DAI\_ASSIGNMENT\_1\_Q1\_Report.pdf

**Dependable AI CSL7370 – Assignment 1**

**Task 1:**

**🡺Importing, Understanding and Pre-Processing the Data :**

First things first, we want to import the Celeb dataset into our workspace. In order to import the celeb dataset into the workspace we need to use google drive space. Using google colab’s inherent library from google.colab import drive Its very easy to mount google drive storage. Once mounted we get access to the storage. Auth token must be pulled when prompted while running it. In addition, the dataset available on the Kaggle has to be copied to the google drive so that we can unzip it and use it as physical files. Luckily, we have Kaggle dataset URL as well as the API tokens generated from the Kaggle such that we can send the Kaggle commands to bash to dump the dataset into our mounted Kaggle folder in Google Drive. On successful completion of the data import, it should look like as in figure 1.

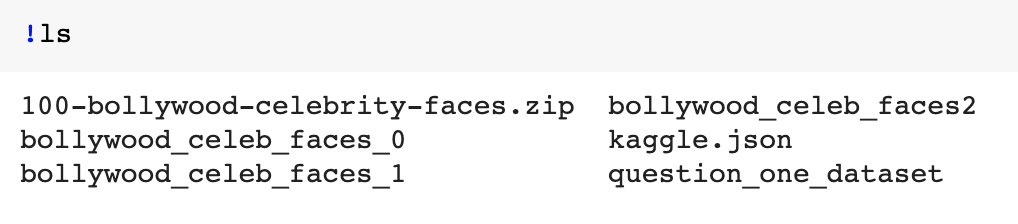


Figure 1 List show of the present working directory

Now that we have the dataset in our hands, import it. For ease, I have created a separate folder for 10 celebrities so to restrict ourselves within the single directory. The directory path looks like:

"/content/gdrive/My Drive/Kaggle/question\_one\_dataset"

Also, we can fetch all the subfolder names in the main directory using os.listdir to list the classes into single variable. In my case its shown in figure 2

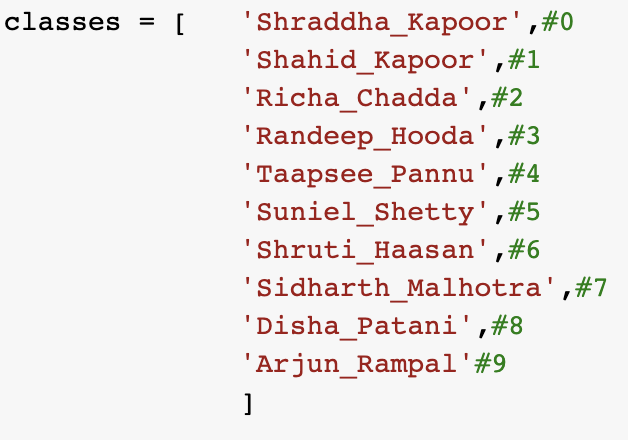


Figure 2 Ten Classes focussed on

One more check just to ensure we are in right place, we shall perform this command:

cv2\_imshow(cv2.imread("/content/gdrive/My Drive/Kaggle/question\_one\_dataset/Shraddha\_Kapoor/1.jpg"))

This results the image display of the file which is passed as path inside cv2.imread(). cv2.imread returns the image display of the file when given path of the file. Also cv2\_imshow is native to google colab which helps the image to render inside the colab notebook. Now that we have all images in jpg format inside the respective directories, we will now pre process them. Preprocessing includes pixel resizing, reading images as numpy arrays and most importantly assigning each image or pixel array of the image with its class label. Class labels are indexed from 0 to 9 for the ease. In order to perform this, the stepwise process is as follows

Initialize empty\_array[]

For each class\_label\_folder inside the present\_directory:

Assign path to the image,

Assign label as the index\_iteration\_number,

For each\_image inside the class\_label\_folder:

Assign image\_path

Actor\_image 🡸 Read each\_image

Actor\_image 🡸 numpy\_array(Actor\_image)

Append\_Inside\_Array[Actor\_image, class\_label]

\*\* reading the image using cv2 process has to be closed within the try catch block because if the file is not present it will return empty buffer.

Now at this stage we are just having the data on which we are going to apply training. Before that, as a best approach to generalize the learning process we can shuffle the dataset (stored inside the empty\_array[]). Finally, in order to divide the dataset (In our case we have 1198 images), perform train\_test\_split from sklearn library. As parameters you need to mention, test\_size which will tell the ratio how much data need to be assigned with test data. The status as of now should look like as in figure 3:

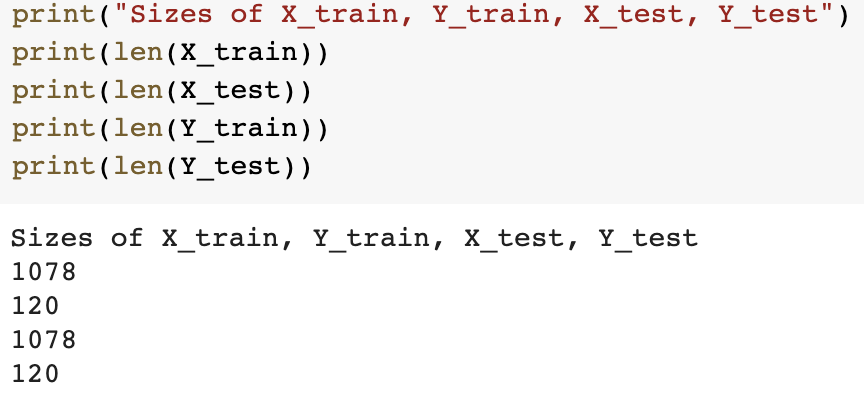


Figure 3 Train-Test-Split from Sklearn

**🡺 Training**

With the help of Sklearn library we can import SVC – Support Vector Classifier. The SVC() function trains the Training data model.fit(X\_train, Y\_train) and with the help of model.predict we can predict the test data. In order to evaluate the model we perform the training on training dataset such that first we evaluate the model on test dataset and also the training dataset itself. It is very intuitive that it will perform well when predicted with training data. And, with low accuracy on test data. One thing to note here is that I used various types of kernels in the SVM Learning process. I tried “rbf”, “linear” and “ploy”. Each of them are describes as: RBF – Radial Basis Functionis kind of gaussian methodology where the data is presented in higher dimension or in gassian space. Also it performs best when classes are separable. Linear Kernel also applied on linearly separable data but have poor optimization. Moreover on checking with multiple kernels the rbf performed better. decision\_function\_shape='ovo' Defines one over one classifier. In other words it identifies the underlying data and creates classifiers classwise, that is the inside classification is being done by checking the class vs class process. In our case since we have total 10 classes, the inner classifier has created a total of 10 choose 2 that is 45 different classifiers separating 2 classes. On evaluation the result we got is Testing accuracy: 0.333 and Training accuracy: 0.7597. This was done for overall accuracy. Now to find the class wise accuracy we will perform the training explicitly giving only 2 classes to the model. So that each time our model learns 2 classes and calculates the accuracy.

Figure 4 Text Box with classwise training results

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Shahid\_Kapoor

------------------------------------------------

0

1

264

Sizes of X\_train, Y\_train, X\_test, Y\_test

237

27

237

27

accuracy: 0.7407407407407407

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Richa\_Chadda

------------------------------------------------

0

1

237

Sizes of X\_train, Y\_train, X\_test, Y\_test

213

24

213

24

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Randeep\_Hooda

------------------------------------------------

0

1

234

Sizes of X\_train, Y\_train, X\_test, Y\_test

210

24

210

24

accuracy: 0.5833333333333334

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Taapsee\_Pannu

------------------------------------------------

0

1

264

Sizes of X\_train, Y\_train, X\_test, Y\_test

237

27

237

27

accuracy: 0.7407407407407407

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Suniel\_Shetty

------------------------------------------------

0

1

191

Sizes of X\_train, Y\_train, X\_test, Y\_test

171

20

171

20

accuracy: 0.7

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Shruti\_Haasan

------------------------------------------------

0

1

249

Sizes of X\_train, Y\_train, X\_test, Y\_test

224

25

224

25

accuracy: 0.56

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Sidharth\_Malhotra

------------------------------------------------

0

1

219

Sizes of X\_train, Y\_train, X\_test, Y\_test

197

22

197

22

accuracy: 0.5909090909090909

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Disha\_Patani

------------------------------------------------

0

1

271

Sizes of X\_train, Y\_train, X\_test, Y\_test

243

28

243

28

accuracy: 0.8214285714285714

------------------------------------------------

-----------Doing for ---------

Shraddha\_Kapoor and Arjun\_Rampal

------------------------------------------------

0

1

221

Sizes of X\_train, Y\_train, X\_test, Y\_test

198

23

198

23

accuracy: 0.5652173913043478

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Shraddha\_Kapoor

------------------------------------------------

0

1

264

Sizes of X\_train, Y\_train, X\_test, Y\_test

237

27

237

27

accuracy: 0.6666666666666666

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Richa\_Chadda

------------------------------------------------

0

1

263

Sizes of X\_train, Y\_train, X\_test, Y\_test

236

27

236

27

accuracy: 0.6666666666666666

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Randeep\_Hooda

------------------------------------------------

0

1

260

Sizes of X\_train, Y\_train, X\_test, Y\_test

234

26

234

26

accuracy: 0.6153846153846154

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Taapsee\_Pannu

------------------------------------------------

0

1

290

Sizes of X\_train, Y\_train, X\_test, Y\_test

261

29

261

29

accuracy: 0.7931034482758621

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Suniel\_Shetty

------------------------------------------------

0

1

217

Sizes of X\_train, Y\_train, X\_test, Y\_test

195

22

195

22

accuracy: 0.4090909090909091

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Shruti\_Haasan

------------------------------------------------

0

1

275

Sizes of X\_train, Y\_train, X\_test, Y\_test

247

28

247

28

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Sidharth\_Malhotra

------------------------------------------------

0

1

245

Sizes of X\_train, Y\_train, X\_test, Y\_test

220

25

220

25

accuracy: 0.56

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Disha\_Patani

------------------------------------------------

0

1

297

Sizes of X\_train, Y\_train, X\_test, Y\_test

267

30

267

30

accuracy: 0.7

------------------------------------------------

-----------Doing for ---------

Shahid\_Kapoor and Arjun\_Rampal

------------------------------------------------

0

1

247

Sizes of X\_train, Y\_train, X\_test, Y\_test

222

25

222

25

accuracy: 0.6

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Shraddha\_Kapoor

------------------------------------------------

0

1

237

Sizes of X\_train, Y\_train, X\_test, Y\_test

213

24

213

24

accuracy: 0.7083333333333334

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Shahid\_Kapoor

------------------------------------------------

0

1

263

Sizes of X\_train, Y\_train, X\_test, Y\_test

236

27

236

27

accuracy: 0.7037037037037037

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Randeep\_Hooda

------------------------------------------------

0

1

233

Sizes of X\_train, Y\_train, X\_test, Y\_test

209

24

209

24

accuracy: 0.7083333333333334

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Taapsee\_Pannu

------------------------------------------------

0

1

263

Sizes of X\_train, Y\_train, X\_test, Y\_test

236

27

236

27

accuracy: 0.5185185185185185

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Suniel\_Shetty

------------------------------------------------

0

1

190

Sizes of X\_train, Y\_train, X\_test, Y\_test

171

19

171

19

accuracy: 0.631578947368421

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Shruti\_Haasan

------------------------------------------------

0

1

248

Sizes of X\_train, Y\_train, X\_test, Y\_test

223

25

223

25

accuracy: 0.68

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Sidharth\_Malhotra

------------------------------------------------

0

1

218

Sizes of X\_train, Y\_train, X\_test, Y\_test

196

22

196

22

accuracy: 0.6818181818181818

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Disha\_Patani

------------------------------------------------

0

1

270

Sizes of X\_train, Y\_train, X\_test, Y\_test

243

27

243

27

accuracy: 0.6296296296296297

------------------------------------------------

-----------Doing for ---------

Richa\_Chadda and Arjun\_Rampal

------------------------------------------------

0

1

220

Sizes of X\_train, Y\_train, X\_test, Y\_test

198

22

198

22

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Shraddha\_Kapoor

------------------------------------------------

0

1

234

Sizes of X\_train, Y\_train, X\_test, Y\_test

210

24

210

24

accuracy: 0.5833333333333334

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Shahid\_Kapoor

------------------------------------------------

0

1

260

Sizes of X\_train, Y\_train, X\_test, Y\_test

234

26

234

26

accuracy: 0.5769230769230769

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Richa\_Chadda

------------------------------------------------

0

1

233

Sizes of X\_train, Y\_train, X\_test, Y\_test

209

24

209

24

accuracy: 0.5833333333333334

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Taapsee\_Pannu

------------------------------------------------

0

1

260

Sizes of X\_train, Y\_train, X\_test, Y\_test

234

26

234

26

accuracy: 0.6538461538461539

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Suniel\_Shetty

------------------------------------------------

0

1

187

Sizes of X\_train, Y\_train, X\_test, Y\_test

168

19

168

19

accuracy: 0.47368421052631576

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Shruti\_Haasan

------------------------------------------------

0

1

245

Sizes of X\_train, Y\_train, X\_test, Y\_test

220

25

220

25

accuracy: 0.76

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Sidharth\_Malhotra

------------------------------------------------

0

1

215

Sizes of X\_train, Y\_train, X\_test, Y\_test

193

22

193

22

accuracy: 0.45454545454545453

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Disha\_Patani

------------------------------------------------

0

1

267

Sizes of X\_train, Y\_train, X\_test, Y\_test

240

27

240

27

accuracy: 0.7777777777777778

------------------------------------------------

-----------Doing for ---------

Randeep\_Hooda and Arjun\_Rampal

------------------------------------------------

0

1

217

Sizes of X\_train, Y\_train, X\_test, Y\_test

195

22

195

22

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Shraddha\_Kapoor

------------------------------------------------

0

1

264

Sizes of X\_train, Y\_train, X\_test, Y\_test

237

27

237

27

accuracy: 0.7777777777777778

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Shahid\_Kapoor

------------------------------------------------

0

1

290

Sizes of X\_train, Y\_train, X\_test, Y\_test

261

29

261

29

accuracy: 0.6551724137931034

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Richa\_Chadda

------------------------------------------------

0

1

263

Sizes of X\_train, Y\_train, X\_test, Y\_test

236

27

236

27

accuracy: 0.6296296296296297

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Randeep\_Hooda

------------------------------------------------

0

1

260

Sizes of X\_train, Y\_train, X\_test, Y\_test

234

26

234

26

accuracy: 0.6923076923076923

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Suniel\_Shetty

------------------------------------------------

0

1

217

Sizes of X\_train, Y\_train, X\_test, Y\_test

195

22

195

22

accuracy: 0.7727272727272727

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Shruti\_Haasan

------------------------------------------------

0

1

275

Sizes of X\_train, Y\_train, X\_test, Y\_test

247

28

247

28

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Sidharth\_Malhotra

------------------------------------------------

0

1

245

Sizes of X\_train, Y\_train, X\_test, Y\_test

220

25

220

25

accuracy: 0.72

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Disha\_Patani

------------------------------------------------

0

1

297

Sizes of X\_train, Y\_train, X\_test, Y\_test

267

30

267

30

accuracy: 0.4666666666666667

------------------------------------------------

-----------Doing for ---------

Taapsee\_Pannu and Arjun\_Rampal

------------------------------------------------

0

1

247

Sizes of X\_train, Y\_train, X\_test, Y\_test

222

25

222

25

accuracy: 0.56

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Shraddha\_Kapoor

------------------------------------------------

0

1

191

Sizes of X\_train, Y\_train, X\_test, Y\_test

171

20

171

20

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Shahid\_Kapoor

------------------------------------------------

0

1

217

Sizes of X\_train, Y\_train, X\_test, Y\_test

195

22

195

22

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Richa\_Chadda

------------------------------------------------

0

1

190

Sizes of X\_train, Y\_train, X\_test, Y\_test

171

19

171

19

accuracy: 0.5789473684210527

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Randeep\_Hooda

------------------------------------------------

0

1

187

Sizes of X\_train, Y\_train, X\_test, Y\_test

168

19

168

19

accuracy: 0.6842105263157895

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Taapsee\_Pannu

------------------------------------------------

0

1

217

Sizes of X\_train, Y\_train, X\_test, Y\_test

195

22

195

22

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Shruti\_Haasan

------------------------------------------------

0

1

202

Sizes of X\_train, Y\_train, X\_test, Y\_test

181

21

181

21

accuracy: 0.7142857142857143

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Sidharth\_Malhotra

------------------------------------------------

0

1

172

Sizes of X\_train, Y\_train, X\_test, Y\_test

154

18

154

18

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Disha\_Patani

------------------------------------------------

0

1

224

Sizes of X\_train, Y\_train, X\_test, Y\_test

201

23

201

23

accuracy: 0.7391304347826086

------------------------------------------------

-----------Doing for ---------

Suniel\_Shetty and Arjun\_Rampal

------------------------------------------------

0

1

174

Sizes of X\_train, Y\_train, X\_test, Y\_test

156

18

156

18

accuracy: 0.7222222222222222

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Shraddha\_Kapoor

------------------------------------------------

0

1

249

Sizes of X\_train, Y\_train, X\_test, Y\_test

224

25

224

25

accuracy: 0.6

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Shahid\_Kapoor

------------------------------------------------

0

1

275

Sizes of X\_train, Y\_train, X\_test, Y\_test

247

28

247

28

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Richa\_Chadda

------------------------------------------------

0

1

248

Sizes of X\_train, Y\_train, X\_test, Y\_test

223

25

223

25

accuracy: 0.84

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Randeep\_Hooda

------------------------------------------------

0

1

245

Sizes of X\_train, Y\_train, X\_test, Y\_test

220

25

220

25

accuracy: 0.72

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Taapsee\_Pannu

------------------------------------------------

0

1

275

Sizes of X\_train, Y\_train, X\_test, Y\_test

247

28

247

28

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Suniel\_Shetty

------------------------------------------------

0

1

202

Sizes of X\_train, Y\_train, X\_test, Y\_test

181

21

181

21

accuracy: 0.5238095238095238

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Sidharth\_Malhotra

------------------------------------------------

0

1

230

Sizes of X\_train, Y\_train, X\_test, Y\_test

207

23

207

23

accuracy: 0.6086956521739131

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Disha\_Patani

------------------------------------------------

0

1

282

Sizes of X\_train, Y\_train, X\_test, Y\_test

253

29

253

29

accuracy: 0.6551724137931034

------------------------------------------------

-----------Doing for ---------

Shruti\_Haasan and Arjun\_Rampal

------------------------------------------------

0

1

232

Sizes of X\_train, Y\_train, X\_test, Y\_test

208

24

208

24

accuracy: 0.75

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Shraddha\_Kapoor

------------------------------------------------

0

1

219

Sizes of X\_train, Y\_train, X\_test, Y\_test

197

22

197

22

accuracy: 0.7727272727272727

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Shahid\_Kapoor

------------------------------------------------

0

1

245

Sizes of X\_train, Y\_train, X\_test, Y\_test

220

25

220

25

accuracy: 0.56

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Richa\_Chadda

------------------------------------------------

0

1

218

Sizes of X\_train, Y\_train, X\_test, Y\_test

196

22

196

22

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Randeep\_Hooda

------------------------------------------------

0

1

215

Sizes of X\_train, Y\_train, X\_test, Y\_test

193

22

193

22

accuracy: 0.4090909090909091

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Taapsee\_Pannu

------------------------------------------------

0

1

245

Sizes of X\_train, Y\_train, X\_test, Y\_test

220

25

220

25

accuracy: 0.48

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Suniel\_Shetty

------------------------------------------------

0

1

172

Sizes of X\_train, Y\_train, X\_test, Y\_test

154

18

154

18

accuracy: 0.5555555555555556

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Shruti\_Haasan

------------------------------------------------

0

1

230

Sizes of X\_train, Y\_train, X\_test, Y\_test

207

23

207

23

accuracy: 0.6956521739130435

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Disha\_Patani

------------------------------------------------

0

1

252

Sizes of X\_train, Y\_train, X\_test, Y\_test

226

26

226

26

accuracy: 0.6538461538461539

------------------------------------------------

-----------Doing for ---------

Sidharth\_Malhotra and Arjun\_Rampal

------------------------------------------------

0

1

202

Sizes of X\_train, Y\_train, X\_test, Y\_test

181

21

181

21

accuracy: 0.6190476190476191

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Shraddha\_Kapoor

------------------------------------------------

0

1

271

Sizes of X\_train, Y\_train, X\_test, Y\_test

243

28

243

28

accuracy: 0.7857142857142857

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Shahid\_Kapoor

------------------------------------------------

0

1

297

Sizes of X\_train, Y\_train, X\_test, Y\_test

267

30

267

30

accuracy: 0.8

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Richa\_Chadda

------------------------------------------------

0

1

270

Sizes of X\_train, Y\_train, X\_test, Y\_test

243

27

243

27

accuracy: 0.48148148148148145

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Randeep\_Hooda

------------------------------------------------

0

1

267

Sizes of X\_train, Y\_train, X\_test, Y\_test

240

27

240

27

accuracy: 0.8888888888888888

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Taapsee\_Pannu

------------------------------------------------

0

1

297

Sizes of X\_train, Y\_train, X\_test, Y\_test

267

30

267

30

accuracy: 0.6666666666666666

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Suniel\_Shetty

------------------------------------------------

0

1

224

Sizes of X\_train, Y\_train, X\_test, Y\_test

201

23

201

23

accuracy: 0.782608695652174

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Shruti\_Haasan

------------------------------------------------

0

1

282

Sizes of X\_train, Y\_train, X\_test, Y\_test

253

29

253

29

accuracy: 0.6206896551724138

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Sidharth\_Malhotra

------------------------------------------------

0

1

252

Sizes of X\_train, Y\_train, X\_test, Y\_test

226

26

226

26

accuracy: 0.6153846153846154

------------------------------------------------

-----------Doing for ---------

Disha\_Patani and Arjun\_Rampal

------------------------------------------------

0

1

254

Sizes of X\_train, Y\_train, X\_test, Y\_test

228

26

228

26

accuracy: 0.7692307692307693

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Shraddha\_Kapoor

------------------------------------------------

0

1

221

Sizes of X\_train, Y\_train, X\_test, Y\_test

198

23

198

23

accuracy: 0.6956521739130435

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Shahid\_Kapoor

------------------------------------------------

0

1

247

Sizes of X\_train, Y\_train, X\_test, Y\_test

222

25

222

25

accuracy: 0.76

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Richa\_Chadda

------------------------------------------------

0

1

220

Sizes of X\_train, Y\_train, X\_test, Y\_test

198

22

198

22

accuracy: 0.7272727272727273

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Randeep\_Hooda

------------------------------------------------

0

1

217

Sizes of X\_train, Y\_train, X\_test, Y\_test

195

22

195

22

accuracy: 0.5

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Taapsee\_Pannu

------------------------------------------------

0

1

247

Sizes of X\_train, Y\_train, X\_test, Y\_test

222

25

222

25

accuracy: 0.68

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Suniel\_Shetty

------------------------------------------------

0

1

174

Sizes of X\_train, Y\_train, X\_test, Y\_test

156

18

156

18

accuracy: 0.7777777777777778

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Shruti\_Haasan

------------------------------------------------

0

1

232

Sizes of X\_train, Y\_train, X\_test, Y\_test

208

24

208

24

accuracy: 0.6666666666666666

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Sidharth\_Malhotra

------------------------------------------------

0

1

202

Sizes of X\_train, Y\_train, X\_test, Y\_test

181

21

181

21

accuracy: 0.6190476190476191

------------------------------------------------

-----------Doing for ---------

Arjun\_Rampal and Disha\_Patani

------------------------------------------------

0

1

254

Sizes of X\_train, Y\_train, X\_test, Y\_test

228

26

228

26

accuracy: 0.7307692307692307

On further checking singleton results the results were not so good. The model mis-classifies most of the stars. Even on multiple EPOCHS, the training results were not so good.

**🡺 Results and Inferences:**

Overall Training Accuracy = ~76%

Overall Testing Accuracy = ~33%

Class-wise Accuracies are attached as text box in figure 4.



Figure 5 Output of Classwise training and accuracies

Two Metrics used are:

1. Classification report (from sklearn library)
2. Confusion Matrix (from sklearn library)

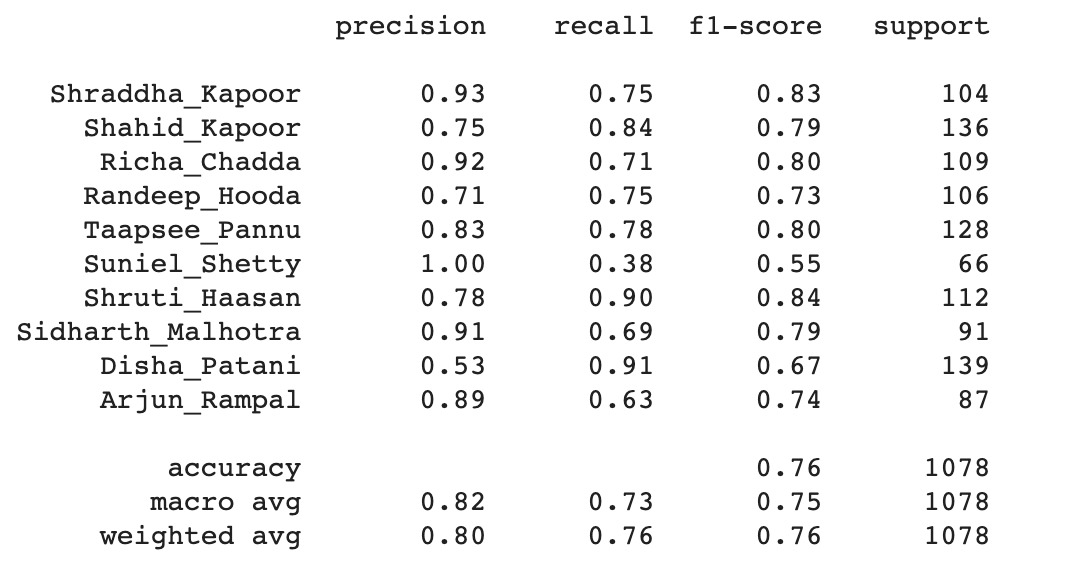


Figure 6 Training Metrics

From figure 6 defines the key metrics of the classification problem of training dataset. To define the terminologies, **precision** is the accuracy within class or in other words, how many are correctly classified within that class. That is you predicted positive and it turned out to be positive in labels as well. Precision = number of positives / total positive labels.

For example, for Shraddha Kapoor 93 examples out of 100 samples are correctly classified or for Richa Chadda 92 examples out of 100 samples are correctly classified where 100 samples are from their respective classes. Accuracy = # of correct classifications / total examples. And **recall** can be defined as number of positives given that the labels are positive or the ground truth is also positive and the predicted output is also positive.

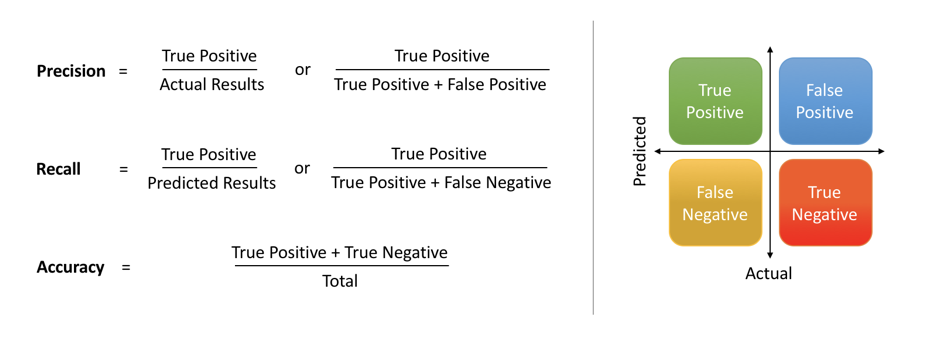


Figure 7 Precision Recall chart

In more simpler words, **Precision** is the number of times the actor (say Shraddha Kapoor) got correctly classified divided by number of times the model assumed it was Shraddha Kapoor. However **Recall** is number of times model classified it was Shraddha Kapoor divided by the number of times it was actually Shraddha Kapoor. **F1** score is the average of precision and recall. In addition, support is the amount of percentage of that class present in the dataset (here it is training set).

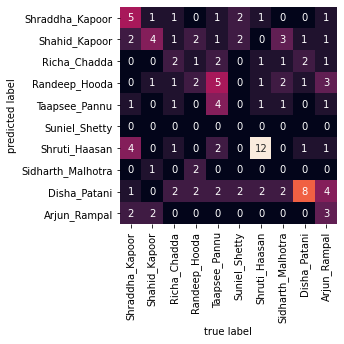


Figure 8 Confusion Matrix for Celeb data

**Confusion Matrix** is also one of the metrics to check how our model performed overall. When analyzed, it is clearly seen as the model performed best on Shruti Haasan. We can find following conclusions:

1. The model predicted Shruti Haasan 12 times correctly. That is the model said the sample is Shruti Haasan and in ground truth it is being Shruti Haasan.
2. The model predicted Disha 8 times correctly. That is the model said the sample is Disha and in ground truth it is being Disha.
3. The model predicted Tapsee Pannu as Randeep Hooda 5 times. Which is more than its true positive (that is 4).
4. Even if you calculate the prediction row of Disha Patani the row sum is highest of Disha Patani. Additionally, in the training data we had the highest number of Disha Patani examples of images. This is clearly a bias towards Disha Patani being predicted for multiple actors.
5. In second observation, of Shruti Haasan, there is a bias towards it. This can be checked as Shraddha Kapoor being classified as Shruti Haasan.
6. Seems like For Suniel Shetty the model learned nothing. The model predicted Suneil Shetty being Suniel Shetty is 0 and model predicted Suniel Shetty being Disha Patani is 2. Clearly a bias towards Disha Patani.

Of course there are numerous observations that can be carried from the data. I have mentioned the highest possibilities of bias towards some classes.