Ensemble modelling

August 15, 2023

1 Dataset

https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000

Cases include a representative collection of all important diagnostic categories in the realm of pigmented lesions: - Actinic keratoses and intraepithelial carcinoma / Bowen's disease (AKIEC), - basal cell carcinoma (BCC) - benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, BKL) - dermatofibroma (DF) - melanoma (MEL) - melanocytic nevi (NV) - vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, VASC).

The classes and the target labels have meen listed below

Category	Class ID
melanoma (MEL)	0
melanocytic nevi (NV)	1
basal cell carcinoma (BCC)	2
Actinic keratoses and intraepithelial carcinoma / Bowen's disease (AKIEC)	3
benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, BKL)	4
dermatofibroma (DF)	5
vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, VASC)	6

2 Classification features

The LBP calculated from the images are being used as features for image classification.

3 Untransformed

Generating classification report and consufion matrix of images that were not log transformed before segmentation
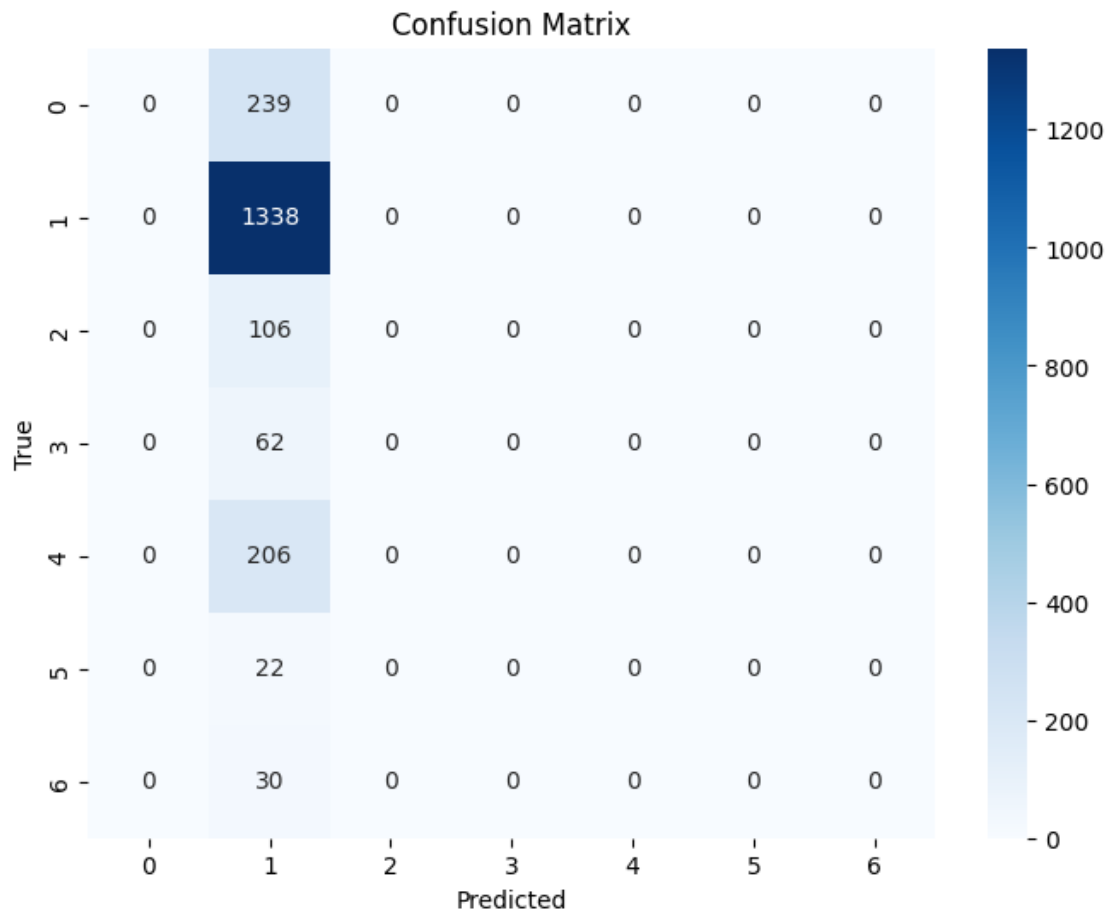
3.1 SVM classification

Train test split is 80% and 20% respectively.

3.1.1 10-fold cross validation

```
Fold 1: Accuracy = 0.67
Fold 2: Accuracy = 0.67
Fold 3: Accuracy = 0.68
Fold 4: Accuracy = 0.67
Fold 5: Accuracy = 0.67
Fold 6: Accuracy = 0.68
Fold 7: Accuracy = 0.68
Fold 8: Accuracy = 0.66
Fold 9: Accuracy = 0.66
Fold 10: Accuracy = 0.67
Mean Accuracy: 0.67
```

[8]: SVC()

3.1.2 Confusion Matrix

3.1.3 Classification Report

```
Sensitivity: 1.00
Specificity: 0.00
          precision    recall  f1-score   support

         0       0.00      0.00      0.00       239
         1       0.67      1.00      0.80      1338
         2       0.00      0.00      0.00       106
         3       0.00      0.00      0.00        62
         4       0.00      0.00      0.00       206
         5       0.00      0.00      0.00        22
         6       0.00      0.00      0.00        30

  accuracy                           0.67      2003
 macro avg       0.10      0.14      0.11      2003
weighted avg     0.45      0.67      0.54      2003
```

```
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```

3.2 Random forrest classifier
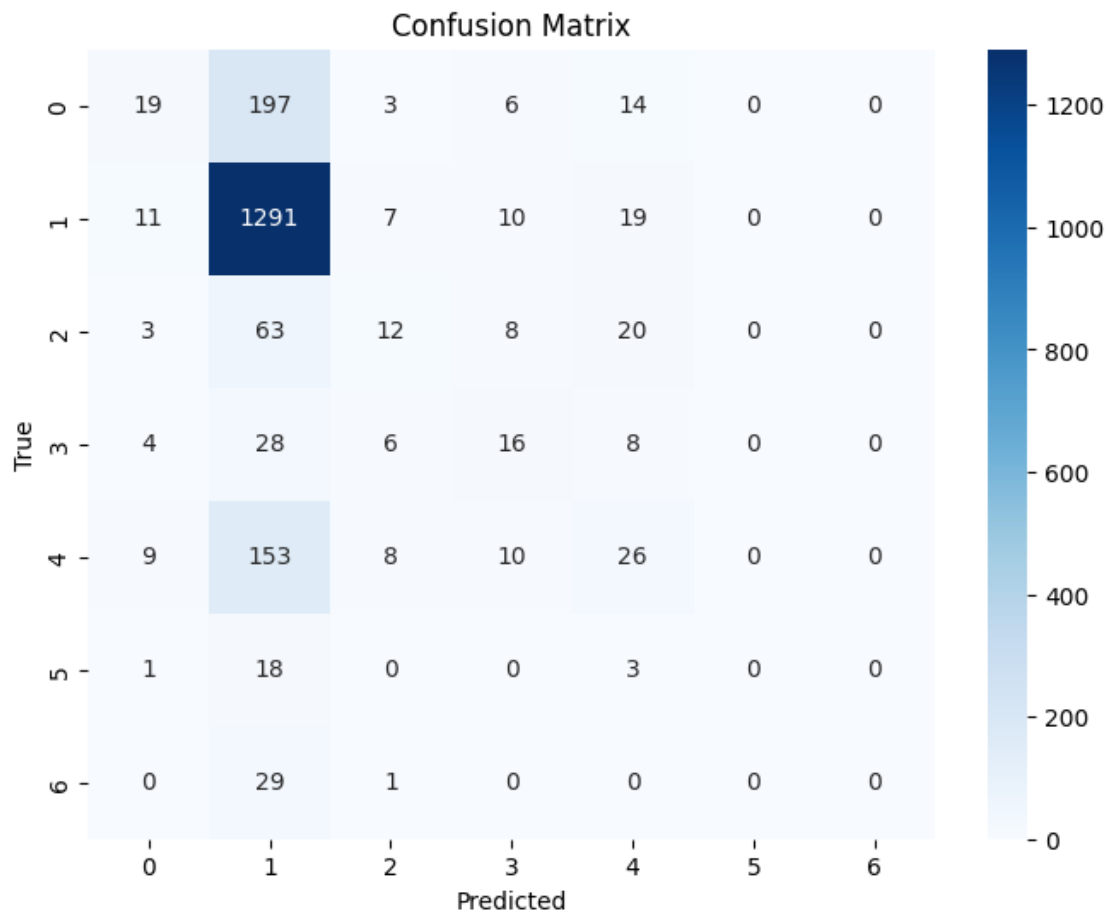
Train test split is 80% and 20% respectively.

3.2.1 10-fold cross validation

```
Fold 1: Accuracy = 0.68
Fold 2: Accuracy = 0.68
Fold 3: Accuracy = 0.69
Fold 4: Accuracy = 0.68
Fold 5: Accuracy = 0.69
Fold 6: Accuracy = 0.69
Fold 7: Accuracy = 0.70
Fold 8: Accuracy = 0.67
```

```
Fold 9: Accuracy = 0.65
Fold 10: Accuracy = 0.67
Mean Accuracy: 0.68
```

[11]: RandomForestClassifier()

3.2.2 Confusion matrix

Confusion Matrix

3.2.3 Classification Report

```
Sensitivity: 0.99
Specificity: 0.09
              precision    recall  f1-score   support

           0       0.40      0.08      0.13       239
           1       0.73      0.96      0.83      1338
           2       0.32      0.11      0.17       106
           3       0.32      0.26      0.29        62
           4       0.29      0.13      0.18       206
```

5	0.00	0.00	0.00	22
6	0.00	0.00	0.00	30
accuracy			0.68	2003
macro avg	0.29	0.22	0.23	2003
weighted avg	0.59	0.68	0.60	2003

```
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```

3.3   Adaboost

Train test split is 80% and 20% respectively.
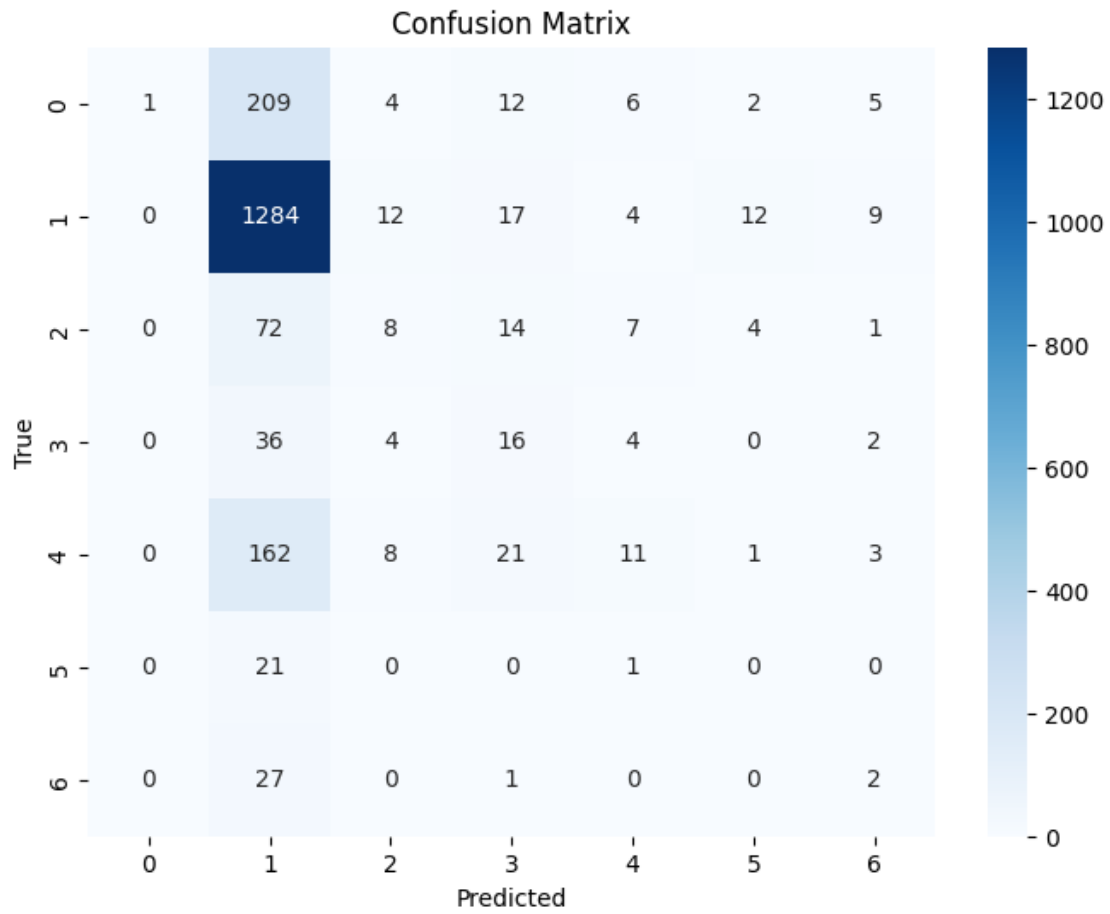
3.3.1   10-fold cross validation

```
Fold 1: Accuracy = 0.60
Fold 2: Accuracy = 0.59
Fold 3: Accuracy = 0.68
Fold 4: Accuracy = 0.62
Fold 5: Accuracy = 0.68
Fold 6: Accuracy = 0.67
Fold 7: Accuracy = 0.55
Fold 8: Accuracy = 0.62
Fold 9: Accuracy = 0.64
Fold 10: Accuracy = 0.64
Mean Accuracy: 0.63
```

[14]: AdaBoostClassifier()

3.3.2 Confusion Matrix

Confusion Matrix

3.3.3 Classification Report

```
Sensitivity: 1.00
Specificity: 0.00
              precision     recall   f1-score     support

           0       1.00       0.00       0.01         239
           1       0.71       0.96       0.82        1338
           2       0.22       0.08       0.11         106
           3       0.20       0.26       0.22          62
           4       0.33       0.05       0.09         206
           5       0.00       0.00       0.00          22
           6       0.09       0.07       0.08          30

    accuracy                             0.66        2003
   macro avg       0.36       0.20       0.19        2003
weighted avg       0.65       0.66       0.57        2003
```

4 Transformed

Generating classification report and consufion matrix of images that were log transformed before segmentation
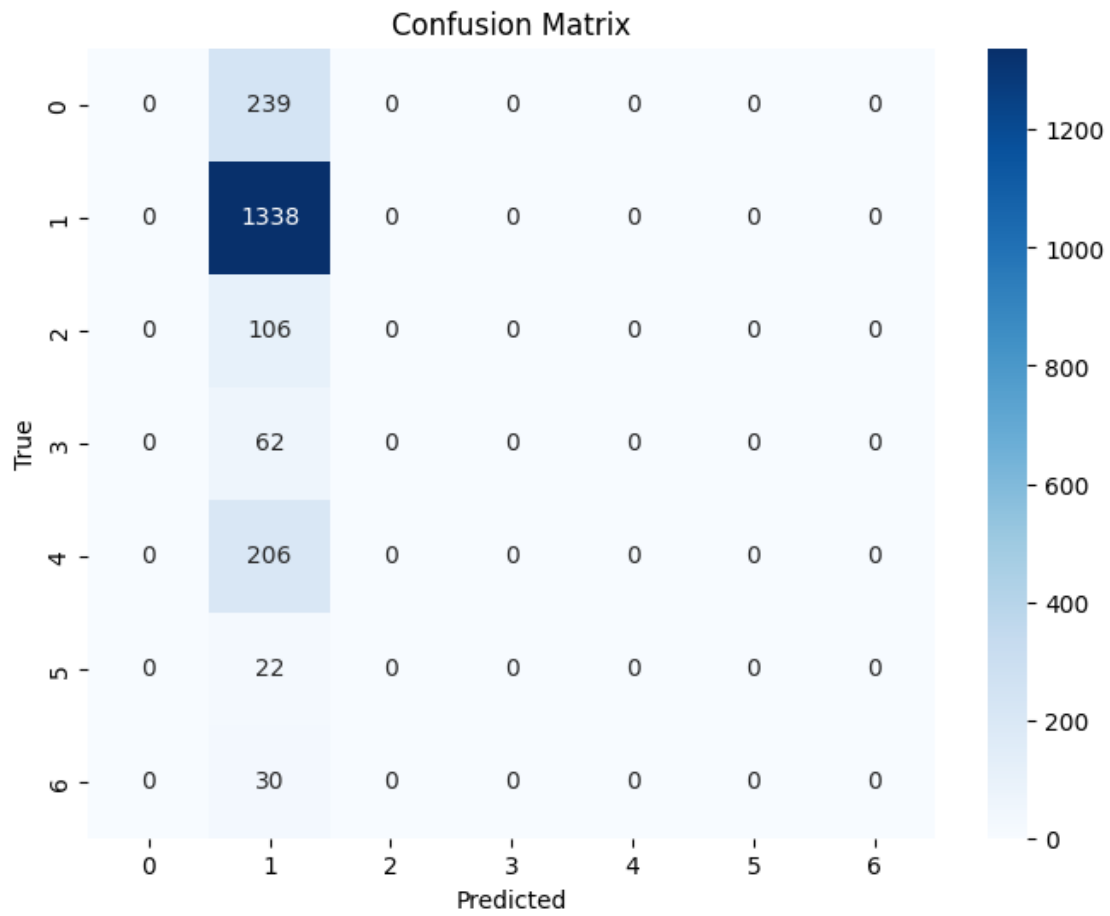
4.1 SVM classification

Train test split is 80% and 20% respectively.

4.1.1 10-fold cross validation

```
Fold 1: Accuracy = 0.67
Fold 2: Accuracy = 0.67
Fold 3: Accuracy = 0.68
Fold 4: Accuracy = 0.67
Fold 5: Accuracy = 0.67
Fold 6: Accuracy = 0.68
Fold 7: Accuracy = 0.68
Fold 8: Accuracy = 0.66
Fold 9: Accuracy = 0.66
Fold 10: Accuracy = 0.67
Mean Accuracy: 0.67
```

[23]: SVC()

4.1.2 confusion Matrix

Confusion Matrix

4.1.3 Classification Report

```
Sensitivity: 1.00
Specificity: 0.00
              precision    recall   f1-score    support

           0       0.00      0.00       0.00        239
           1       0.67      1.00       0.80       1338
           2       0.00      0.00       0.00        106
           3       0.00      0.00       0.00         62
           4       0.00      0.00       0.00        206
           5       0.00      0.00       0.00         22
           6       0.00      0.00       0.00         30

    accuracy                           0.67       2003
   macro avg       0.10      0.14       0.11       2003
weighted avg       0.45      0.67       0.54       2003
```

```
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```

4.2 Random forrest classifier

4.2.1 10-fold cross validation

```
Fold 1: Accuracy = 0.68
Fold 2: Accuracy = 0.69
Fold 3: Accuracy = 0.70
Fold 4: Accuracy = 0.68
Fold 5: Accuracy = 0.69
Fold 6: Accuracy = 0.68
Fold 7: Accuracy = 0.69
Fold 8: Accuracy = 0.68
Fold 9: Accuracy = 0.68
Fold 10: Accuracy = 0.67
Mean Accuracy: 0.68
```

4.2.2 Confusion Matrix

Confusion Matrix
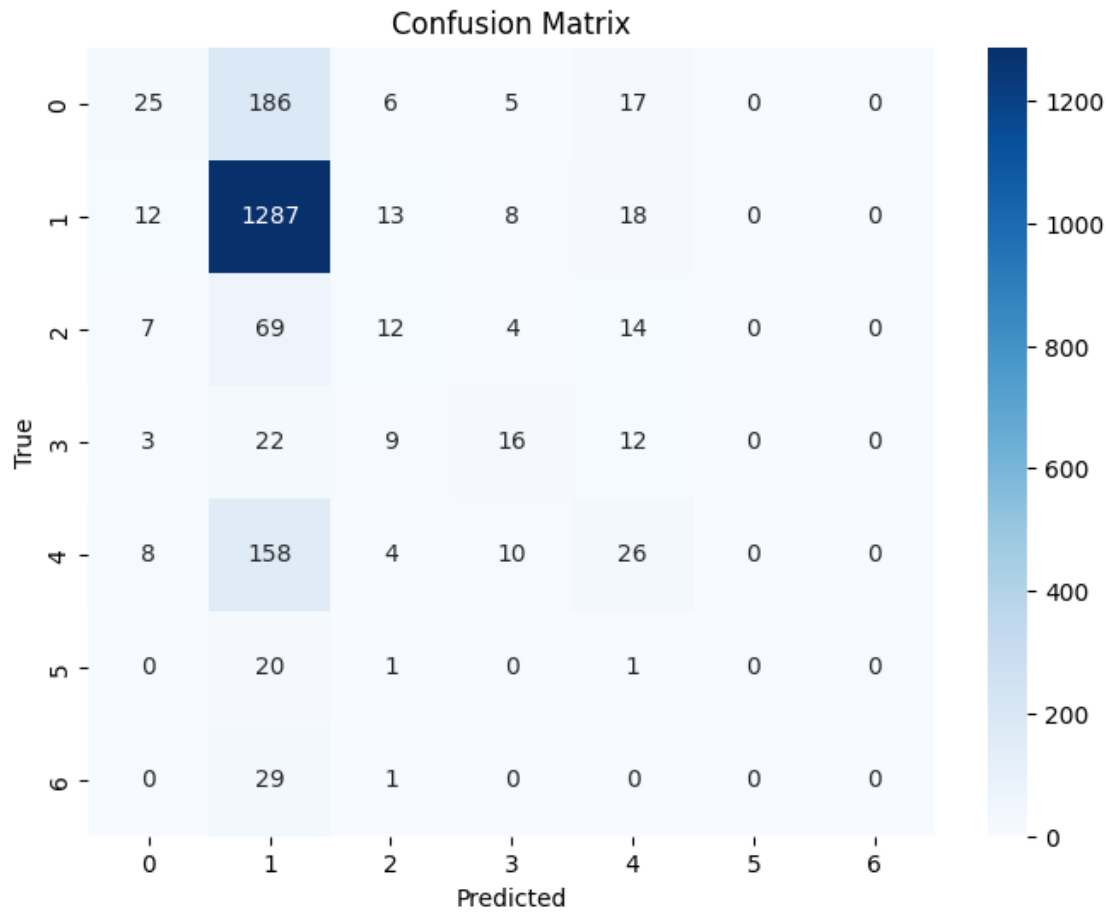
4.2.3 Classification Report

```
Sensitivity: 0.99
Specificity: 0.12
              precision    recall  f1-score   support

           0       0.45      0.10      0.17       239
           1       0.73      0.96      0.83      1338
           2       0.26      0.11      0.16       106
           3       0.37      0.26      0.30        62
           4       0.30      0.13      0.18       206
           5       0.00      0.00      0.00        22
           6       0.00      0.00      0.00        30

    accuracy                           0.68      2003
   macro avg       0.30      0.22      0.23      2003
weighted avg       0.60      0.68      0.61      2003
```

```
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```

4.3 Adaboost

Train test split is 80% and 20% respectively.
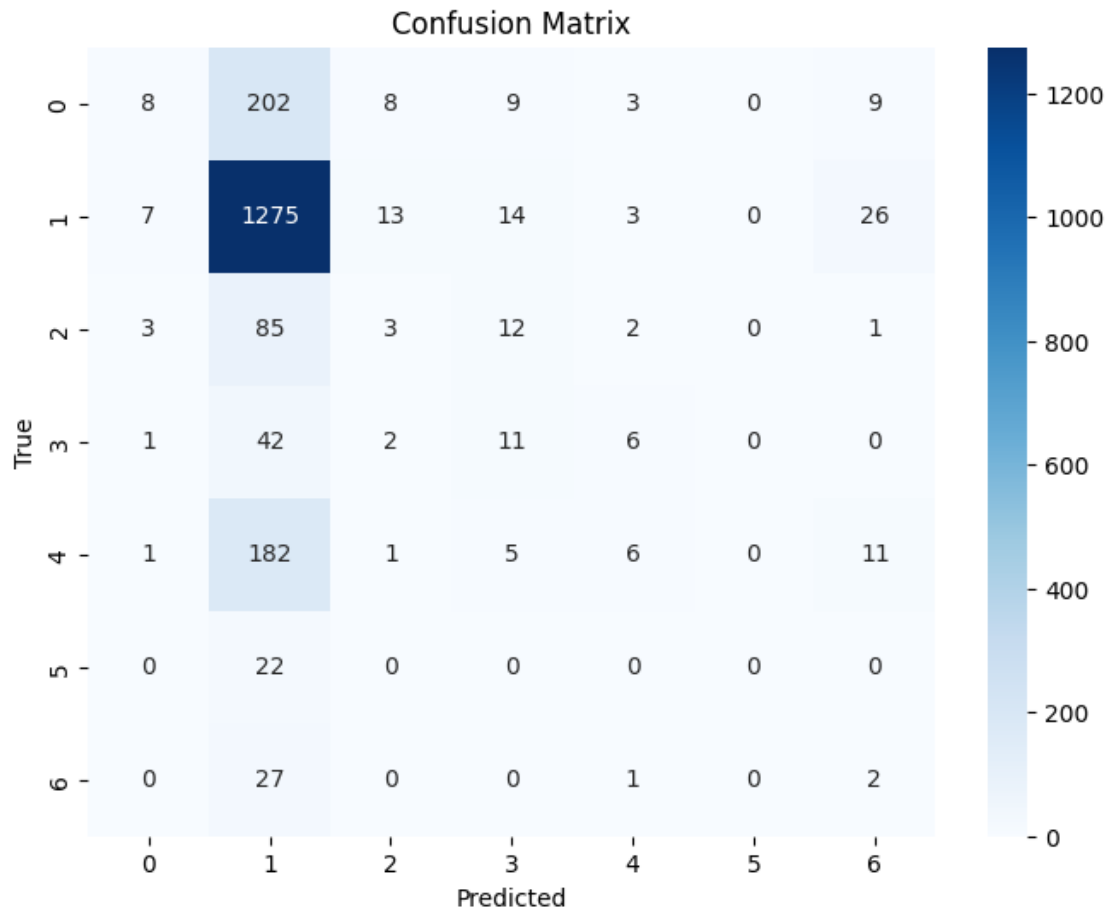
4.3.1 10-fold cross validation

```
Fold 1: Accuracy = 0.65
Fold 2: Accuracy = 0.66
Fold 3: Accuracy = 0.66
Fold 4: Accuracy = 0.67
Fold 5: Accuracy = 0.65
Fold 6: Accuracy = 0.66
Fold 7: Accuracy = 0.64
Fold 8: Accuracy = 0.64
Fold 9: Accuracy = 0.65
Fold 10: Accuracy = 0.63
Mean Accuracy: 0.65
```

```
[29]: AdaBoostClassifier()
```

4.3.2 Confusion Matrix

Confusion Matrix

4.3.3 Classification report

```
Sensitivity: 0.99
Specificity: 0.04
              precision    recall  f1-score   support

           0       0.40      0.03      0.06       239
           1       0.69      0.95      0.80      1338
           2       0.11      0.03      0.05       106
           3       0.22      0.18      0.19        62
           4       0.29      0.03      0.05       206
           5       0.00      0.00      0.00        22
           6       0.04      0.07      0.05        30

    accuracy                           0.65      2003
   macro avg       0.25      0.18      0.17      2003
weighted avg       0.55      0.65      0.56      2003
```

```
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/rajdeep/projects/en/lib/python3.10/site-
packages/sklearn/metrics/_classification.py:1344: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```