Face\_recognition Project Report

Yusif Imamverdiyev M1DSAI

Introduction to data science

Project name : **Face recognition [hard]**

For loading our data i used **fetch\_lfw\_people** function from sklearn . Face recognition is the supervised classification task of identifying a person from an image. Our data set consist 6 peoples images. Our images are small. A feature vector that encodes the intensity of every pixel will have high dimensions. Training from such high dimensional data colud require many samples to avoid over-fitting . Instead , we will use PCA to compactly represent the images in terms of a small number of principal components .

I then randomly split images into train and test sets and fit the PCA object on training set X\_train,X\_test,y\_train,y\_test=train\_test\_split(X,y,random\_state=11)

pca=PCA(n\_components=150,whiten=True)

I reduce all of the instances to 150 dimensions and train a logistic regression classifier. The data set contains classes, Scikit learn automatically creates binary classifier using the one versus all strategy behind the scens

X\_train\_PCA=pca.fit\_transform(X\_train)

X\_test\_PCA=pca.transform(X\_test)

print(X\_train.shape)

print(X\_train\_reduced.shape)

Finally we evaluate the performance of the classifier using cross-validation and a test set. The avereage per-vlass F1-score of the classifier trained on the full data was 0.77 , but required significantly more time to train and could be prohibitively slow in an application with more training instances.

Follow output of the script .

