Question 5

22b0913 22b0949 22b1030
Umesh Gnana Mahesh Gupta

Contents

1 Introduction 2

2 Precision 2

3 Recall 3

4 F1 Score 4

5 Accuracy Score 5

6 Matthews Correlation Coefficient 5

7 Conclusions 6

8 Instructions to run 6

1 Introduction

To find where there was no corresponding identity, we used a threshold technique. For each image, we computed minimum squared difference between its eigen coefficients and those of all training images, if this difference > threshold then we classified it has no identity.

```
        for j = 1:size(X_test, 2)
            % For all images in the X_test
            error = sum((eigen_coef - test_coef(:, j)).^2); % Calculate
                square error
            [m, index] = min(error); % Find minimum value and index
            if m <= threshold % Then it is correct prediction
                if Y_test(j) == 0 % If prediction is from non trained part
                     False Postive
                    temp_FP  = temp_FP  + 1;
                else % True Positive
                    temp_TP = temp_TP + 1;
                    if Y_train(index) == Y_test(j)
                        rc = rc + 1;
                    end
                end
            else % Then our prediction is wrong
                if Y_test(j) == 0 % if prediction is wrong from non
                    trained part then True Negative
                    temp_TN = temp_TN + 1;
                else % False negative
                    temp_FN = temp_FN + 1;
                end
            end
```

In this question we used $k = 75$ since from the question 4 we can clearly say that there was a significant change in the images from $k = 75$. That is it has very good recognition rate. Hence we used that k value in this question. Now coming to finding the best threshold part, we are containing a range 100 evenly spaced values between 70 and 300.

```
    threshold_values = linspace(70, 300, 100);
```

Now I will fix a measuring metrics, the measuring metrics we used in this question are `f1 score`, `accuracy score`, `matthews correlation coefficient`, `precision` and `recall`. For a particular metrics we are maintaining a variable, `best_score` and considering the best `threshold` as of then

- `False Positive (FP)` - Incorrectly classifying a negative instance as positive.
- `True Positive (TP)` - Correctly classifying a positive instance as positive
- `True Negative (TN)` - Correctly classifying a negative instance as negative.
- `False Negative (FN)` - Incorrectly classifying a positive instance as negative.

2 Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

```
With measuring metrics prec
Accuracy: 0.675000
F1 Score: 0.745098
MCC: 0.475595
Best Threshold: 70.000000
Recall: 0.593750
Confusion matrix:
TP: 76   FP: 0
FN: 52   TN: 32
Recognition rate: 0.475000
```

precision

Precision measures the accuracy of positive predictions by a model. It is the ratio of true positives to the total predicted positives (true positives + false positives), indicating how many of the predicted positives are actually correct.

- `threshold` - 70.00
- `recognition rate` - 0.475

3 Recall

$$\text{Recall} = \frac{TP}{TP + FN}$$

```
With measuring metrics recc
Accuracy: 0.812500
F1 Score: 0.895105
MCC: 0.225018
Best Threshold: 172.222222
Recall: 1.000000
Confusion matrix:
TP: 128 FP: 30
FN: 0   TN: 2
Recognition rate: 0.737500
```

recall

Recall (or sensitivity) measures the ability of a model to correctly identify all positive instances. It is the ratio of true positives to the total actual positives (true positives + false negatives).

We can see the results when recall is considered as a metric

- `threshold` = 172.22
- `recognition rate` = 0.7375

4 F1 Score

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \cdot \frac{TP}{2TP + FP + FN}$$

```
With measuring metrics f1
Accuracy: 0.812500
F1 Score: 0.895105
MCC: 0.225018
Best Threshold: 172.222222
Recall: 1.000000
Confusion matrix:
TP: 128 FP: 30
FN: 0   TN: 2
Recognition rate: 0.737500
```

f1

The F1 score is the harmonic mean of precision and recall, providing a balanced measure between the two. It is useful when there is an uneven class distribution, as it focuses on the balance between false positives and false negatives.

Taking f1 score as metrics we will try to maximize the f1 score by iterating through all possible thresholds.

- `Threshold` = 172.222
- `f1` = 0.895
- `recognition rate` = 0.737

5 Accuracy Score

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

```
With measuring metrics acc
Accuracy: 0.812500
F1 Score: 0.895105
MCC: 0.225018
Best Threshold: 172.222222
Recall: 1.000000
Confusion matrix:
TP: 128 FP: 30
FN: 0   TN: 2
Recognition rate: 0.737500
```

accuracy score

Accuracy score measures the proportion of correctly classified instances (both true positives and true negatives) out of the total instances. It provides an overall assessment of the model's performance but can be misleading when classes are imbalanced.

- `recognition rate` - 0.737
- `threshold` = 172.22

6 Matthews Correlation Coefficient

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

```
With measuring metrics mcc
Accuracy: 0.756250
F1 Score: 0.826667
MCC: 0.492498
Best Threshold: 86.262626
Recall: 0.726562
Confusion matrix:
TP: 93  FP: 4
FN: 35  TN: 28
Recognition rate: 0.581250
```

Matthews Correlation Coefficient

The Matthews Correlation Coefficient (MCC) is a comprehensive measure of the quality of binary classifications, taking into account all four categories: true positives, true negatives, false positives, and false negatives. It yields a value between -1 and +1, where +1 indicates perfect predictions, 0 represents random predictions, and -1 denotes total disagreement between predicted and actual classes. MCC is particularly useful in evaluating models on imbalanced datasets, as it provides a balanced view of performance across all classes.

- `threshold` - 86.26
- `recognition rate` - 0.568

7 Conclusions

Through our rigorous testing, we found 128 True Positives, 30 False Positives and 2 true negatives in many cases i.e, when we used Accuracy score, F1 score, Recall as metrics. There are cases of few metrics like MCC and precision which shows lower (0.475) recognition rates than previous metrics (0.737). Hence, depending on the application we can choose any of the thresholds mentioned above. However, for general applications we would like to have a low number of False Positives and False Negatives. Therefore we got,

- `threshold` = 172.222
- **TP**: 128
- **FP**: 30
- **FN**: 0
- **TN**: 2

So, when we test our system on images of people which were not part of the training set we will face situations where unknown faces may be misclassified, Poor accuracy on unseen data and Occurrence of False positives (correctly identifying an unknown person as someone from the training set).

8 Instructions to run

`q5.m` file contains the code for finding the best threshold, Make sure that `ORL/` directory is in the same directory as `q5.m`