CSE4088 Introduction to Machine Learning Homework 4

Berk Kırtay - 150118043

Polynomial Kernels

For all the sections, hw4.py is implemented. More details and comments can be found in the source code. The functions q2(), q3(), q4() and $q5_q6()$ must be run in the hw4.py to acquire the answers to the questions of this section.

Results after running the mentioned functions:

```
$ py hw4.py

Q2) The classifier 0 has the highest Ein among others with Ein: 10.588396653408317%

Q3) The classifier 1 has the lowest Ein among others with Ein: 1.4401316691811772%

Q4) Difference between the number of support vectors of these two classifiers is 1793

Q5) For C=0.001: Ein=0.4484304932735439%, Eout=1.650943396226412, Number of support vectors:76

Q5) For C=0.01: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:34

Q5) For C=0.1: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:24

Q6) For C=1: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:24

Q6) For C=0.0001 and Q=2: Ein=0.4484304932735439%, Eout=1.650943396226412, Number of support vectors:34

Q6) For C=0.01 and Q=2: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:34

Q6) For C=0.1 and Q=2: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:24

Q6) For C=1 and Q=2: Ein=0.34030749519538215%, Eout=1.8867924528301883, Number of support vectors:24

Q6) For C=0.0001 and Q=5: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:24

Q6) For C=0.001 and Q=5: Ein=0.3484304932735439%, Eout=1.28667924528301883, Number of support vectors:26

Q6) For C=0.001 and Q=5: Ein=0.4484304932735439%, Eout=1.28667924528301883, Number of support vectors:26

Q6) For C=0.001 and Q=5: Ein=0.3484368994234463%, Eout=2.1226415094339646, Number of support vectors:25

Q7) For C=0.01 and Q=5: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:25

Q8) For C=1 and Q=5: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:25

Q8) For C=1 and Q=5: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:25

Q8) For C=1 and Q=5: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:25

Q8) For C=1 and Q=5: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:25

Q9) For C=1 and Q=5: Ein=0.32030749519538215%,
```

2.

The classifier 0 has the highest Ein among others with Ein: 10.588396653408317So, the answer is A.

3.

The classifier 1 has the lowest Ein among others with Ein: 1.4401316691811772So, the answer is A.

4.

Difference between the number of support vectors of these two classifiers is 1793. It is the closest to the 1800, so, the answer is C.

5.

Answer dd must be correct because the maximum C achieves the lowest Ein and all other values do not strictly decrease as it is mentioned in the other answers. The output of this question:

```
Q5) For C=0.001: Ein=0.4484304932735439%, Eout=1.650943396226412, Number of support vectors:76 Q5) For C=0.01: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:34 Q5) For C=0.1: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:24 Q5) For C=1: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:24
```

So, the answer is D.

6.

All other choices are incorrect except B. The number of support vectors is lower at Q=5 for C=0.001. The output of this question:

```
Q6) For C=0.0001 and Q=2: Ein=0.8968609865470878%, Eout=1.650943396226412, Number of support vectors:236 Q6) For C=0.001 and Q=2: Ein=0.4484304932735439%, Eout=1.650943396226412, Number of support vectors:76 Q6) For C=0.01 and Q=2: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:34 Q6) For C=0.1 and Q=2: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:24 Q6) For C=1 and Q=2: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:24 Q6) For C=0.0001 and Q=5: Ein=0.4484304932735439%, Eout=1.8867924528301883, Number of support vectors:26 Q6) For C=0.001 and Q=5: Ein=0.4484304932735439%, Eout=2.1226415094339646, Number of support vectors:25 Q6) For C=0.01 and Q=5: Ein=0.384368994234463%, Eout=2.1226415094339646, Number of support vectors:23 Q6) For C=0.1 and Q=5: Ein=0.32030749519538215%, Eout=1.8867924528301883, Number of support vectors:25 Q6) For C=1 and Q=5: Ein=0.32030749519538215%, Eout=2.1226415094339646, Number of support vectors:25 Q6) For C=1 and Q=5: Ein=0.32030749519538215%, Eout=2.1226415094339646, Number of support vectors:25
```

So, the answer is B.

Cross Validation

The functions q7() and q8() must be run in the hw4.py to acquire the answers to the questions of this section.

Results after running the mentioned functions:

```
$ py hw4.py
Q7) The selection numbers of C's:
{0.0001: 25, 0.001: 37, 0.01: 21, 0.1: 9, 1: 8}
Q8) The average value of Ecv over the 100 runs is 0.00499999999999982
```

7.

As we can see from the output, C=0.001 is most selected with 37 times over 100 runs. So, the answer is B.

8.

As we can see from the output, the average value of Ecv over the 100 runs is 0.0049 which is the closest to 0.005. So, the answer is C

RBF Kernel

The function q9()_q10() must be run in the hw4.py to acquire the answers to the questions of this section.

Results after running the mentioned function:

```
$ py hw4.py
Q9) 1000000 gives the lowest value of Ein (0.06406149903908087).
Q10) 100 gives the lowest value of Eout (1.8867924528301883).
```

9.

 $1000000=10^6$ gives the lowest value of Ein. So, the answer is E.

10.

100 gives the lowest value of Ein. So, the answer is C.