FACE MASK DETECTION

Vansh Khandelwal B19EE088

ABSTRACT

Here we have a Dataset which consists of a large number of images which consists of both masked image and unmasked images and we have to classify the images accordingly with the help of different classifiers like Decision Tree Classifier and MLP Classifier and Logistics Regression Classifier.

INTRODUCTION

Face recognition represents one of the most interesting modalities of biometric. Due to his low intrusiveness and to the constant decrease in image acquisition cost, it's particularly suitable for a wide number of real time applications. In this paper we propose a very fast image pre-processing by the introduction of a linearly shaded elliptical mask centered over the faces. Facial recognition studies the features around the eye, nose, mouth, and ears to identify an individual to determine whether the individual is wearing a mask or not.

RELATED WORK

- First we uploaded the data on drive and then mounted the drive on collab.
- Extracting the data of masked and unmasked faces from zip file.
- Resize all images in 100 x 100 so that all images have the same size.
- Splitting the dataset in training and testing half and reducing the data so we can easily train models.
- Reducing the features so we can cross validate the results.
- We used PCA for feature reduction.

  Principal Component Analysis or PCA is a widely used technique for dimensionality

reduction of the large data set. Reducing the number of components or features costs some accuracy and on the other hand, it makes the large data set simpler, easy to explore and visualize. Also, it reduces the computational complexity of the model which makes machine learning algorithms run faster. It is always a question and debatable how much accuracy it is sacrificing to get a less complex and reduced dimensions data set.

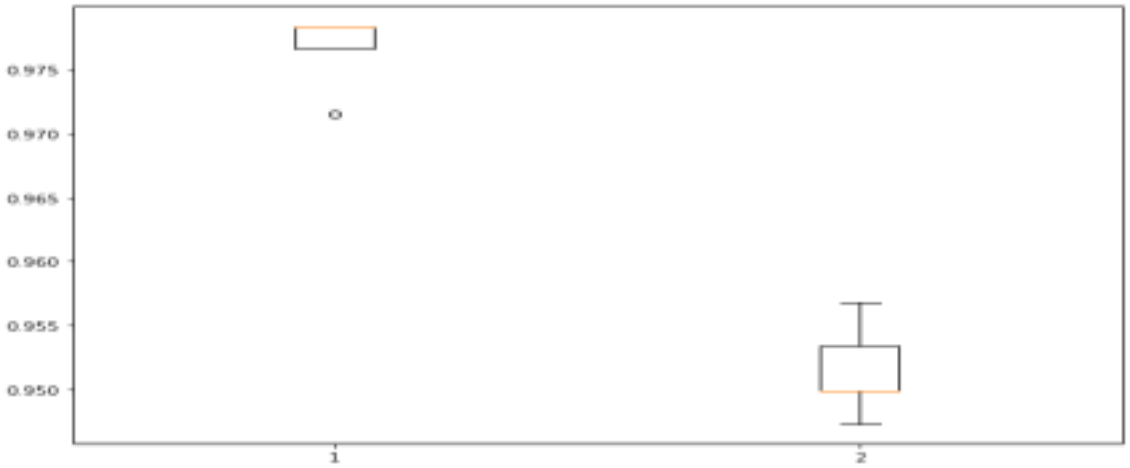
METHODS

We have used 3 different models:

1) **Decision Tree** is a flowchart-like tree structure where an internal node represents a feature(or attribute), the branch represents a decision rule, and each leaf node represents the outcome. The topmost node in a decision tree is known as the root node. It learns to partition on the basis of the attribute value. It partitions the tree in a recursive manner called recursive partitioning. This flowchart-like structure helps you in decision making. Decision Tree is a white box type of ML algorithm. It shares internal decision-making logic, which is not available in the black box type of algorithms such as Neural Network.

Acurracy : `0.9694406077348067`
Accuracy with Reduced Features : `0.9471685082872928`

Boxplot of cross validation score:



2) **Multilayer Perceptron** (MLP) is a class of feedforward artificial neural network (ANN). The term MLP is used ambiguously, sometimes loosely to any feedforward ANN, sometimes strictly to refer to networks composed of multiple layers of perceptrons.
It is composed of more than one perceptron. They are composed of an input layer to receive the signal, an output layer that makes a decision or prediction about the input, and in between those two, an arbitrary number of hidden layers that are the true computational engine of the MLP. MLPs with one hidden layer are capable of approximating any continuous function.

Acurracy : 0.9884323204419889
Accuracy with Reduced Features : 0.9457872928176796

Boxplot of cross validation score:

3) **Logistic regression** is a predictive analysis. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.Logistic regression is used to obtain odds ratio in the presence of more than one explanatory variable. The procedure is quite similar to multiple linear regression, with the exception that the response variable is binomial. The result is the impact of each variable on the odds ratio of the observed event of interest.

Acurracy : 0.9875690607734806
Accuracy with Reduced Features : 0.9720303867403315

Boxplot of cross validation score:

CONCLUSION

- On comparing our results we get that the best model is MLP Model according to the accuracy `0.9884323204419889`
- The best model Logistic Regression Model according to the reduced feature accuracy `0.9720303867403315` .

This face mask detector can be deployed in many areas like shopping malls, airports and other heavy traffic places to monitor the public and to avoid the spread of the disease by checking who is following basic rules and who is not.

REFRENCES

- https://www.researchgate.net/publication/342757061_Efficient_Masked_Face_Recognition_Method_during_the_COVID-19_Pandemic
- https://www.leewayhertz.com/face-mask-detection-system/
- https://www.pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning/