CENG 499 - Introduction to Machine Learning
Fall 2022
Homework 3

Necat Kılıçarslan

December 2022

1 PART2

1.1 Dataset1

Here is my configuration for part2-dataset1:

- SVC(kernel="rbf", C=1)

- SVC(kernel="rbf", C=100)

- SVC(kernel="linear", C=1)

- SVC(kernel="linear", C=100)

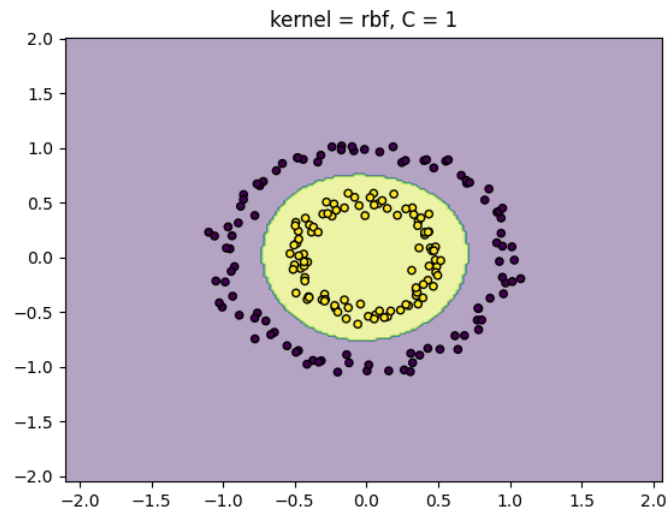Below is the plot result of each configurations:
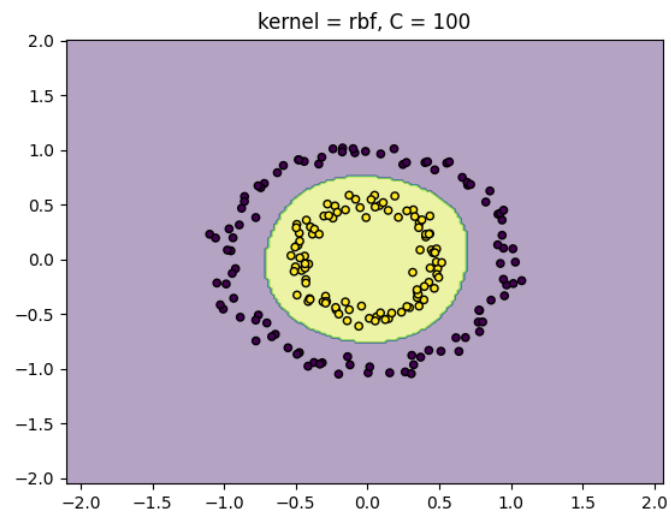

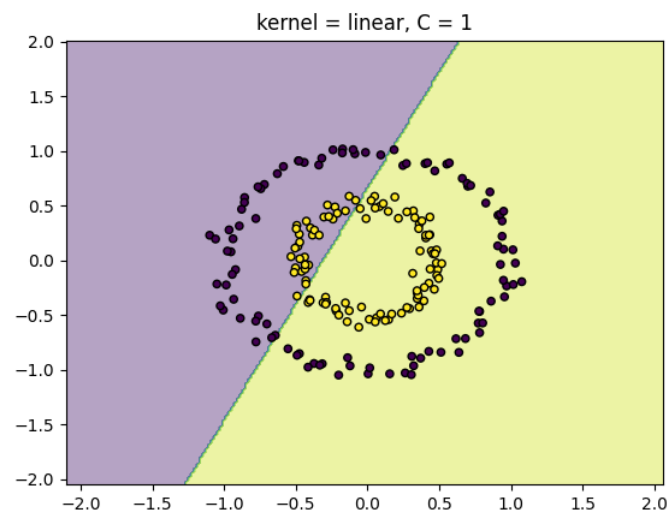Figure 1: Plot result
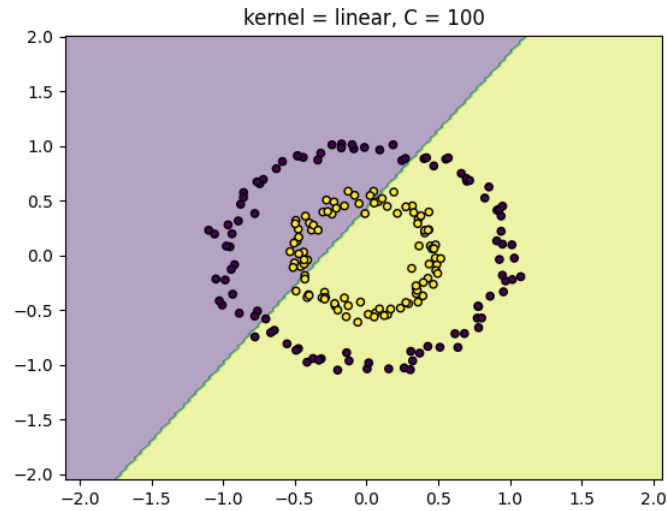
Figure 2: Plot result

Figure 3: Plot result

Figure 4: Plot result

1.2 Dataset2

Here is my configuration and the results:

Figure 5: Configuration Results

2 PART 3 - Report

In this part, I tried 3 different configurations for each classifier. I print the configuration accuracy, mean, standard deviation, and confidence interval for each configuration and for each classifier. Then I printed the test accuracy, mean, standard deviation, and confidence interval for each classifier. Finally, I printed the f1 score, mean, standard deviation, and confidence interval for each classifier. Now one by one I put the screenshot of each result into my report file. Firstly I want to add configuration accuracy, mean, standard deviation, mean, and confidence interval for each configuration for each classifier.

Figure 6: Configuration scores

```
SVM: kernel =  rbf , C =   100
Mean:  0.7056918415441589
Std:  0.004566659503906971
Confidence Interval:  0.0023110485709737184

Classifier:  DecisionTreeClassifier()
Decision Tree: max_depth =  1 , criterion =  gini
Mean:  0.6989763962144167
Std:  0.002742712543241927
Confidence Interval:  0.0013880040537789294

Decision Tree: max_depth =  5 , criterion =  gini
Mean:  0.7021835184977369
Std:  0.013390899384882879
Confidence Interval:  0.006776730093629696

Decision Tree: max_depth =  10 , criterion =  gini
Mean:  0.6862662626716044
Std:  0.011641468014055512
Confidence Interval:  0.005891395667862194

Classifier:  RandomForestClassifier()
Random Forest: n_estimators =  10 , max_depth =  1
Mean:  0.7000965099315454
Std:  0.00024264628888771703
Confidence Interval:  0.0001227759646884718

Random Forest: n_estimators =  10 , max_depth =  5
Mean:  0.7228863202783077
Std:  0.007935769500907556
```

Figure 7: Configuration scores

```
Random Forest: n_estimators =  10 , max_depth =  5
Mean:  0.7228863202783077
Std:  0.007935769500907556
Confidence Interval:  0.004016053473870479

Random Forest: n_estimators =  10 , max_depth =  10
Mean:  0.7272149029289642
Std:  0.010747946707114493
Confidence Interval:  0.005439211497403691
```

Figure 8: Configuration scores

As we can understand from the configuration scores, the best result is given by $RandomForestClassifier()$ with n-estimator = 10, and max-depth = 10 parameters.

Now I want to put the screenshot results of the test accuracies, mean, standard deviation, and confidence interval for each configuration.

```
Test accuracy
Classifier:  KNeighborsClassifier()
Mean:  0.6573945202687718
Std:  0.02136625659709858
Confidence Interval:  0.010812817713590668

Classifier:  SVC()
Mean:  0.7084018149886414
Std:  0.00549606984431514
Confidence Interval:  0.0027813950982790423

Classifier:  DecisionTreeClassifier()
Mean:  0.699005592418766
Std:  0.017655108676793044
Confidence Interval:  0.008934717739078333

Classifier:  RandomForestClassifier()
Mean:  0.7373882864900829
Std:  0.022601674772091573
Confidence Interval:  0.011438025571857711
```

Figure 9: Test scores

As we can seen from the test scores, the best result is given by $RandomForestClassifier()$. Now finally, I want to add the f1 scores, mean, standard deviation and confidence interval for each configuration.

```
F1 score accuracy
Classifier:  KNeighborsClassifier()
Mean:  0.7737342468376055
Std:  0.016307249761233454
Confidence Interval:  0.008252607015033065

Classifier:  SVC()
Mean:  0.8255021735758642
Std:  0.0037219910966797176
Confidence Interval:  0.0018835873788705824

Classifier:  DecisionTreeClassifier()
Mean:  0.8016535726610812
Std:  0.022289618232023595
Confidence Interval:  0.011280103173577282

Classifier:  RandomForestClassifier()
Mean:  0.8281868329546556
Std:  0.015628626131712765
Confidence Interval:  0.007909176074344142
```

Figure 10: Test scores

Now again best result is given by $RandomForestClassifier$.