

Detection of whether the person is wearing mask or not.

Akshaykumar Kanani(B19EE008), Student, IIT Jodhpur,
Jyani Akshay Jagdishbhai(B19EE041), Student, IIT Jodhpur

Abstract: Today's world is suffering from huge crisis SARS-CoV. So, wearing masks has become a necessity of today's world. According to senior scholar and infectious disease expert at the Johns Hopkins Center for Health Security, Adalja.Amesh A, MD, "Face masks can help protect against many respiratory infections that are spread through the air, droplet route, and that includes flu and also coronavirus. So, in this project we have used a dataset containing nearly more than 8000 images containing images with and without Mask, on which we have performed preprocessing, HaarCascades, feature selection, then applied different classifiers to predict the accuracy.

Index Terms: Image preprocessing, Feature Selection, Cross Validation, Opencv



1. Introduction

This project has been done as a part of our course of Pattern Recognition and Machine Learning(PRML) at Indian Institute of Technology Jodhpur, under the supervision of Dr. Richa Singh. In this project we have determined whether the person is wearing masks or not.

2. Working of the Code

Under this section, we will discuss the concepts involved and the working of the code.

A. Libraries and preprocessing of data

First we have imported all the necessary libraries such as pandas, numpy, matplotlib.pyplot and all the libraries required to run different classifiers and all. After that we have uploaded the dataset containing 2000 images of the person in which 1000 are wearing mask and remaining are not.. Then we assigned the labels to for target,

Then we convert the images into the array using opencv. And we are ready to make the models.

B. Extracting only Face Image from Whole Pictures

We use Viola-Jones Face Detection Technique(also known as Haar Cascades)

C. Data split in train and test

We split the data in 50-50 % as per the instruction.

D. Classifiers with and without Feature selection(PCA)

- **Decision Tree Classifier:** Decision Tree Classifier is a simple and widely used classification technique. It applies a straightforward idea to solve the classification problem. Decision Tree Classifier poses a series of carefully crafted questions about the attributes of the test record. Each time it receives an answer, a follow-up question is asked until a conclusion about the class label of the record is reached.
- **Random Forest Classifier:** Random forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees. In this we have used `n_estimators=1000` and `random_state=42`.
- **Logistic Regression:** Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression is estimating the parameters of a logistic model.
- **SVM:** SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes.

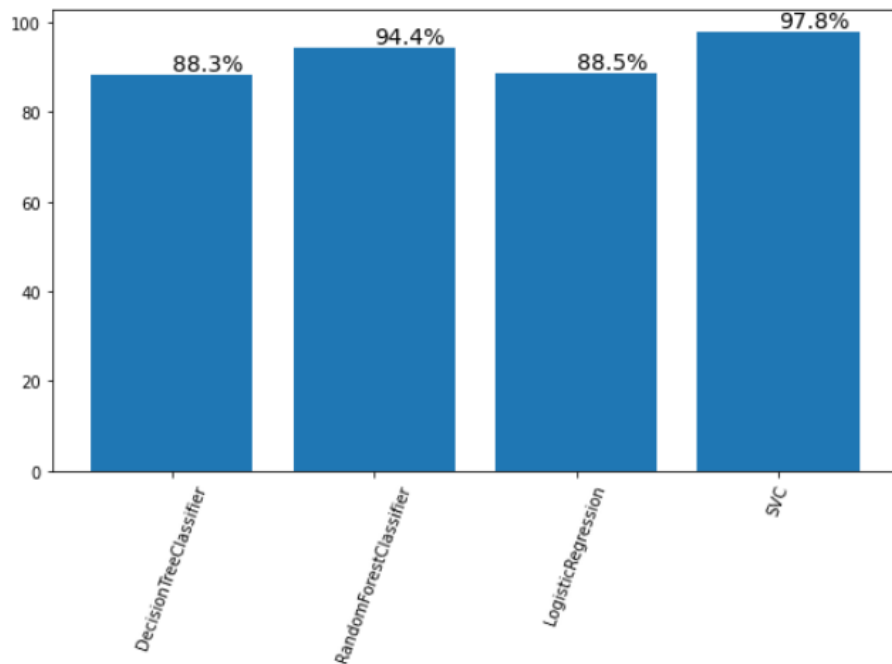
3. Conclusion

Before Feature Selection

After applying Four different classification algorithms we got the following results as shown in the table:

Classifier	Precision		Recall		F1 Score		Accuracy
	No-Mask	Mask	No-Mask	Mask	No-Mask	Mask	
Decision Tree	0.88	0.89	0.90	0.87	0.89	0.88	88.3
Random Forest	0.96	0.93	0.93	0.96	0.94	0.94	94.4
Logistic Regression	0.82	0.99	0.98	0.77	0.90	0.87	88.5
SVC	0.98	0.98	0.98	0.98	0.98	0.98	97.8

Then we plotted the graph of accuracies of different classifiers used:



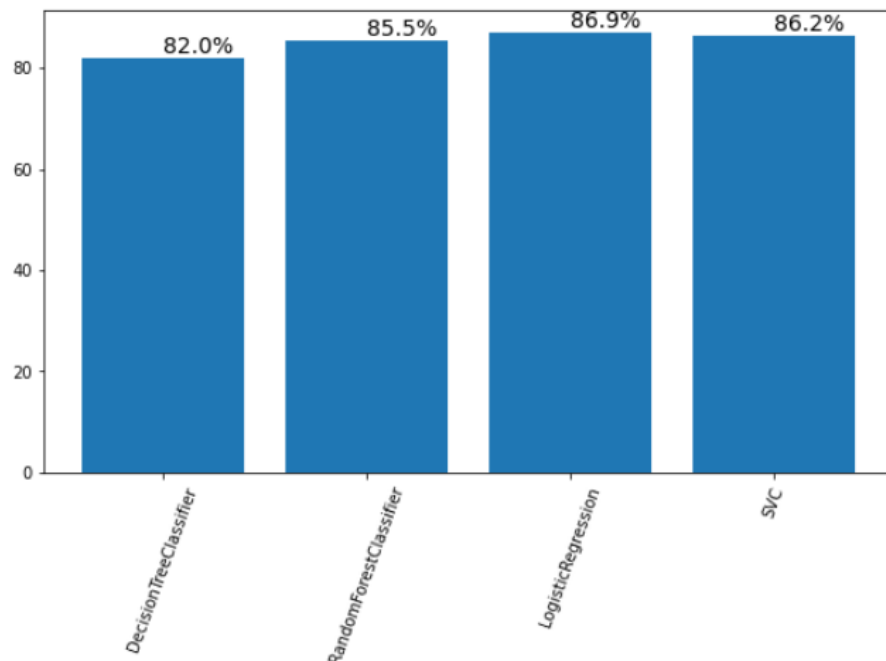
From the above table and plot we can see that SVM has the highest accuracy score of 97.8% and Decision Tree classifier has the least accuracy score of 83.3%. And the mean accuracy score of all the classifiers used is 92.25%.

After Feature Selection(PCA) :

After applying Four different classification algorithms we got the following results as shown in the table:

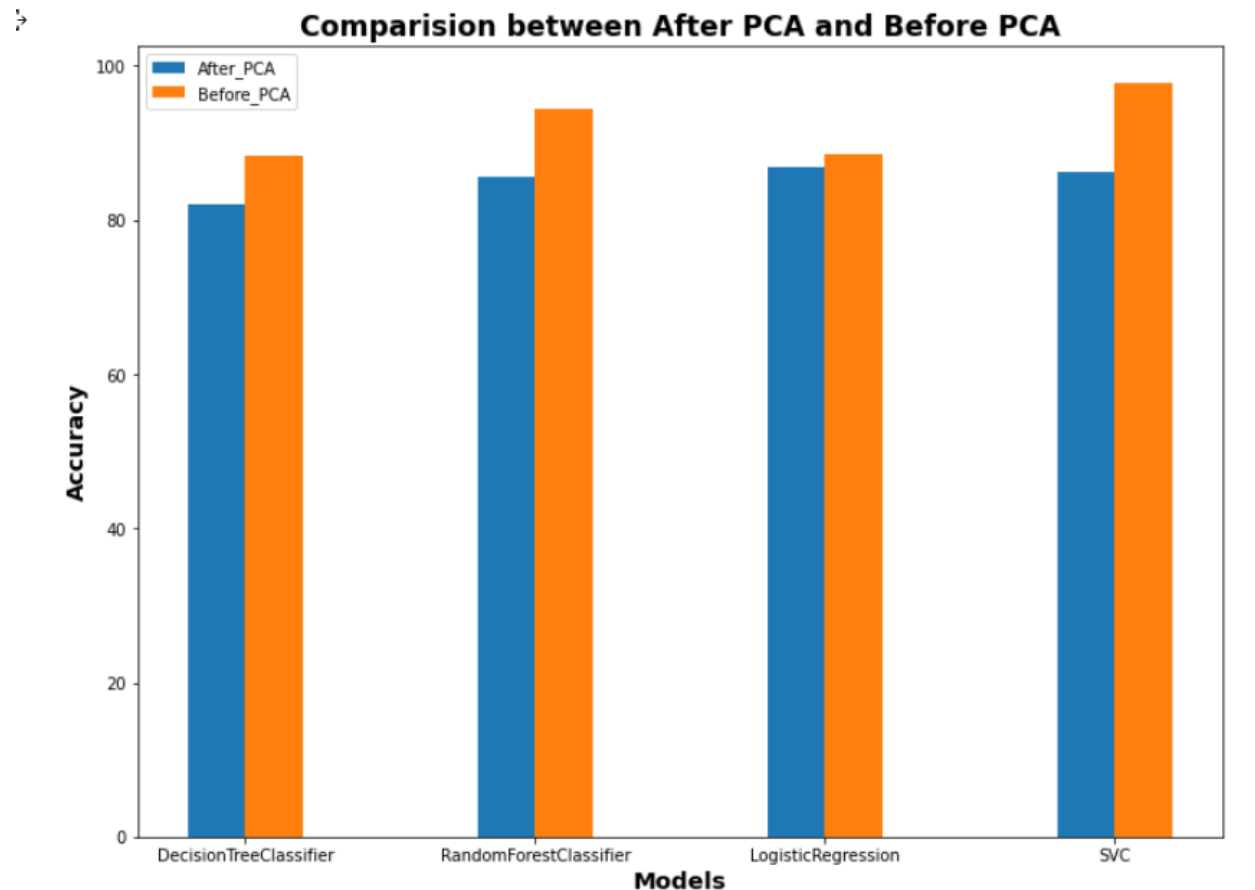
Classifier	Precision		Recall		F1 Score		Accuracy
	No-Mask	Mask	No-Mask	Mask	No-Mask	Mask	
Decision Tree	0.84	0.80	0.80	0.84	0.82	0.82	82.0
Random Forest	0.87	0.84	0.84	0.87	0.86	0.85	85.5
Logistic Regression	0.87	0.86	0.87	0.87	0.87	0.87	86.9
SVC	0.88	0.85	0.85	0.88	0.86	0.86	86.2

Then we plotted the graph of accuracies of different classifiers used:



From the above table and plot we can see that Logistic Regression has the highest accuracy score of 86.9% and Decision Tree classifier has the least accuracy score of 82.0%. And the mean accuracy score of all the classifiers used is 85.15%.

For better visualization we have:



Acknowledgement

We would like to express our gratitude to our course instructor Dr Richa Singh for her immense guidance and support throughout the course of Pattern Recognition and Machine Learning. We would also like to thank our lab instructors Dr Romi Banerjee and Dr Yashaswi Verma for their guidance and support throughout the lab sessions of the course and any other time. We would also like to express our gratitude towards all the teaching assistants involved in this course for helping us whenever we needed. And last but not the least, we would like to thank our batchmates for their immense help and support. Without their support and guidance this project would not have been completed.

References

- <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>
- <https://towardsdatascience.com/face-detection-with-haar-cascade-727f68dafd08>
- <https://towardsdatascience.com/pca-is-not-feature-selection-3344fb764ae6#:~:text=The%20only%20way%20PCA%20is,this%20is%20usually%20not%20true.&text=Once%20you've%20completed%20PCA,combination%20of%20the%20old%20variables.>

Contribution of Member

Akshaykumar Kanani (B19EE008) & Jyani Akshay Jagdishbhai (B19EE041)

We both of us has written the code for has written the code for model training and testing, Image Processing, Extracting Face Image from whole Pictures, DecisionTreeClassifier, RandomForestClassifier, LogisticRegression, SVC ,and also use the dimensionality reduction technique (PCA) and finally compare the Accuracy, F1 score and Classification Report of Before PCA and After PCA result by drawing Bar plot, We both did all work together by joining Gmeet.

Code debugging and finding the references were done by all the members of the Team.