4 1 1 2008/2009 1 ... 3.75 5.50 1.44 3.75 6.50

4 5 1 1 2008/2009 1 ... 3.50 1.65 4.75 3.30 1.67

[5 rows x 115 columns]

Displaying top-5 rows of League data

id country\_id name

0 1 1 Belgium Jupiler League

1 1729 1729 England Premier League

2 4769 4769 France Ligue 1

3 7809 7809 Germany 1. Bundesliga

4 10257 10257 Italy Serie A

Displaying top-5 rows of Team data

id team\_api\_id team\_fifa\_api\_id team\_long\_name team\_short\_name

0 1 9987 673.0 KRC Genk GEN

1 2 9993 675.0 Beerschot AC BAC

2 3 10000 15005.0 SV Zulte-Waregem ZUL

3 4 9994 2007.0 Sporting Lokeren LOK

4 5 9984 1750.0 KSV Cercle Brugge CEB

Displaying descriptive analysis of Match data

id country\_id ... BSD BSA

count 25979.000000 25979.000000 ... 14161.000000 14161.000000

mean 12990.000000 11738.630317 ... 3.660742 4.405663

std 7499.635658 7553.936759 ... 0.868272 3.189814

min 1.000000 1.000000 ... 1.330000 1.120000

25% 6495.500000 4769.000000 ... 3.250000 2.500000

50% 12990.000000 10257.000000 ... 3.400000 3.400000

75% 19484.500000 17642.000000 ... 3.750000 5.000000

max 25979.000000 24558.000000 ... 13.000000 34.000000

[8 rows x 105 columns]

Displaying descriptive analysis of League data

id country\_id

count 11.000000 11.000000

mean 12452.090909 12452.090909

std 8215.308472 8215.308472

min 1.000000 1.000000

25% 6289.000000 6289.000000

50% 13274.000000 13274.000000

75% 18668.000000 18668.000000

max 24558.000000 24558.000000

Displaying the rows and columns

Countries data: (11, 2)

League data: (11, 3)

Match data: (25979, 115)

Team data: (299, 5)

Displaying the final dataset

id league\_id home\_team\_api\_id ... season total\_goals result

0 1 1 9987 ... 2008/2009 2 0

1 2 1 10000 ... 2008/2009 0 0

2 3 1 9984 ... 2008/2009 3 -1

3 4 1 9991 ... 2008/2009 5 1

4 5 1 7947 ... 2008/2009 4 -1

[5 rows x 9 columns]

Displaying the rows and columns count of the final dataset : (25979, 9)

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

Displaying the rows and columns count of the feature dataset dataset :(4064, 12)

The shape of the train and test datasets are (3234, 10) and (830, 10)

Training initiated

Logistic Regression model is getting trained

SVM model is getting trained

Random Forest model is getting trained

K Neighbors Nneighbors model is getting trained

Multi Layer Perceptron model is getting trained

Gradient Boosting model is getting trained

Decision Tree model is getting trained

prediction initiated

Logistic Regression model is getting analyzed

SVM model is getting analyzed

Random Forest model is getting analyzed

K Neighbors Nneighbors model is getting analyzed

Multi Layer Perceptron model is getting analyzed

Gradient Boosting model is getting analyzed

Decision Tree model is getting analyzed

classification performance report of Logistic Regression

precision recall f1-score support

-1 0.471545 0.522523 0.495726 222

0 0.384615 0.080645 0.133333 248

1 0.541353 0.800000 0.645740 360

accuracy 0.510843 830

macro avg 0.465838 0.467723 0.424933 830

weighted avg 0.475849 0.510843 0.452511 830

classification performance report of SVM

precision recall f1-score support

-1 0.000000 0.000000 0.000000 222

0 0.000000 0.000000 0.000000 248

1 0.433735 1.000000 0.605042 360

accuracy 0.433735 830

macro avg 0.144578 0.333333 0.201681 830

weighted avg 0.188126 0.433735 0.262428 830

classification performance report of Random Forest

precision recall f1-score support

-1 0.504386 0.518018 0.511111 222

0 0.316901 0.181452 0.230769 248

1 0.550000 0.702778 0.617073 360

accuracy 0.497590 830

macro avg 0.457096 0.467416 0.452985 830

weighted avg 0.468151 0.497590 0.473306 830

classification performance report of K Neighbors Nneighbors

precision recall f1-score support

-1 0.335616 0.441441 0.381323 222

0 0.287081 0.241935 0.262582 248

1 0.513678 0.469444 0.490566 360

accuracy 0.393976 830

macro avg 0.378792 0.384274 0.378157 830

weighted avg 0.398346 0.393976 0.393226 830

classification performance report of Multi Layer Perceptron

precision recall f1-score support

-1 0.000000 0.000000 0.000000 222

0 0.000000 0.000000 0.000000 248

1 0.433735 1.000000 0.605042 360

accuracy 0.433735 830

macro avg 0.144578 0.333333 0.201681 830

weighted avg 0.188126 0.433735 0.262428 830

classification performance report of Gradient Boosting

precision recall f1-score support

-1 0.523605 0.549550 0.536264 222

0 0.395349 0.137097 0.203593 248

1 0.571429 0.811111 0.670494 360

accuracy 0.539759 830

macro avg 0.496794 0.499252 0.470117 830

weighted avg 0.506025 0.539759 0.495083 830

classification performance report of Decision Tree

precision recall f1-score support

-1 0.336100 0.364865 0.349892 222

0 0.312253 0.318548 0.315369 248

1 0.464286 0.433333 0.448276 360

accuracy 0.380723 830

macro avg 0.370879 0.372249 0.371179 830

weighted avg 0.384573 0.380723 0.382249 830

Accuracy score of the models

Accuracy Score of Logistic Regression is 0.5108433734939759

Accuracy Score of SVM is 0.43373493975903615

Accuracy Score of Random Forest is 0.4975903614457831

Accuracy Score of K Neighbors Nneighbors is 0.39397590361445783

Accuracy Score of Multi Layer Perceptron is 0.43373493975903615

Accuracy Score of Gradient Boosting is 0.5397590361445783

Accuracy Score of Decision Tree is 0.38072289156626504

Chart, bar chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated A picture containing chart

Description automatically generated

Thank you !!