



















**Explanation:**

In the first step we are getting the data set and visualizing it to our needs. We are getting it types and information and printing it. In the next step the dataset is trained with 75% and tested for the rest at random and the table for that is displayed with the addition of the “isTrain” Boolean column. The next step has us import the DecisionTreeClassifier method and create a decision tree that uses entropy for criterion attribute and only splits at 20 matches. Then a confusion matrix is created for the testing dataset. After this we create a dotfile for the visualization of the Decision tree and save it else separately at the defined location. Now we start again but this time we train the model with 80% of the values. We now run the model 10 times and get the mean score which turned out to be 0.9333333333333333. This is followed by testing the model with remaining 20% data and the accuracy and confusion matrix is printed out. Accuracy being 0.9666666666666667. We have to be careful not to add the labels to the command since we are supposed to print it out without labels. Now we can visualize the confusion matrix with the seaborn library.