Lab4 Report:

Task: Create your own dataset for Image Classification Problem. Use the workflow as discussed in the Tutorial 4 Session using Decision Tree Algorithm. Report the accuracy and confusion matrix obtained.

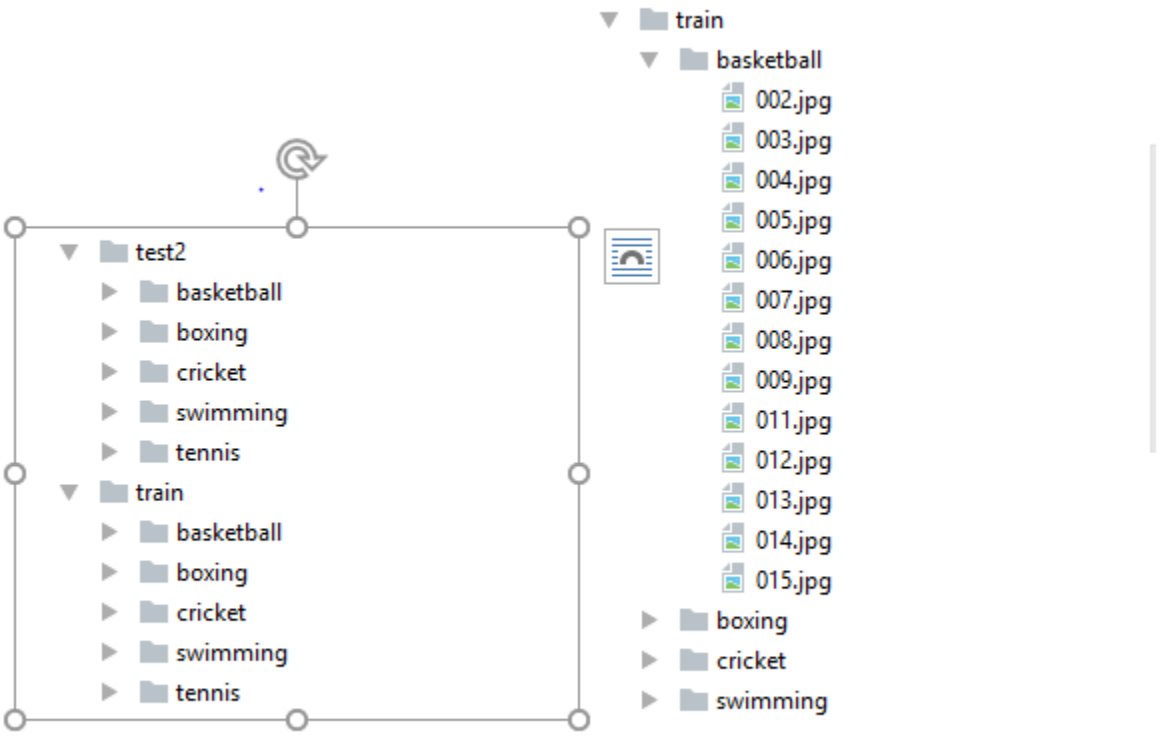
Solution:

In this Tutorial, I implemented the Image Classification for images in sports Category. I considered 5 different sports images as 5 classes and generated model using Decision tree as well as Random Forest. The 5 Classes are:

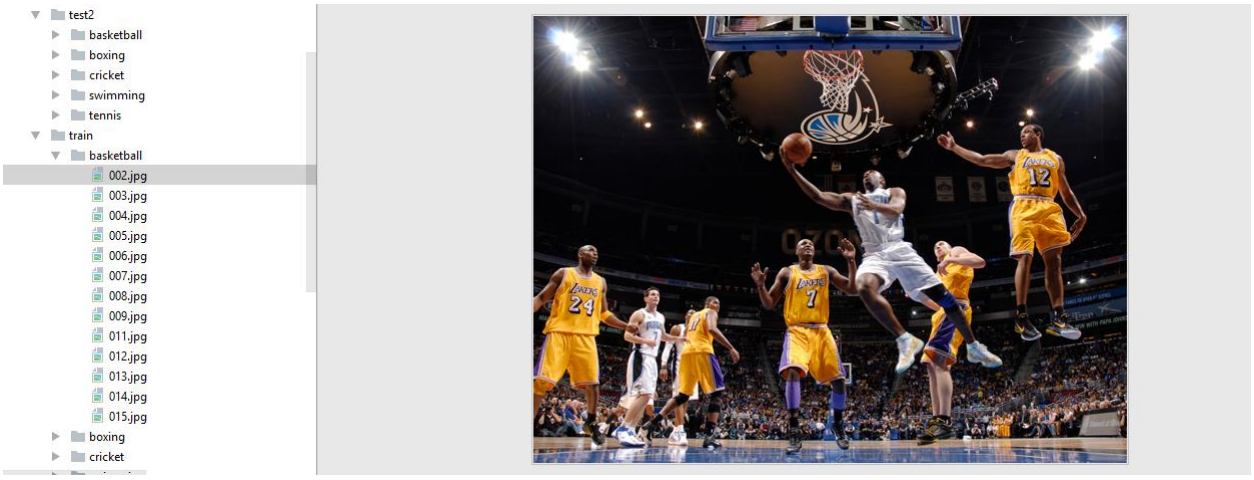
1. Basket Ball
2. Boxing
3. Cricket
4. Swimming
5. Tennis

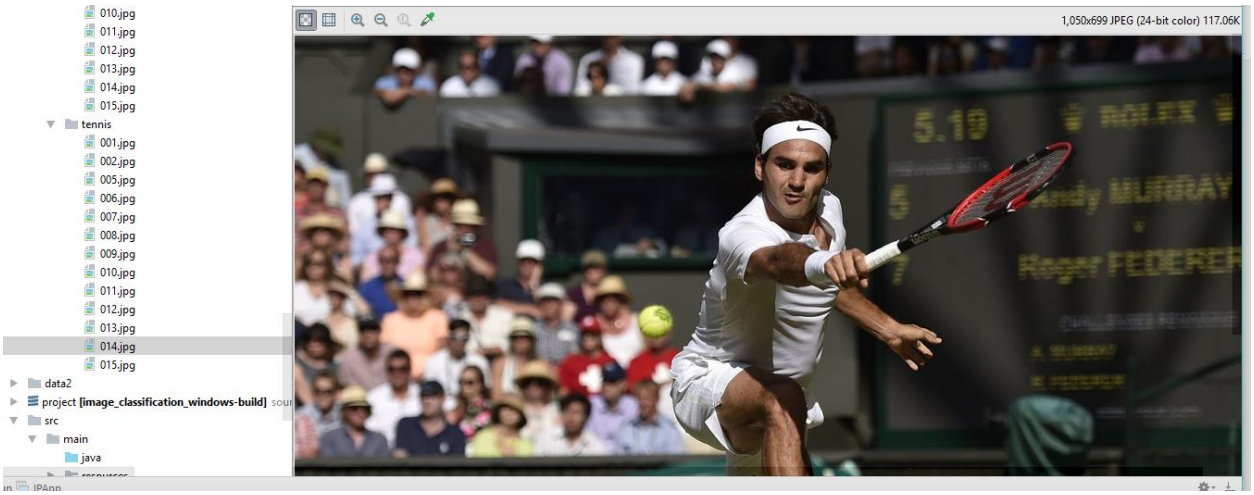
Training and Testing the Model:

Took 15 images in each category for training the model and 5 images in each category for testing the model.

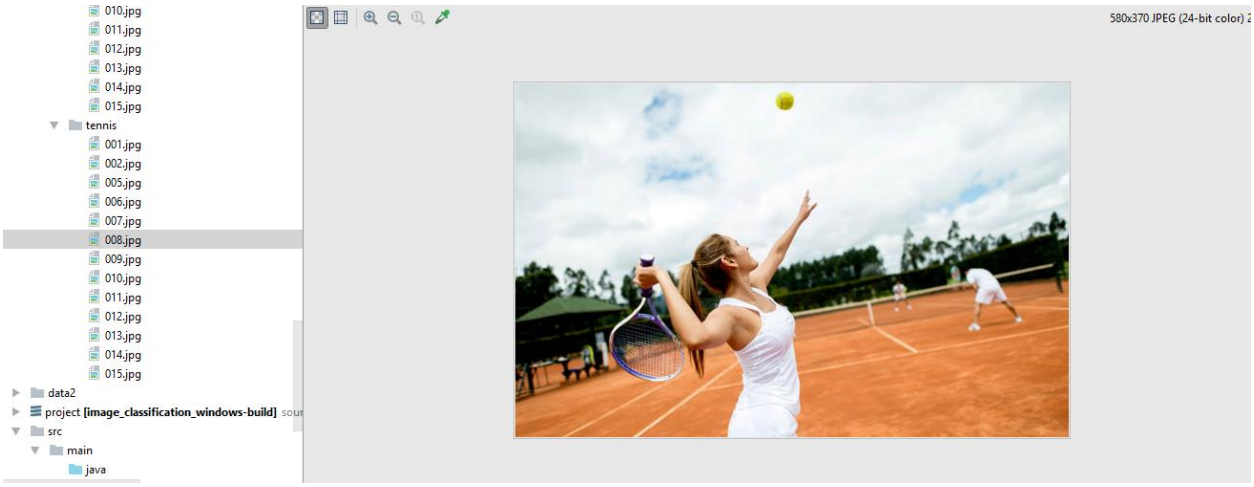


Sample Screen Shots of the training Data:

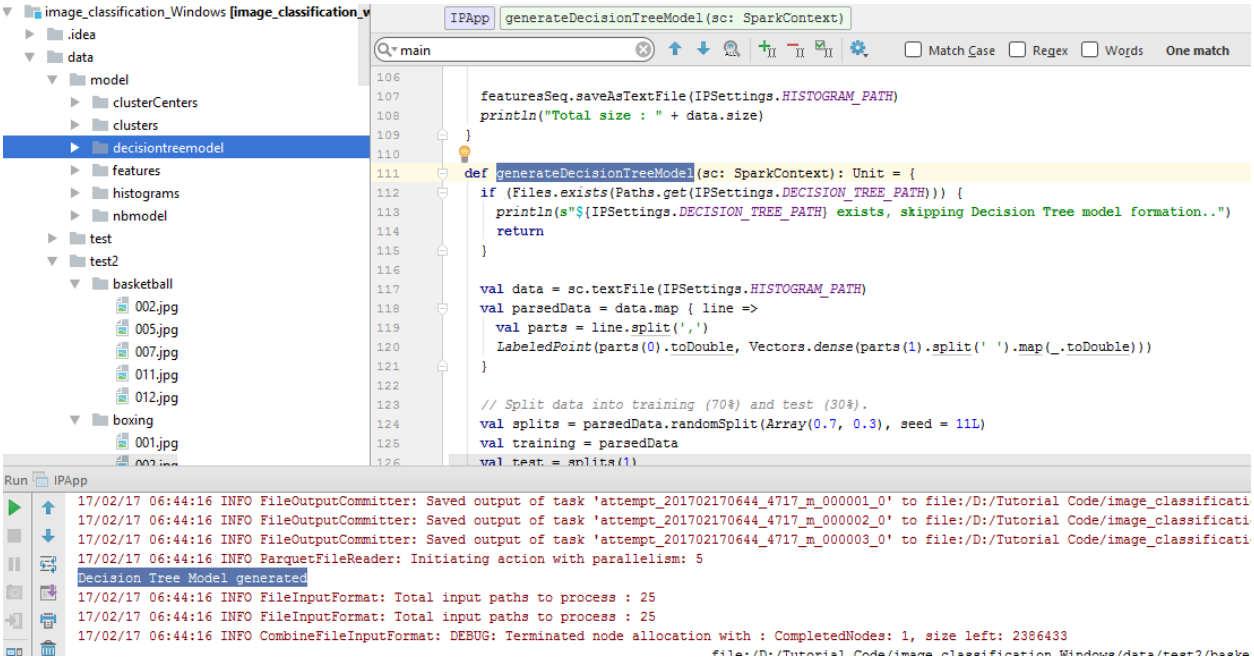



010.jpg
011.jpg
012.jpg
013.jpg
014.jpg
015.jpg
tennis
001.jpg
002.jpg
005.jpg
006.jpg
007.jpg
008.jpg
009.jpg
010.jpg
011.jpg
012.jpg
013.jpg
014.jpg
015.jpg
data2
project [image_classification_windows-build]
src
main
java
1,050x699 JPEG (24-bit color) 117.06K


012.jpg
013.jpg
014.jpg
015.jpg
swimming
001.jpg
002.jpg
003.jpg
004.jpg
005.jpg
006.jpg
007.jpg
008.jpg
009.jpg
010.jpg
011.jpg
012.jpg
013.jpg
014.jpg
015.jpg
tennis
data2
project [image_classification_windows-build]
src



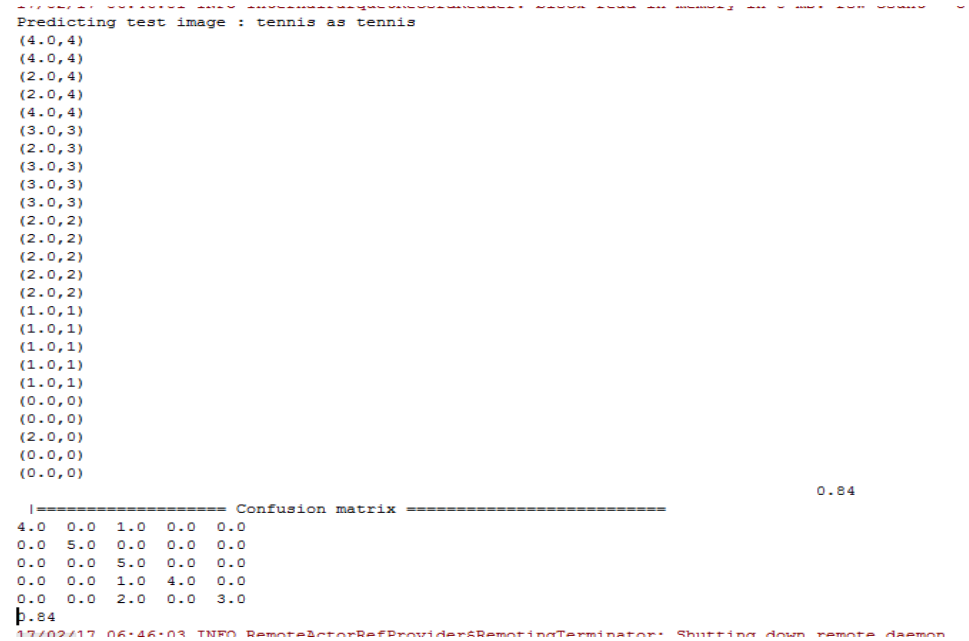
Screen shot of the generated model using decision Tree:



link for Source code:
https://github.com/gt784/BigDataAnalytictcsProject/tree/master/Source/Lab4/ImageClassification%20Using%20Decision%20Tree

Results:(Accuracy and Confusion Matrix)

Decision Tree: The model generated using decision tree has obtained 84% Accuracy. Below is the screen sot for accuracy and confusion Matrix



Random Forest

The model generated using random forest has obtained 72% accuracy. Below is the screen shot for confusion matrix and accuracy.