

# README

## CS3244 Homework 3

Thenaesh Elango, Lee Yi Min, Lim Jia Yee

October 30, 2016

### Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Description</b>	<b>2</b>
2.1	Data Sets . . . . .	2
2.2	Learning Models Used . . . . .	2
2.2.1	Initial Approach: Support Vector Machine (SVM) only . . . . .	2
2.3	Second Approach: Recursive Feature Elimination with Cross-Validation (RFECV) . . . . .	4
2.4	Final Approach: Principal Component Analysis (PCA), SVM, and Neural Networks (NN) . . . . .	5
2.5	Other Approaches: Bagging, Decision Trees, and Adaptive Boosting (Adaboost) . . . . .	6
<b>3</b>	<b>File Listing</b>	<b>6</b>
<b>4</b>	<b>Statement of Teams' Independent Work</b>	<b>7</b>

# 1 Introduction

We can recognise faces after seeing them for a few times. We have essentially “learnt” how to recognise these faces, especially from the more distinctive features of these faces. They include the shape of the face, the curvature of the eyes, the slope of the nose, the thickness of the lips, and so on. We aim to extend this “learning” to computer programs.

## 2 Description

In the field of machine learning, this problem of face recognition, given previously labelled faces, is a classification problem. These previously labelled data will be referred to as the training data.

The learner, essentially a computer program, attempts to create a general formula based on the training data to “recognise” other face images not in the training data.

### 2.1 Data Sets

There are 966 training samples and 322 test samples. Each sample has 1850 features, where each feature is simply a pixel. The faces belong to a total of 7 people, and thus there are 7 classes.

### 2.2 Learning Models Used

#### 2.2.1 Initial Approach: Support Vector Machine (SVM) only

**Model** The initial models used were SVMs. All 966 training samples and their 1850 features were passed into the SVM. This means there was no feature dimensionality reduction and no cross-validation.

There were a total of more than 80 SVM models with various combinations of arguments, namely kernels, costs, gammas (applicable for radial basis function (RBF) and polynomial kernel only), degrees (applicable for polynomial kernel only).

The kernels used were the linear kernel, the polynomial kernel of degrees 2 to 4 inclusive, and the RBF kernel.

The following ranges are of step size power of 10: The range of costs used was 0.01 to 100 inclusive. The range of gammas used was 0.0001 to 1 inclusive.

**Classification** For the SVM models using the RBF kernel, all of the test samples were classified under the same class, which seemed to be suspicious. On the other hand, for the SVM models using the linear or the polynomial kernels, the

test samples had a diverse classification, and the SVM models generally agree on the classification for each of the test samples for most of the combinations of the arguments.

A possible explanation for the uniform classification under the RBF kernel could be that RBF is intolerant to slack [1] and thus could not separate the data points which were likely to be close together due to the sheer number of features, many of which could be similar as the human face does not deviate drastically.

As for the linear and polynomial kernels, the classifications were agreeable amongst these models due to:

1. Rather low degrees compared to the number of features
2. Arguments for costs, gamma (if applicable) are within a small range

For any disagreements, the mode of the classification was chosen.

The simple SVM models were decent as the first step, but could have been improved in the following:

1. Ensure the images were oriented in the same angle
2. Gather the more significant features (or pixels) for classification to reduce noise, time, and space complexity
3. Need to try out other models and also think about whether the possible natures of the data set could affect the learning of the different models

As such, in the next batch of models, a new approach with the improvements above in mind will be explored.

### **2.3 Second Approach: Recursive Feature Elimination with Cross-Validation (RFECV)**

## **2.4 Final Approach: Principal Component Analysis (PCA), SVM, and Neural Networks (NN)**

## **2.5 Other Approaches: Bagging, Decision Trees, and Adaptive Boosting (Adaboost)**

## **3 File Listing**

## **4 Statement of Teams' Independent Work**

## References

- [1] Martin C., *What is an RBF Kernel?*, 2012, Feb 6.