CSET228(P)-Lab10-E23CSEU0055

By - Arihant Gupta, E23CSEU0055, EB02

Question 1: @

Q1. Implementing Support Vector Machine (SVM) for Classification. You will train an SVM model on the **Wine Quality dataset**, which is available on the **UCI Machine Learning Repository**.

Dataset Link: https://archive.ics.uci.edu/ml/datasets/wine+quality

Follow the steps below:

- · Download and Load the Dataset
- · Perform Data Pre-processing to check missing values and normalize the datset
- · Split the dataset in 80:20 ratio
- Train SVM model
- Evaluate the performance of the model
- Compute accuracy, precision, recall, and F1-score.
- Experiment with different kernels (RBF, polynomial) and compare results.

Code: ℰ

```
1 import pandas as pd
2 from sklearn.model selection import train test split
3 from sklearn.preprocessing import StandardScaler
4 from sklearn.svm import SVC
5 from sklearn.metrics import accuracy_score, precision_score, recall_score, fl_score, classification_report
6
7 df = pd.read_csv("white.csv", sep=';')
8
9 X = df.drop("quality", axis=1)
10 y = df["quality"]
11 scaler = StandardScaler()
12 X scaled = scaler.fit transform(X)
13
14 X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42, stratify=y)
15
16 def evaluate_svm(kernel_type):
17
       model = SVC(kernel=kernel_type, random_state=42)
18
       model.fit(X train, y train)
19
       y_pred = model.predict(X_test)
20
21
       print(f"SVM with {kernel type} kernel")
22
       print("Accuracy:", accuracy_score(y_test, y_pred))
23
       print("Precision:", precision_score(y_test, y_pred, average='weighted', zero_division=0))
24
       print("Recall:", recall_score(y_test, y_pred, average='weighted'))
25
       print("F1 Score:", f1_score(y_test, y_pred, average='weighted'))
       print("\nClassification Report:\n", classification_report(y_test, y_pred))
26
27
28 evaluate_svm('rbf')
29
30 evaluate_svm('poly')
31
```

Output: \mathscr{O}

1	SVM with rbf k	ernel			
2	Accuracy: 0.628125				
3	Precision: 0.6117222764269392				
4	Recall: 0.628125				
5	F1 Score: 0.6053464705720344				
6					
7	Classification Report:				
8	precision recall f1-score support				
9					
10	3	0.00	0.00	0.00	2
11	4	0.00	0.00	0.00	11
12	5	0.65	0.76	0.70	136
13	6	0.58	0.66	0.62	128
14	7	0.82	0.35	0.49	40
15	8	0.00	0.00	0.00	3
16					
17				0.63	320
18	•	0.34	0.29	0.30	320
19		0.61	0.63	0.61	320
20					
	SVM with poly	kernel			
	Accuracy: 0.60625				
23	•				
24					
	5 F1 Score: 0.5846716295122099				
26					
27					
28		precision	recal1	f1-score	support
29					
30		0.00	0.00	0.00	2
31		0.00	0.00	0.00	11
32		0.63	0.79	0.70	136
33		0.57	0.73	0.57	128
34		0.76	0.33	0.46	40
35		0.00	0.00	0.00	3
36		0.00	0.00	0.00	5
37				0.61	320
38	•	0.33	0.28	0.01	320
39	3	0.59	0.28	0.58	320
29	weighted avg	0.39	0.01	0.30	320

Question 2: @

Q2. Implementing ANN & MLP for breast cancer classification. The dataset contains features extracted from **digitized images of breast cancer cell nuclei**. The goal is to classify tumors as **malignant (0) or benign (1)**.

Dataset link:

https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic

School of Computer Science Engineering and Technology

Follow the steps below to implement **both ANN (using TensorFlow/Keras) and MLP (using Scikit-learn)** to classify breast cancer tumors.

Step 1: Load and Explore the Dataset

- i. Import the dataset using load_breast_cancer from sklearn.datasets.
- ii. Convert the dataset into a Pandas DataFrame.
- Display the first five rows using df.head().
- iv. Print dataset statistics using df.describe().

Step 2: Data Pre-processing

- i. Check for missing values using df.isnull().sum().
- ii. Standardize features using StandardScaler from sklearn.preprocessing.
- iii. Convert target labels (Malignant = 0, Benign = 1) to categorical format (only for ANN).

Step 3: Splitting the Data

Use train_test_split: 80% training data, 20% testing data

Step 4: Build the ANN Model

- i. Import Sequential from tensorflow.keras.models.
- ii. Add two hidden layers using Dense with ReLU activation.
- iii. Use sigmoid activation for the output layer (since it's a binary classification problem).

Step 5: Compile and Train the ANN Model

- i. Use binary_crossentropy as the loss function
- ii. Use adam as the optimizer
- iii. Train the model using model.fit()

Epochs: 50 Batch size: 32

Step 6: Build the MLP Model

- i. Import MLPClassifier from sklearn.neural network.
- ii. Set the following parameters:

Activation Function: ReLU

Optimizer: Adam

Step 7: Train the MLP Model: Train the model using fit().

Step 8: Performance Evaluation

For ANN (Keras Model):

Evaluate the model on test data using model.evaluate().

Compute precision, recall, and F1-score using classification_report()

Plot loss and accuracy curves using matplotlib.

For MLP (Scikit-learn Model):

Predict labels for the test data.

Compute accuracy, precision, recall, and F1-score using classification_report.

Compare the ANN and MLP results.

Code: ℰ

```
from sklearn.datasets import load_breast_cancer
import pandas as pd
import numpy as np
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.utils import to_categorical

data = load_breast_cancer()
```

```
10 df = pd.DataFrame(data.data, columns=data.feature_names)
11 df['target'] = data.target
12
13 print(df.head())
14
15 print(df.describe())
16
17 print("Missing values:\n", df.isnull().sum())
18
19 # ii. Standardize features
20 from sklearn.preprocessing import StandardScaler
21 scaler = StandardScaler()
22 X scaled = scaler.fit transform(df.drop('target', axis=1))
23
24
25 y_ann = to_categorical(df['target'])
y = df['target']
27
28 from sklearn.model_selection import train_test_split
29
30 X train ann, X test ann, y train ann, y test ann = train test split(
31
       X_scaled, y_ann, test_size=0.2, random_state=42, stratify=y)
32
33 X_train_mlp, X_test_mlp, y_train_mlp, y_test_mlp = train_test_split(
34
       X_scaled, y, test_size=0.2, random_state=42, stratify=y)
35
37 ann_model = Sequential()
38 ann_model.add(Dense(16, activation='relu', input_shape=(X_scaled.shape[1],)))
39 ann_model.add(Dense(8, activation='relu'))
40 ann_model.add(Dense(2, activation='sigmoid'))
41
42 ann_model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
43
44 history = ann model.fit(
45
       X_train_ann, y_train_ann, epochs=50, batch_size=32,
46
        validation_data=(X_test_ann, y_test_ann), verbose=0)
47
48 from sklearn.neural_network import MLPClassifier
50 mlp = MLPClassifier(hidden_layer_sizes=(16, 8), activation='relu', solver='adam', max_iter=500,
    random state=42)
51
52 mlp.fit(X_train_mlp, y_train_mlp)
53
54 from sklearn.metrics import classification_report, accuracy_score
55 import matplotlib.pyplot as plt
57 loss, acc = ann_model.evaluate(X_test_ann, y_test_ann, verbose=0)
58 print(f"\n ✓ ANN Test Accuracy: {acc:.4f}, Loss: {loss:.4f}")
60 y_pred_ann = ann_model.predict(X_test_ann)
61 y_pred_ann_labels = np.argmax(y_pred_ann, axis=1)
62 y_true_ann_labels = np.argmax(y_test_ann, axis=1)
63
64 print("\nClassification Report (ANN):\n", classification_report(y_true_ann_labels, y_pred_ann_labels))
65
```

```
66 plt.figure(figsize=(12,5))
67 plt.subplot(1,2,1)
68 plt.plot(history.history['loss'], label='Train Loss')
69 plt.plot(history.history['val loss'], label='Val Loss')
70 plt.title('Loss Curve')
71 plt.legend()
72
73 plt.subplot(1,2,2)
74 plt.plot(history.history['accuracy'], label='Train Acc')
75 plt.plot(history.history['val_accuracy'], label='Val Acc')
76 plt.title('Accuracy Curve')
77 plt.legend()
78 plt.show()
79
80 y_pred_mlp = mlp.predict(X_test_mlp)
81 print("\n

MLP Test Accuracy:", accuracy_score(y_test_mlp, y_pred_mlp))
82 print("\nClassification Report (MLP):\n", classification_report(y_test_mlp, y_pred_mlp))
83
84 print("\n Final Comparison:")
85 print(f"ANN Accuracy: {acc:.4f}")
86 print(f"MLP Accuracy: {accuracy_score(y_test_mlp, y_pred_mlp):.4f}")
```

Output: @

```
1 mean radius mean texture mean perimeter mean area mean smoothness \
2 0
        17.99 10.38 122.80 1001.0
                                                    0.11840
                                132.90 1326.0
3 1
          20.57
                     17.77
                                                       0.08474
4 2
         19.69
                    21.25
                                130.00 1203.0
                                                      0.10960
5 3
         11.42
                    20.38
                                 77.58
                                          386.1
                                                       0.14250
6 4
          20.29
                     14.34
                                 135.10 1297.0
                                                        0.10030
7
8
     mean compactness mean concavity mean concave points mean symmetry \
                        0.3001
9 0
          0.27760
                                          0.14710
                                                       0.2419
10 1
            0.07864
                         0.0869
                                          0.07017
                                                        0.1812
11 2
            0.15990
                        0.1974
                                          0.12790
                                                        0.2069
12 3
            0.28390
                         0.2414
                                           0.10520
                                                        0.2597
13 4
            0.13280
                          0.1980
                                          0.10430
                                                        0.1809
14
mean fractal dimension ... worst texture worst perimeter worst area \
16 0
                0.07871 ... 17.33
                                              184.60
                                                          2019.0
17 1
                0.05667 ...
                                  23.41
                                               158.80
                                                         1956.0
                                  25.53
                                               152.50
18 2
                0.05999 ...
                                                         1709.0
19 3
                 0.09744 ...
                                   26.50
                                                98.87
                                                          567.7
20 4
                 0.05883 ...
                                  16.67
                                               152.20
                                                          1575.0
21
22
     worst smoothness worst compactness worst concavity worst concave points \
         0.1622
23 0
                      0.6656
                                          0.7119
                                                           0.2654
24 1
             0.1238
                           0.1866
                                           0.2416
                                                            0.1860
25 2
             0.1444
                            0.4245
                                           0.4504
                                                            0.2430
             0.2098
                                                            0.2575
26 3
                            0.8663
                                           0.6869
27 4
             0.1374
                            0.2050
                                           0.4000
                                                            0.1625
28
29
   worst symmetry worst fractal dimension target
30 0
           0.4601
                             0.11890
31 1
            0.2750
                               0.08902
                                           0
32 2
           0.3613
                               0.08758
```

```
33 3
               0.6638
                                        0.17300
                                                      0
34 4
               0.2364
                                        0.07678
                                                      0
35
36
   [5 rows x 31 columns]
37
           mean radius mean texture mean perimeter
                                                         mean area \
38
   count
            569.000000
                          569.000000
                                           569.000000
                                                        569.000000
39
    mean
             14.127292
                           19.289649
                                            91.969033
                                                        654.889104
                            4.301036
40
   std
              3.524049
                                            24.298981
                                                        351.914129
                            9.710000
                                            43.790000
                                                        143.500000
41
   min
              6.981000
42
   25%
             11.700000
                           16.170000
                                            75.170000
                                                        420.300000
43
   50%
             13.370000
                           18.840000
                                            86.240000
                                                        551.100000
44
   75%
             15.780000
                           21.800000
                                           104.100000
                                                        782.700000
             28.110000
                           39.280000
                                           188.500000 2501.000000
45
   max
46
47
           mean smoothness mean compactness mean concavity mean concave points \
                569.000000
                                  569.000000
                                                   569.000000
                                                                         569.000000
48
   count
                  0.096360
                                     0.104341
                                                     0.088799
                                                                           0.048919
49
   mean
50
   std
                  0.014064
                                     0.052813
                                                     0.079720
                                                                           0.038803
51
   min
                  0.052630
                                     0.019380
                                                     0.000000
                                                                           0.000000
52
   25%
                  0.086370
                                     0.064920
                                                     0.029560
                                                                           0.020310
53
   50%
                  0.095870
                                     0.092630
                                                     0.061540
                                                                           0.033500
54
   75%
                  0.105300
                                     0.130400
                                                     0.130700
                                                                           0.074000
55
   max
                  0.163400
                                     0.345400
                                                     0.426800
                                                                           0.201200
56
57
           mean symmetry mean fractal dimension ... worst texture \
58
   count
              569.000000
                                      569.000000
                                                           569,000000
                0.181162
                                         0.062798
                                                            25.677223
59
   mean
60
   std
                0.027414
                                         0.007060
                                                             6.146258
61
                0.106000
                                         0.049960
                                                            12.020000
   min
62 25%
                0.161900
                                         0.057700
                                                            21.080000
                                                   . . .
63
   50%
                0.179200
                                         0.061540
                                                            25.410000
64
   75%
                                                            29.720000
                0.195700
                                         0.066120
65
                0.304000
                                         0.097440
                                                            49.540000
   max
66
67
           worst perimeter
                             worst area worst smoothness worst compactness \
                569.000000
                             569.000000
                                                569.000000
                                                                    569.000000
68
   count
                107.261213
69
   mean
                             880.583128
                                                  0.132369
                                                                     0.254265
70
   std
                 33.602542
                             569.356993
                                                  0.022832
                                                                     0.157336
71
    min
                 50.410000
                             185.200000
                                                  0.071170
                                                                      0.027290
                 84.110000
72
   25%
                             515.300000
                                                  0.116600
                                                                     0.147200
73
   50%
                 97.660000
                             686.500000
                                                  0.131300
                                                                      0.211900
74
   75%
                125.400000
                             1084.000000
                                                  0.146000
                                                                      0.339100
75
                251.200000
                            4254.000000
                                                  0.222600
                                                                      1.058000
   max
76
77
           worst concavity
                            worst concave points worst symmetry \
78
   count
                569.000000
                                       569.000000
                                                       569.000000
79
    mean
                  0.272188
                                         0.114606
                                                         0.290076
80
   std
                  0.208624
                                         0.065732
                                                         0.061867
81
   min
                  0.000000
                                         0.000000
                                                         0.156500
82
   25%
                  0.114500
                                         0.064930
                                                         0.250400
83
   50%
                  0.226700
                                         0.099930
                                                         0.282200
                  0.382900
                                                         0.317900
84
    75%
                                         0.161400
85
   max
                  1.252000
                                         0.291000
                                                         0.663800
86
87
           worst fractal dimension
                                         target
88
                        569.000000 569.000000
   count
89
   mean
                          0.083946
                                       0.627417
90
   std
                          0.018061
                                       0.483918
```

```
91 min
                             0.055040
                                           0.000000
                0.071460 0.000000
0.080040 1.000000
0.092080 1.000000
0.207500 1.000000
 92 25%
 93 50%
 94 75%
 95 max
 97 [8 rows x 31 columns]
 98 Missing values:
99 mean radius100 mean texture
                              0
0
101 mean perimeter
                             Θ
102 mean area
mean area 0

103 mean smoothness 0

104 mean compactness 0

105 mean concavity 0

106 mean concave points 0

107 mean symmetry 0
108 mean fractal dimension 0
109 radius error 0
110 texture error
                                0
111 perimeter error
112 area error
                               0
0
                               Θ
Θ
113 smoothness error
114 compactness error
115 concavity error
116 concave points error 0
117 symmetry error 0
118 fractal dimension error 0
119 worst radius 0
0
122 worst area
worst smoothness 0
124 worst compactness 0
125 worst concavity 0
126 worst concave points 0
127 worst symmetry 0
128 worst fractal dimension 0
129 target
                                   0
130 dtype: int64
```

```
1 MLP Test Accuracy: 0.956140350877193
2
3 Classification Report (MLP):
4
                  precision
                              recall f1-score
                                               support
5
6
                     0.91
                               0.98
                                         0.94
                                                    42
             0
7
              1
                     0.99
                               0.94
                                         0.96
                                                    72
8
9
                                         0.96
                                                   114
      accuracy
10
      macro avg
                     0.95
                               0.96
                                         0.95
                                                   114
11 weighted avg
                     0.96
                               0.96
                                         0.96
                                                   114
12
13
14 Final Comparison:
15 ANN Accuracy: 0.9474
16 MLP Accuracy: 0.9561
```

