# 1. Perceptron and Dual Perceptron

## 1.2 Implementation

1. Implement the perceptron algorithm and test it on the PerceptronData dataset using ten fold cross validation.
2. Implement the dual perceptron algorithm and test it on the PerceptronData dataset using ten fold cross validation.

**PerceptronData:**

**perceptron algorithm:**

```
          1    2    3    4    5       6       7    8    9      10 \
train  0.992222  1.0  1.0  1.0  1.0  0.994444  0.986667  1.00  1.0  0.974444
test   0.970000  1.0  1.0  1.0  1.0  1.000000  0.980000  0.99  1.0  0.970000

      mean accuracy  std accuracy
train      0.994778     0.008495
test       0.991000     0.012867
          1    2    3    4    5       6       7       8    9 \
train  0.992389  1.0  1.0  1.0  1.0  0.994401  0.987234  1.000000  1.0
test   0.969289  1.0  1.0  1.0  1.0  1.000000  0.981132  0.990196  1.0

          10  mean recall  std recall
train  0.975118    0.994914    0.008251
test   0.970000    0.991062    0.012894
          1    2    3    4    5       6       7    8    9      10 \
train  0.992031  1.0  1.0  1.0  1.0  0.994477  0.986425  1.00  1.0  0.974111
test   0.970612  1.0  1.0  1.0  1.0  1.000000  0.979592  0.99  1.0  0.970188

      mean precision  std precision
train       0.994704      0.008616
test        0.991039      0.012763
```

**dual perceptron algorithm:**

```
          1       2    3       4    5    6       7       8    9 \
train  0.988889  0.995556  1.0  0.996667  1.00  1.0  0.993333  0.996667  1.0
test   0.990000  1.000000  1.0  1.000000  0.99  1.0  0.980000  0.990000  1.0

       10  mean accuracy  std accuracy
train  1.0      0.997111     0.003748
test   1.0      0.995000     0.007071
          1       2    3       4       5    6       7       8 \
train  0.988966  0.995402  1.0  0.996767  1.000000  1.0  0.993065  0.996802
test   0.989362  1.000000  1.0  1.000000  0.988889  1.0  0.979968  0.990741
```

```
        9   10  mean recall  std recall
train  1.0  1.0    0.997100    0.003765
test   1.0  1.0    0.994896    0.007173
          1        2   3        4        5   6        7        8  \
train  0.988782  0.995736  1.0  0.996583  1.000000  1.0  0.993593  0.996544
test   0.990741  1.000000  1.0  1.000000  0.991071  1.0  0.979968  0.989362

        9   10  mean precision  std precision
train  1.0  1.0       0.997124       0.003741
test   1.0  1.0       0.995114       0.007002
```

3. Compare the performance of the two algorithms on the PerceptronData dataset and make sure that they have (almost) identical performance.

--From the tables above, it's obvious that they have almost identical performance in recall, precision and accuracy.

## 1.3 Kernelizing Dual Perceptron

1. Run the dual perceptron with the linear kernel on the Two Spiral dataset and show that the data is not separable using ten-fold cross validation.
2. Run the dual perceptron with the Gaussian (RBF) kernel on the Two Spiral dataset and show that the data is separable using ten-fold cross validation.

**Two Spiral:**

We can see that for linear kernel, the final metrics scores are around 0.65 or 0.64, it's low and it differs a lot each time I run this algorithm. So it's not separable.

**linear kernel:**

```
          1        2        3        4        5        6    7  \
train  0.638889  0.661111  0.667778  0.616667  0.653333  0.632222  0.66
test   0.680000  0.700000  0.600000  0.670000  0.680000  0.550000  0.64

          8        9       10  mean accuracy  std accuracy
train  0.657778  0.652222  0.646667       0.648667      0.015492
test   0.680000  0.670000  0.680000       0.655000      0.046248
          1        2        3        4        5        6        7  \
train  0.638910  0.661059  0.66783  0.616656  0.65342  0.632222  0.660058
test   0.678844  0.696507  0.59735  0.669697  0.67776  0.550000  0.637175

          8        9       10  mean recall  std recall
train  0.657736  0.652142  0.646782    0.648681    0.015497
test   0.684848  0.667677  0.687500    0.654736    0.046903
          1        2        3        4        5        6        7  \
train  0.638906  0.661103  0.667852  0.616644  0.653405  0.632223  0.660033
```

test  0.678844  0.699714  0.597980  0.668269  0.676329  0.550505  0.636071

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.657736 | 0.652172 | 0.646782 | 0.648685 | 0.015505 |
| test | 0.683661 | 0.667069 | 0.685990 | 0.654443 | 0.046729 |

For the RBF kernel it's separable because the metrics scores are around 0.73-0.75 now, which improves a lot compared to linear kernel. And when I set gamma as 5, the scores will reach as high as 1.

**Gaussian (RBF) kernel:**

(nested k folds: search for gamma within {0.25, 0.2, 0.1, 0})

the best gamma for fold 1 is 0.25
the best gamma for fold 2 is 0.25
the best gamma for fold 3 is 0.2
the best gamma for fold 4 is 0.25
the best gamma for fold 5 is 0.25
the best gamma for fold 6 is 0.25
the best gamma for fold 7 is 0.1
the best gamma for fold 8 is 0.25
the best gamma for fold 9 is 0.2
the best gamma for fold 10 is 0.25

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.737778 | 0.716667 | 0.745556 | 0.751111 | 0.737778 | 0.736667 | 0.742222 |
| test | 0.810000 | 0.700000 | 0.790000 | 0.740000 | 0.690000 | 0.770000 | 0.730000 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.695556 | 0.734444 | 0.772222 | 0.737 | 0.020219 |
| test | 0.680000 | 0.820000 | 0.690000 | 0.742 | 0.052662 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.737694 | 0.717632 | 0.746016 | 0.750993 | 0.73752 | 0.739124 | 0.739820 |
| test | 0.809295 | 0.696970 | 0.787315 | 0.741546 | 0.69686 | 0.750812 | 0.750505 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.693833 | 0.734444 | 0.774778 | 0.737185 | 0.020982 |
| test | 0.695652 | 0.820000 | 0.665239 | 0.741419 | 0.052619 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.738070 | 0.723994 | 0.796705 | 0.751216 | 0.738424 | 0.775187 | 0.796641 |
| test | 0.809916 | 0.696970 | 0.811632 | 0.740000 | 0.699510 | 0.788920 | 0.784926 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.728819 | 0.734446 | 0.806051 | 0.758955 | 0.03158 |
| test | 0.723346 | 0.822061 | 0.692857 | 0.757014 | 0.05192 |

**Gaussian (RBF) kernel: (with gamma always being 5)**

the best gamma for fold 1 is 5
the best gamma for fold 2 is 5

the best gamma for fold 3 is 5
the best gamma for fold 4 is 5
the best gamma for fold 5 is 5
the best gamma for fold 6 is 5
the best gamma for fold 7 is 5
the best gamma for fold 8 is 5
the best gamma for fold 9 is 5
the best gamma for fold 10 is 5

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | mean accuracy | \ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.00 | 1.00 | 1.0 | 1.0 | 1.0 | 1.00 | 1.00 | 1.00 | 1.0 | 1.0 | 1.000 | |
| test | 0.98 | 0.98 | 1.0 | 1.0 | 1.0 | 0.99 | 0.99 | 0.99 | 1.0 | 1.0 | 0.993 | |

| | std accuracy |
|---|---|
| train | 0.000000 |
| test | 0.008233 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | \ |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.000000 | 1.000000 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.000000 | 1.000000 | 1.0 | |
| test | 0.973684 | 0.981132 | 1.0 | 1.0 | 1.0 | 0.989796 | 0.991667 | 0.990741 | 1.0 | |

| | 10 | mean recall | std recall |
|---|---|---|---|
| train | 1.0 | 1.000000 | 0.000000 |
| test | 1.0 | 0.992702 | 0.009281 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | \ |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.000000 | 1.000000 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.000000 | 1.000000 | 1.0 | |
| test | 0.984375 | 0.979592 | 1.0 | 1.0 | 1.0 | 0.990385 | 0.987805 | 0.989362 | 1.0 | |

| | 10 | mean precision | std precision |
|---|---|---|---|
| train | 1.0 | 1.000000 | 0.000000 |
| test | 1.0 | 0.993152 | 0.007789 |

## 2. Regularized Logistic Regression

1. Do you think that w0 should be included in the regularization?
2. Calculate the gradient of the objective with respect to the model parameters.
5. Is it possible to kernelize Equation-2? If yes, then what would be the dual objective function for regularized logistic regression?

1. No. $w_0$ should not be included in the regularization. Because $w_0$ corresponds to new added all-1 column and this column doesn't have an effect on penalty or regularization.

2. $\min_w \frac{\lambda}{2}\|w\|^2 + \sum_{i=1}^{N} \ln(1+e^{-y_i w^T x_i})$ (2).

$$\frac{\partial(2)}{\partial w_0} = \sum_{i=1}^{N} (h_w(x^{(i)}) - y^{(i)}) x_0^{(i)}$$

$$\frac{\partial(2)}{\partial w_j} = \sum_{i=1}^{N} (h_w(x^{(i)}) - y^{(i)}) x_j^{(i)} + \lambda w_j \quad (j = 1,2,3,\cdots,m)$$

$$h_w(x) = \frac{1}{1+e^{-w^T x}}$$

5. Yes.

Based on the Representer theorem we can safely replace $w = \sum_i d_i y_i x_i$, so the dual version of KLR:

$$\min_d \frac{\lambda}{2} \sum_{i=1}^{N}\sum_{j=1}^{N} d_i d_j y_i y_j k(x_i, x_j) + \sum_{j=1}^{N} \ln(1+\exp(-y_i \sum_{d=1}^{N} d_i y_i k(x_i, x_j)))$$

3. Develop a gradient descent algorithm for learning the parameters from given training data.
4. Contrast the performance of Logistic Regression with Regularized Logistic Regression for the Spambase, Diabetes and Breast Cancer datasets using ten fold cross validation.

Over all, the performances of Logistic Regression and Regularized Logistic Regression are similar according to the metrics tables.

**Spambase:**

**Logistic Regression:**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.925380 | 0.925139 | 0.923931 | 0.928278 | 0.927554 | 0.926829 | 0.926588 |
| test | 0.917391 | 0.932609 | 0.926087 | 0.932609 | 0.923913 | 0.897826 | 0.939130 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.926829 | 0.931418 | 0.927071 | 0.926902 | 0.00203 |
| test | 0.923913 | 0.902174 | 0.930435 | 0.922609 | 0.01337 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.917994 | 0.917531 | 0.915602 | 0.921255 | 0.920055 | 0.918514 | 0.918944 |
| test | 0.908269 | 0.924423 | 0.918843 | 0.924945 | 0.916305 | 0.888894 | 0.933469 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.919029 | 0.924342 | 0.918979 | 0.919225 | 0.002338 |
| test | 0.915230 | 0.893156 | 0.926754 | 0.915029 | 0.014482 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.925469 | 0.925356 | 0.924064 | 0.927757 | 0.927928 | 0.927574 | 0.927153 |
| test | 0.914392 | 0.931649 | 0.929357 | 0.935689 | 0.921257 | 0.897577 | 0.935625 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.927142 | 0.931543 | 0.927226 | 0.927121 | 0.002000 |
| test | 0.925120 | 0.900799 | 0.930553 | 0.922202 | 0.013772 |

**Regularized Logistic Regression:**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.924656 | 0.927312 | 0.926346 | 0.925863 | 0.925622 | 0.929969 | 0.925380 |
| test | 0.926087 | 0.910870 | 0.926087 | 0.939130 | 0.936957 | 0.917391 | 0.923913 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.926829 | 0.927795 | 0.928037 | 0.926781 | 0.001560 |
| test | 0.900000 | 0.919565 | 0.919565 | 0.921957 | 0.011524 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.916671 | 0.920299 | 0.918581 | 0.917938 | 0.917708 | 0.923113 | 0.917887 |
| test | 0.916376 | 0.904746 | 0.919614 | 0.931385 | 0.934130 | 0.908798 | 0.913690 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.918452 | 0.919988 | 0.920403 | 0.919104 | 0.001867 |

test 0.895522 0.914087 0.907345 0.914569 0.011737

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.924868 | 0.927336 | 0.926791 | 0.926232 | 0.925761 | 0.929811 | 0.925286 |
| test | 0.929177 | 0.903753 | 0.923409 | 0.940812 | 0.934889 | 0.916227 | 0.926110 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.927407 | 0.928303 | 0.928098 | 0.926989 | 0.001518 |
| test | 0.898063 | 0.916642 | 0.927940 | 0.921702 | 0.013303 |

**Diabetes:**

**Logistic Regression:**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.781792 | 0.781792 | 0.784682 | 0.781792 | 0.781792 | 0.784682 | 0.777457 |
| test | 0.789474 | 0.776316 | 0.736842 | 0.750000 | 0.763158 | 0.763158 | 0.828947 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.774566 | 0.790462 | 0.774566 | 0.781358 | 0.004867 |
| test | 0.842105 | 0.684211 | 0.842105 | 0.777632 | 0.050180 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.735558 | 0.731975 | 0.739899 | 0.740267 | 0.733296 | 0.738440 | 0.728613 |
| test | 0.751373 | 0.740860 | 0.689231 | 0.665368 | 0.708333 | 0.730655 | 0.805385 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.726983 | 0.743168 | 0.729046 | 0.734725 | 0.005602 |
| test | 0.817308 | 0.650458 | 0.810980 | 0.736995 | 0.060171 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.769459 | 0.765789 | 0.771268 | 0.769998 | 0.767994 | 0.770308 | 0.764525 |
| test | 0.762920 | 0.800000 | 0.707071 | 0.682950 | 0.768199 | 0.748397 | 0.811373 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.761863 | 0.773800 | 0.759488 | 0.767449 | 0.004470 |
| test | 0.817308 | 0.712861 | 0.825267 | 0.763635 | 0.050356 |

**Regularized Logistic Regression:**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.784682 | 0.784682 | 0.790462 | 0.781792 | 0.770231 | 0.784682 | 0.777457 |
| test | 0.776316 | 0.763158 | 0.697368 | 0.736842 | 0.868421 | 0.815789 | 0.828947 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.776012 | 0.783237 | 0.778902 | 0.781214 | 0.005708 |
| test | 0.802632 | 0.736842 | 0.802632 | 0.782895 | 0.050866 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.731953 | 0.738440 | 0.739288 | 0.741257 | 0.728209 | 0.735951 | 0.731273 |
| test | 0.763043 | 0.708333 | 0.670168 | 0.621137 | 0.789474 | 0.778261 | 0.801176 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.725147 | 0.734619 | 0.733331 | 0.733947 | 0.005039 |
| test | 0.774395 | 0.694791 | 0.771373 | 0.737215 | 0.059932 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.772937 | 0.770308 | 0.775318 | 0.773024 | 0.757361 | 0.769974 | 0.763698 |
| test | 0.766324 | 0.768199 | 0.742273 | 0.621137 | 0.842623 | 0.842857 | 0.807692 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.760784 | 0.768984 | 0.765188 | 0.767757 | 0.005816 |
| test | 0.799679 | 0.729870 | 0.777244 | 0.769790 | 0.064543 |

**Breast Cancer datasets:**

**Logistic Regression:**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.986355 | 0.988304 | 0.990253 | 0.990253 | 0.990253 | 0.990253 | 0.988304 |
| test | 1.000000 | 0.982143 | 0.928571 | 0.982143 | 0.982143 | 0.964286 | 0.982143 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.986355 | 0.988304 | 0.988304 | 0.988694 | 0.001538 |
| test | 0.982143 | 0.982143 | 0.964286 | 0.975000 | 0.019196 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.983922 | 0.986641 | 0.987976 | 0.987771 | 0.986772 | 0.987976 | 0.985294 |
| test | 1.000000 | 0.968750 | 0.923810 | 0.980000 | 0.984848 | 0.961905 | 0.977273 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.984035 | 0.986285 | 0.985546 | 0.986222 | 0.001507 |
| test | 0.987179 | 0.979167 | 0.959064 | 0.972200 | 0.020882 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.986997 | 0.988576 | 0.991187 | 0.991204 | 0.992401 | 0.991187 | 0.989667 |
| test | 1.000000 | 0.987805 | 0.923810 | 0.984375 | 0.979167 | 0.961905 | 0.985714 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.987004 | 0.988507 | 0.989628 | 0.989636 | 0.001862 |
| test | 0.972222 | 0.984848 | 0.959064 | 0.973891 | 0.021467 |

**Regularized Logistic Regression:**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.990253 | 0.986355 | 0.986355 | 0.986355 | 0.990253 | 0.988304 | 0.986355 |
| test | 0.964286 | 1.000000 | 1.000000 | 1.000000 | 0.964286 | 0.964286 | 1.000000 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.986355 | 0.990253 | 0.986355 | 0.987719 | 0.001849 |
| test | 0.946429 | 0.964286 | 1.000000 | 0.980357 | 0.021379 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.988075 | 0.983805 | 0.983746 | 0.984035 | 0.987823 | 0.985546 | 0.984035 |

test 0.960171 1.000000 1.000000 1.000000 0.958333 0.973684 1.000000

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.983625 | 0.987875 | 0.983746 | 0.985231 | 0.001937 |
| test | 0.937500 | 0.963109 | 1.000000 | 0.979280 | 0.023544 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.991176 | 0.986989 | 0.986984 | 0.987004 | 0.991200 | 0.989628 | 0.987004 |
| test | 0.960171 | 1.000000 | 1.000000 | 1.000000 | 0.970588 | 0.950000 | 1.000000 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.986973 | 0.991196 | 0.986984 | 0.988514 | 0.002019 |
| test | 0.957143 | 0.963109 | 1.000000 | 0.980101 | 0.021575 |

# 3. Determining Model Hyper-parameters

1. Report the training and test perfomance of both kernels on the Diabetes, Breast Cancer, and Spambase datasets in a table (include mean and standard deviation of recall, precision, and accuracy). List the best hyper-parameters that you found by optimizing the accuracy over your parameter grid, for each fold.
3. For each dataset, provide a ROC-AUC curve across all k-folds.

**Spambase:**

**Linear kernel:**
the best C for fold 1 is 8
the best C for fold 2 is 16
the best C for fold 3 is 16
the best C for fold 4 is 128
the best C for fold 5 is 8
the best C for fold 6 is 16
the best C for fold 7 is 32
the best C for fold 8 is 16
the best C for fold 9 is 32
the best C for fold 10 is 1024

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.933591 | 0.936247 | 0.935281 | 0.934798 | 0.935281 | 0.935764 | 0.936247 |
| test | 0.934783 | 0.934783 | 0.934783 | 0.932609 | 0.917391 | 0.926087 | 0.932609 |

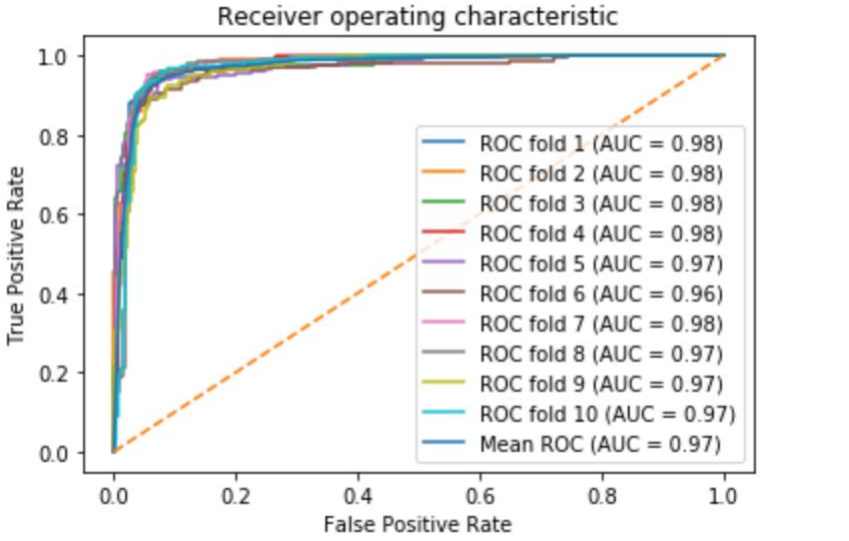
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.935523 | 0.935764 | 0.935281 | 0.935378 | 0.000774 |
| test | 0.934783 | 0.917391 | 0.934783 | 0.930000 | 0.007159 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.926945 | 0.930341 | 0.929288 | 0.928085 | 0.929800 | 0.930513 | 0.930409 |
| test | 0.927305 | 0.934105 | 0.928368 | 0.931344 | 0.903373 | 0.916090 | 0.926890 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.929723 | 0.930091 | 0.929263 | 0.929446 | 0.001138 |
| test | 0.931712 | 0.910175 | 0.930409 | 0.923977 | 0.010403 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.933457 | 0.935553 | 0.934810 | 0.934770 | 0.934292 | 0.934754 | 0.936022 |
| test | 0.936255 | 0.931334 | 0.933959 | 0.929196 | 0.923542 | 0.926348 | 0.927960 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.934620 | 0.935152 | 0.934499 | 0.934793 | 0.000697 |
| test | 0.933169 | 0.913960 | 0.934783 | 0.929051 | 0.006647 |




ROC-AUC curve for Spambase, Linear kernel:

**RBF kernel:**

the best C for fold 1 is 512.0, and the best gamma for fold 1 is 0.001953125
the best C for fold 2 is 1024.0, and the best gamma for fold 2 is 0.0009765625
the best C for fold 3 is 8.0, and the best gamma for fold 3 is 0.0078125
the best C for fold 4 is 32.0, and the best gamma for fold 4 is 0.00390625
the best C for fold 5 is 128.0, and the best gamma for fold 5 is 0.00390625
the best C for fold 6 is 128.0, and the best gamma for fold 6 is 0.001953125
the best C for fold 7 is 512.0, and the best gamma for fold 7 is 0.0009765625
the best C for fold 8 is 1024.0, and the best gamma for fold 8 is 0.001953125
the best C for fold 9 is 128.0, and the best gamma for fold 9 is 0.00390625
the best C for fold 10 is 128.0, and the best gamma for fold 10 is 0.00390625

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.965226 | 0.959672 | 0.955083 | 0.954359 | 0.965467 | 0.956049 | 0.956049 |
| test | 0.921739 | 0.921739 | 0.932609 | 0.952174 | 0.921739 | 0.934783 | 0.947826 |

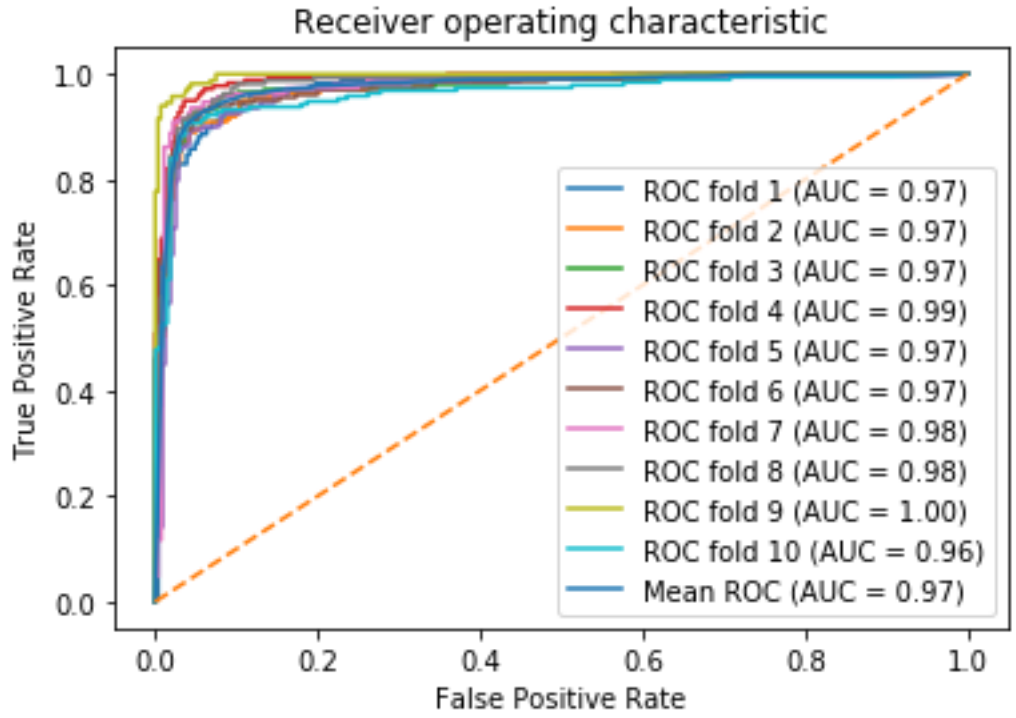
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.969331 | 0.962328 | 0.964501 | 0.960807 | 0.005279 |
| test | 0.936957 | 0.965217 | 0.932609 | 0.936739 | 0.014473 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | \ |
|---|---|---|---|---|---|---|---|---|
| train | 0.961411 | 0.955317 | 0.950364 | 0.949625 | 0.961244 | 0.951458 | 0.951954 | |
| test | 0.915013 | 0.916113 | 0.923010 | 0.953757 | 0.912166 | 0.928649 | 0.942741 | |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.965812 | 0.958312 | 0.960145 | 0.956564 | 0.005600 |
| test | 0.934571 | 0.964921 | 0.926869 | 0.931781 | 0.017421 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | \ |
|---|---|---|---|---|---|---|---|---|
| train | 0.965749 | 0.959725 | 0.955470 | 0.954639 | 0.966260 | 0.956234 | 0.955674 | |
| test | 0.918766 | 0.924816 | 0.934165 | 0.947073 | 0.925778 | 0.935358 | 0.949206 | |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.970160 | 0.962941 | 0.965221 | 0.961207 | 0.005568 |
| test | 0.928857 | 0.960243 | 0.934758 | 0.935902 | 0.012770 |




ROC-AUC curve for Spambase, RBF kernel

**Breast Cancer dataset:**

**Linear kernel:**

the best C for fold 1 is 0.03125
the best C for fold 2 is 0.25
the best C for fold 3 is 0.25
the best C for fold 4 is 0.5
the best C for fold 5 is 0.0625

the best C for fold 6 is 0.0625
the best C for fold 7 is 0.0625
the best C for fold 8 is 0.03125
the best C for fold 9 is 0.03125
the best C for fold 10 is 0.25

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.978558 | 0.984405 | 0.988304 | 0.988304 | 0.988304 | 0.980507 | 0.982456 |
| test | 0.982143 | 1.000000 | 0.946429 | 0.982143 | 0.928571 | 0.982143 | 0.982143 |

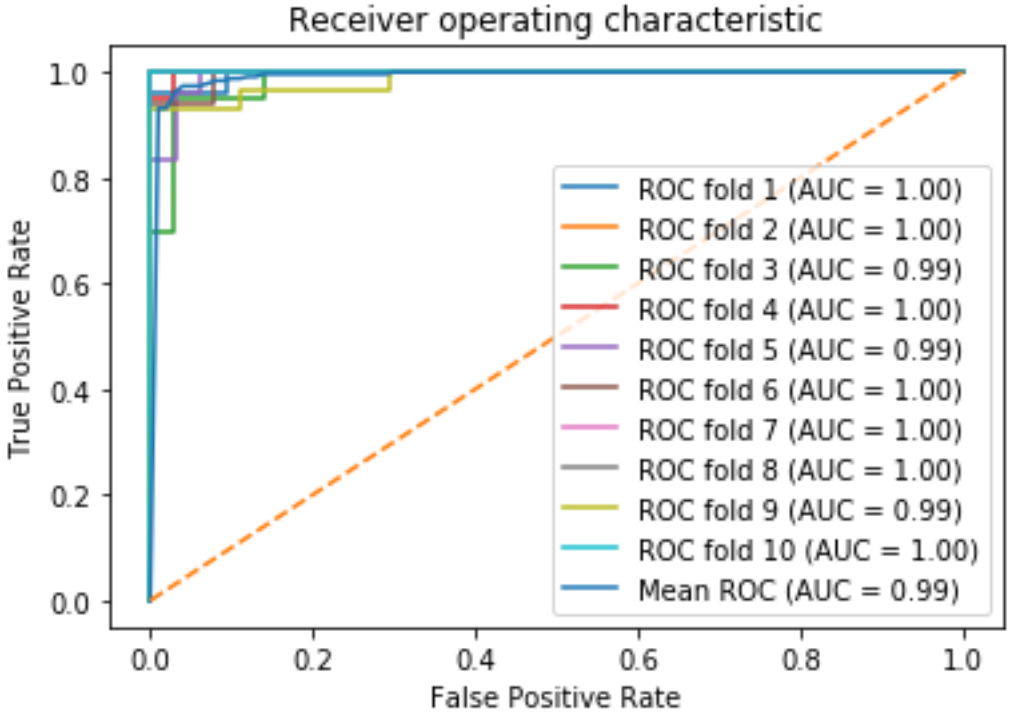
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.980507 | 0.984405 | 0.986355 | 0.984211 | 0.003612 |
| test | 1.000000 | 0.964286 | 0.982143 | 0.975000 | 0.022588 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.971866 | 0.981188 | 0.986468 | 0.984293 | 0.985164 | 0.976343 | 0.977814 |
| test | 0.979167 | 1.000000 | 0.936111 | 0.976190 | 0.921875 | 0.970588 | 0.986842 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.975351 | 0.980576 | 0.984035 | 0.980310 | 0.004830 |
| test | 1.000000 | 0.965517 | 0.987179 | 0.972347 | 0.025671 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.982237 | 0.98545 | 0.988545 | 0.990854 | 0.989684 | 0.982363 | 0.985056 |
| test | 0.984848 | 1.00000 | 0.946657 | 0.986111 | 0.933155 | 0.987500 | 0.973684 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.983523 | 0.985431 | 0.987004 | 0.986015 | 0.002975 |
| test | 1.000000 | 0.965517 | 0.972222 | 0.974970 | 0.021808 |




ROC-AUC curve for Breast Cancer, linear kernel

**RBF kernel:**

the best C for fold 1 is 64.0, and the best gamma for fold 1 is 0.00048828125
the best C for fold 2 is 16.0, and the best gamma for fold 2 is 0.001953125
the best C for fold 3 is 128.0, and the best gamma for fold 3 is 0.001953125
the best C for fold 4 is 16.0, and the best gamma for fold 4 is 0.00390625
the best C for fold 5 is 4.0, and the best gamma for fold 5 is 0.015625
the best C for fold 6 is 8.0, and the best gamma for fold 6 is 0.00390625
the best C for fold 7 is 512.0, and the best gamma for fold 7 is 0.0001220703125
the best C for fold 8 is 4.0, and the best gamma for fold 8 is 0.03125
the best C for fold 9 is 64.0, and the best gamma for fold 9 is 0.0009765625
the best C for fold 10 is 128.0, and the best gamma for fold 10 is 0.0009765625

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.982456 | 0.984405 | 0.988304 | 0.984405 | 0.990253 | 0.984405 | 0.990253 |
| test | 1.000000 | 0.982143 | 0.982143 | 1.000000 | 0.964286 | 0.982143 | 0.928571 |

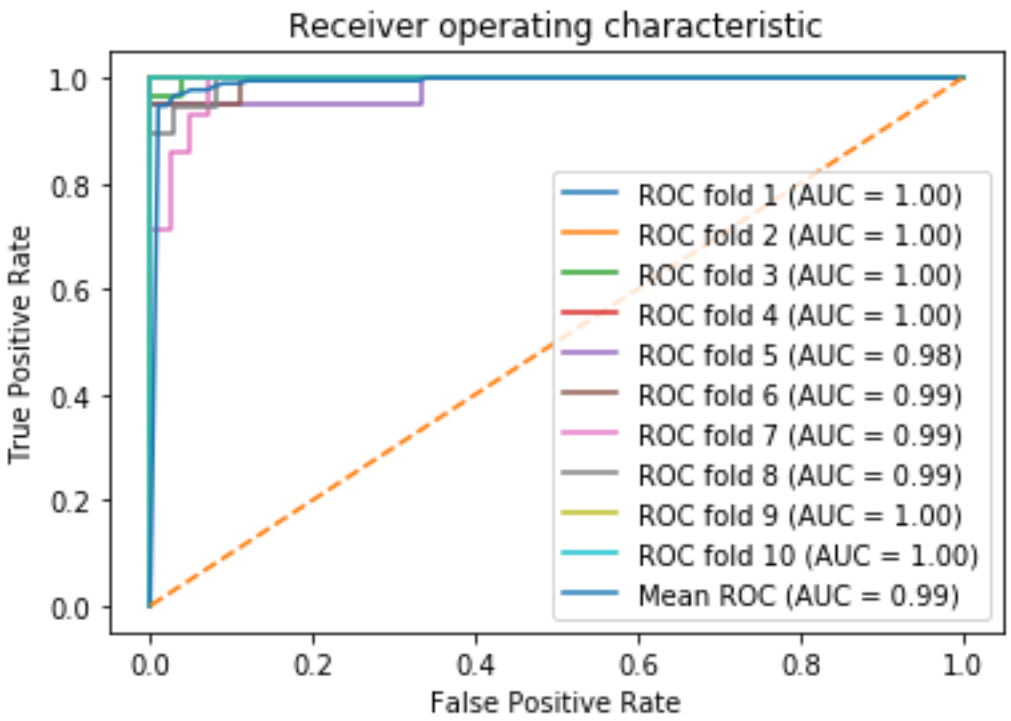
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.990253 | 0.980507 | 0.986355 | 0.986160 | 0.003493 |
| test | 0.964286 | 0.964286 | 1.000000 | 0.976786 | 0.022351 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.977712 | 0.980479 | 0.983607 | 0.980031 | 0.986979 | 0.98126 | 0.989250 |
| test | 1.000000 | 0.987179 | 0.982759 | 1.000000 | 0.961111 | 0.97500 | 0.880952 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.987047 | 0.975542 | 0.982584 | 0.982449 | 0.004353 |
| test | 0.960171 | 0.960000 | 1.000000 | 0.970717 | 0.035564 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.985079 | 0.986552 | 0.991071 | 0.986644 | 0.992331 | 0.985450 | 0.990177 |
| test | 1.000000 | 0.972222 | 0.982143 | 1.000000 | 0.961111 | 0.986486 | 0.924242 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.992308 | 0.982427 | 0.988164 | 0.988020 | 0.003359 |
| test | 0.960171 | 0.969697 | 1.000000 | 0.975607 | 0.023756 |




ROC-AUC curve for Breast Cancer, RBF kernel

**Diabetes:**

**Linear kernel:**

the best C for fold 1 is 4.0
the best C for fold 2 is 1024.0
the best C for fold 3 is 0.25
the best C for fold 4 is 16.0
the best C for fold 5 is 8.0
the best C for fold 6 is 0.25
the best C for fold 7 is 1.0
the best C for fold 8 is 0.5
the best C for fold 9 is 4.0
the best C for fold 10 is 1024.0

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.783237 | 0.774566 | 0.773121 | 0.778902 | 0.778902 | 0.778902 | 0.777457 |
| test | 0.763158 | 0.789474 | 0.815789 | 0.776316 | 0.776316 | 0.736842 | 0.802632 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.780347 | 0.768786 | 0.776012 | 0.777023 | 0.00409 |
| test | 0.684211 | 0.802632 | 0.750000 | 0.769737 | 0.03886 |

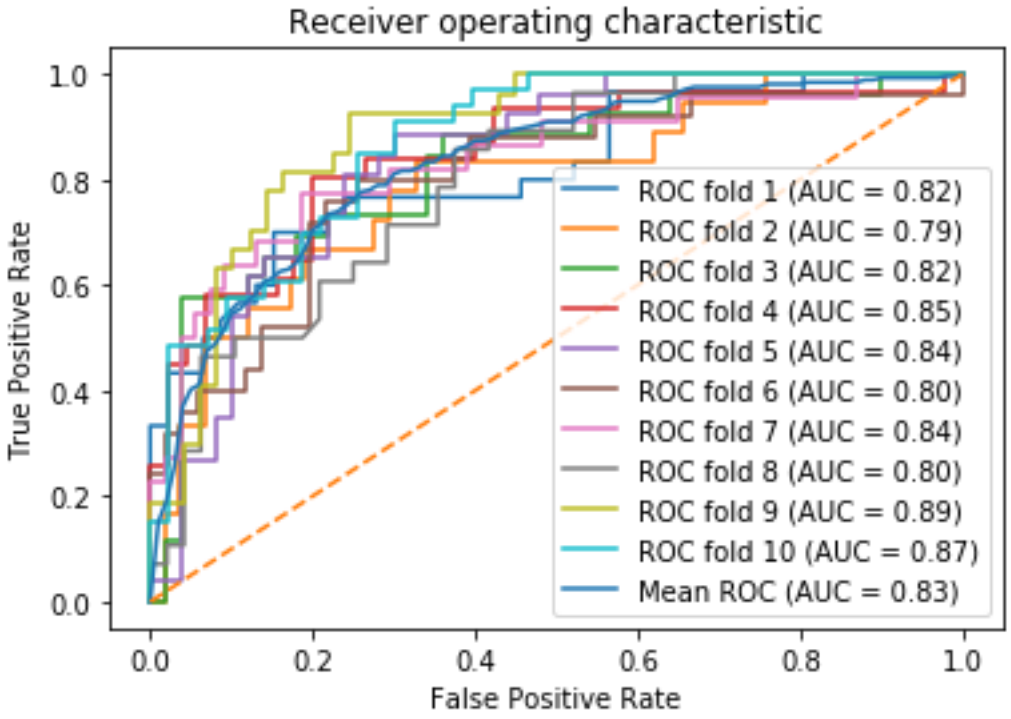
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.735849 | 0.733176 | 0.722415 | 0.727755 | 0.730680 | 0.734275 | 0.732564 |

test  0.728986  0.689655  0.758462  0.740860  0.746923  0.650980  0.753367

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.731213 | 0.718284 | 0.726025 | 0.729224 | 0.005585 |
| test | 0.645833 | 0.763794 | 0.729739 | 0.720860 | 0.043492 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.767184 | 0.763424 | 0.760030 | 0.762752 | 0.766014 | 0.764574 | 0.765792 |
| test | 0.765993 | 0.706250 | 0.821839 | 0.800000 | 0.751765 | 0.710383 | 0.760606 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.766501 | 0.752951 | 0.756185 | 0.762541 | 0.004755 |
| test | 0.657051 | 0.793771 | 0.761218 | 0.752888 | 0.049569 |




ROC-AUC curve for Diabetes, linear kernel

**RBF kernel:**

the best C for fold 1 is 8.0, and the best gamma for fold 1 is 0.03125
the best C for fold 2 is 0.5, and the best gamma for fold 2 is 0.0625
the best C for fold 3 is 128.0, and the best gamma for fold 3 is 0.00048828125
the best C for fold 4 is 1.0, and the best gamma for fold 4 is 0.125
the best C for fold 5 is 0.5, and the best gamma for fold 5 is 0.015625
the best C for fold 6 is 1.0, and the best gamma for fold 6 is 0.0078125
the best C for fold 7 is 32.0, and the best gamma for fold 7 is 0.0009765625
the best C for fold 8 is 64.0, and the best gamma for fold 8 is 0.000244140625
the best C for fold 9 is 32.0, and the best gamma for fold 9 is 0.00390625
the best C for fold 10 is 1.0, and the best gamma for fold 10 is 0.03125

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.823699 | 0.784682 | 0.789017 | 0.819364 | 0.784682 | 0.784682 | 0.778902 |

test  0.684211  0.789474  0.750000  0.723684  0.776316  0.776316  0.763158

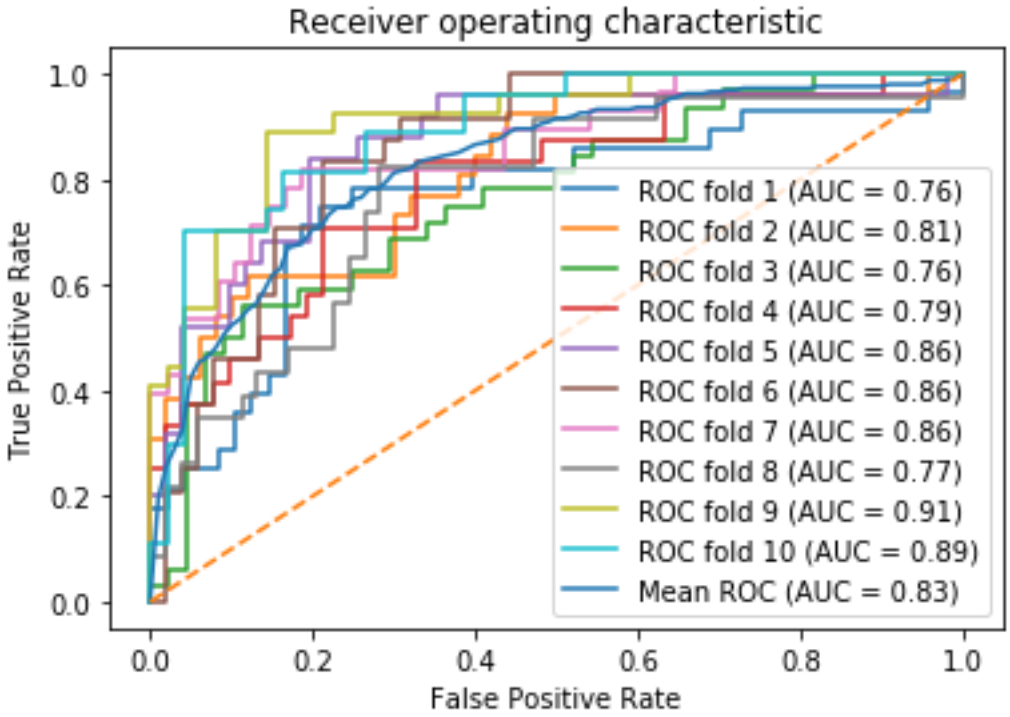
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.781792 | 0.787572 | 0.777457 | 0.791185 | 0.016409 |
| test | 0.723684 | 0.776316 | 0.868421 | 0.763158 | 0.049232 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.776125 | 0.737034 | 0.738719 | 0.777444 | 0.729289 | 0.726398 | 0.730107 |
| test | 0.630952 | 0.738462 | 0.724432 | 0.674679 | 0.710980 | 0.724359 | 0.693452 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.733338 | 0.733662 | 0.725902 | 0.740802 | 0.019414 |
| test | 0.641920 | 0.701814 | 0.831444 | 0.707249 | 0.056145 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.825022 | 0.773341 | 0.774842 | 0.816937 | 0.781913 | 0.787010 | 0.764330 |
| test | 0.657143 | 0.776786 | 0.759229 | 0.678835 | 0.757663 | 0.742424 | 0.799539 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.773387 | 0.781776 | 0.765469 | 0.784403 | 0.020615 |
| test | 0.665709 | 0.807604 | 0.879654 | 0.752458 | 0.070105 |




ROC-AUC curve for Diabetes, RBF kernel

2. Repeat the grid search in the first part but instead of optimizing accuracy, **optimize AUC**. Note: In order calculate AUC will need the SVM to predict class probabilities, instead of predicting labels.
3. For each dataset, provide a ROC-AUC curve across all k-folds.

**Spambase:**

**Linear kernel:**
the best C for fold 1 is 16
the best C for fold 2 is 16
the best C for fold 3 is 8
the best C for fold 4 is 8
the best C for fold 5 is 32
the best C for fold 6 is 16
the best C for fold 7 is 128
the best C for fold 8 is 32
the best C for fold 9 is 64
the best C for fold 10 is 8

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.934798 | 0.936730 | 0.935281 | 0.936247 | 0.936730 | 0.933832 | 0.933591 |
| test | 0.910870 | 0.928261 | 0.930435 | 0.945652 | 0.921739 | 0.939130 | 0.950000 |

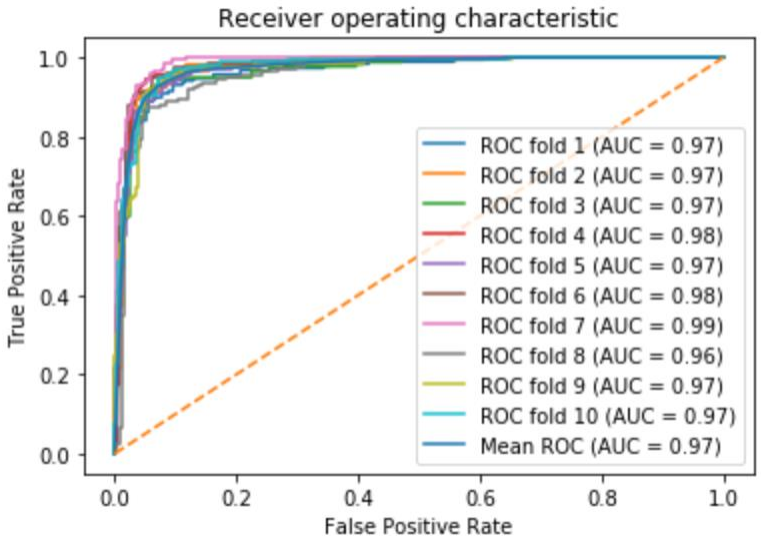
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.937696 | 0.936006 | 0.934557 | 0.935547 | 0.001354 |
| test | 0.913043 | 0.932609 | 0.928261 | 0.930000 | 0.012750 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.928888 | 0.931676 | 0.929159 | 0.930405 | 0.930565 | 0.92807 | 0.927479 |
| test | 0.903542 | 0.917663 | 0.927976 | 0.942935 | 0.916855 | 0.93317 | 0.944356 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.932214 | 0.930502 | 0.928202 | 0.929716 | 0.001595 |
| test | 0.905156 | 0.932381 | 0.922651 | 0.924668 | 0.014163 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.933892 | 0.935549 | 0.934905 | 0.935581 | 0.936400 | 0.932807 | 0.933365 |
| test | 0.911766 | 0.930714 | 0.926318 | 0.943747 | 0.919959 | 0.939949 | 0.947771 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.936639 | 0.935256 | 0.934427 | 0.934882 | 0.001258 |
| test | 0.915574 | 0.926219 | 0.925428 | 0.928745 | 0.011920 |




ROC-AUC curve for Spambase, Linear kernel, optimize AUC

**RBF kernel:**

the best C for fold 1 is 128, and the best gamma for fold 1 is 0.00390625
the best C for fold 2 is 128, and the best gamma for fold 2 is 0.00390625
the best C for fold 3 is 512, and the best gamma for fold 3 is 0.0009765625
the best C for fold 4 is 32, and the best gamma for fold 4 is 0.00390625
the best C for fold 5 is 1024, and the best gamma for fold 5 is 0.0009765625
the best C for fold 6 is 128, and the best gamma for fold 6 is 0.001953125
the best C for fold 7 is 512, and the best gamma for fold 7 is 0.001953125
the best C for fold 8 is 1024, and the best gamma for fold 8 is 0.001953125
the best C for fold 9 is 128, and the best gamma for fold 9 is 0.00390625
the best C for fold 10 is 512, and the best gamma for fold 10 is 0.0009765625

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.962569 | 0.963777 | 0.956291 | 0.955808 | 0.961362 | 0.954842 | 0.965226 |
| test | 0.945652 | 0.945652 | 0.941304 | 0.930435 | 0.934783 | 0.932609 | 0.939130 |

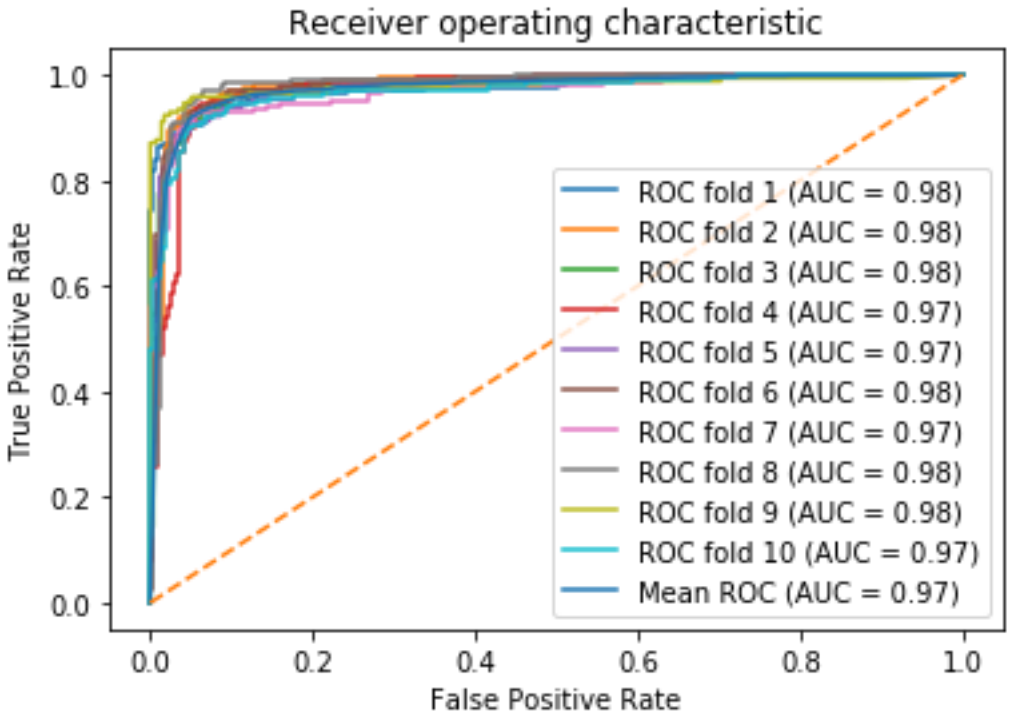
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.968124 | 0.963294 | 0.957981 | 0.960927 | 0.004476 |
| test | 0.947826 | 0.958696 | 0.926087 | 0.940217 | 0.009681 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.958594 | 0.959300 | 0.951939 | 0.951351 | 0.957567 | 0.950679 | 0.961399 |
| test | 0.937393 | 0.941993 | 0.935918 | 0.925202 | 0.928024 | 0.927424 | 0.933285 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.963989 | 0.959354 | 0.953684 | 0.956785 | 0.004593 |
| test | 0.946707 | 0.951608 | 0.923279 | 0.935083 | 0.009491 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.963003 | 0.964505 | 0.956487 | 0.955808 | 0.961156 | 0.954617 | 0.965878 |
| test | 0.946992 | 0.946652 | 0.938961 | 0.929960 | 0.938508 | 0.930175 | 0.935481 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.969136 | 0.963835 | 0.958053 | 0.961248 | 0.004836 |
| test | 0.945205 | 0.959748 | 0.923279 | 0.939496 | 0.010585 |




ROC-AUC curve for Spambase, RBF kernel, optimize AUC

**Breast Cancer:**

**Linear kernel:**

the best C for fold 1 is 2.0
the best C for fold 2 is 0.03125
the best C for fold 3 is 1.0
the best C for fold 4 is 0.125
the best C for fold 5 is 0.125
the best C for fold 6 is 0.125
the best C for fold 7 is 0.5
the best C for fold 8 is 1.0
the best C for fold 9 is 0.0625
the best C for fold 10 is 0.03125

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.986355 | 0.978558 | 0.992203 | 0.984405 | 0.982456 | 0.986355 | 0.986355 |
| test | 0.982143 | 1.000000 | 0.928571 | 0.982143 | 0.982143 | 0.982143 | 0.964286 |

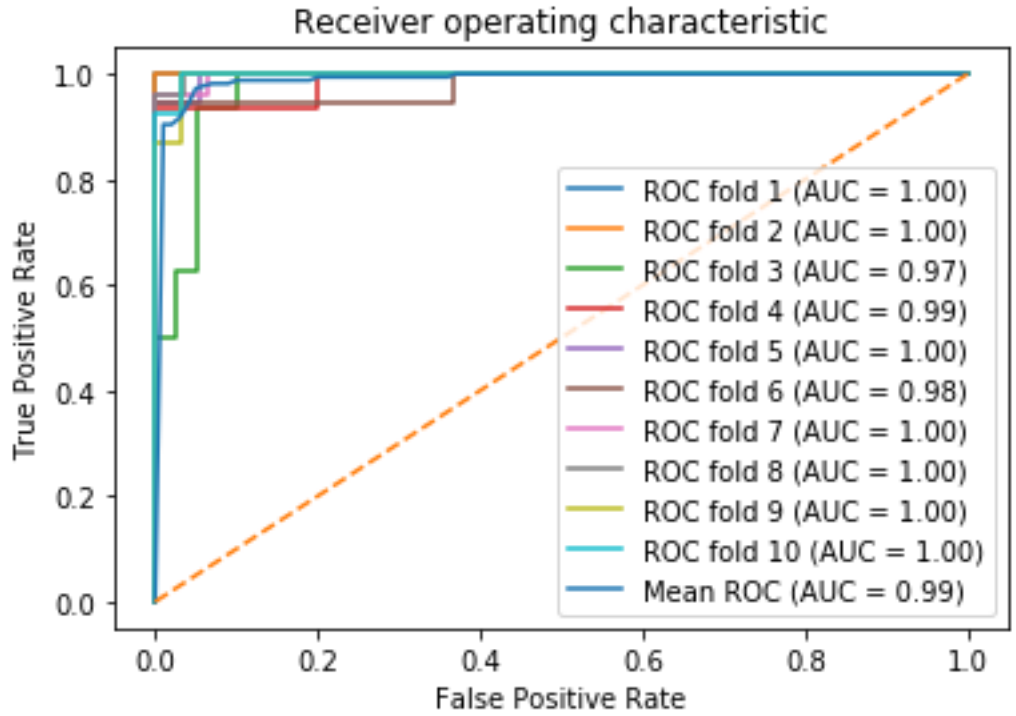
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.988304 | 0.982456 | 0.978558 | 0.984600 | 0.004256 |
| test | 0.946429 | 0.982143 | 0.946429 | 0.969643 | 0.022351 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.984746 | 0.973886 | 0.989796 | 0.982513 | 0.978740 | 0.982969 | 0.982423 |
| test | 0.979167 | 1.000000 | 0.931250 | 0.968750 | 0.973684 | 0.972222 | 0.960000 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.986140 | 0.976190 | 0.971589 | 0.980899 | 0.005709 |
| test | 0.944444 | 0.984848 | 0.942308 | 0.965667 | 0.021235 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.985844 | 0.98082 | 0.993769 | 0.984431 | 0.983912 | 0.988095 | 0.988189 |
| test | 0.984848 | 1.00000 | 0.903509 | 0.987805 | 0.986842 | 0.987179 | 0.969697 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.988475 | 0.986486 | 0.982294 | 0.986232 | 0.003694 |
| test | 0.953125 | 0.979167 | 0.954545 | 0.970672 | 0.027969 |




ROC-AUC curve for Breast Cancer, linear kernel, optimize AUC

**RBF kernel:**

the best C for fold 1 is 32.0, and the best gamma for fold 1 is 0.0009765625
the best C for fold 2 is 128.0, and the best gamma for fold 2 is 0.0001220703125
the best C for fold 3 is 2.0, and the best gamma for fold 3 is 0.0078125
the best C for fold 4 is 8.0, and the best gamma for fold 4 is 0.015625
the best C for fold 5 is 4.0, and the best gamma for fold 5 is 0.0078125
the best C for fold 6 is 32.0, and the best gamma for fold 6 is 0.00390625
the best C for fold 7 is 4.0, and the best gamma for fold 7 is 0.03125
the best C for fold 8 is 16.0, and the best gamma for fold 8 is 0.001953125
the best C for fold 9 is 16.0, and the best gamma for fold 9 is 0.00390625

the best C for fold 10 is 512.0, and the best gamma for fold 10 is 0.000244140625

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.980507 | 0.986355 | 0.982456 | 0.988304 | 0.980507 | 0.992203 | 0.988304 |
| test | 1.000000 | 0.964286 | 0.964286 | 0.946429 | 0.982143 | 0.964286 | 1.000000 |

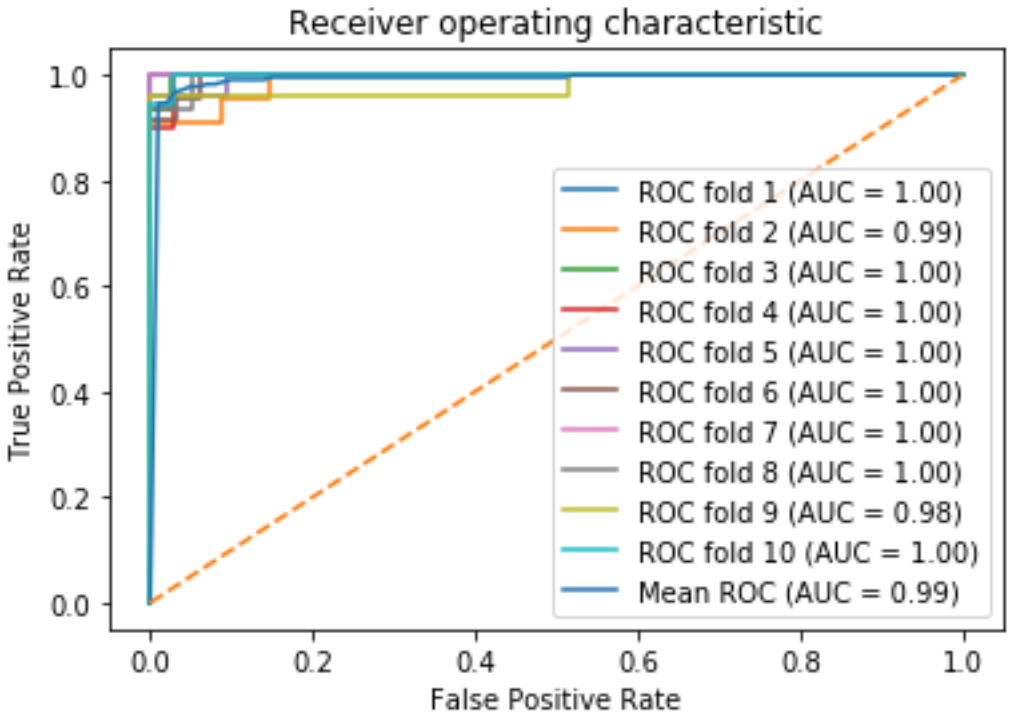
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.982456 | 0.986355 | 0.986355 | 0.985380 | 0.003817 |
| test | 0.964286 | 0.964286 | 0.982143 | 0.973214 | 0.017354 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.976531 | 0.982663 | 0.978015 | 0.984375 | 0.976768 | 0.989418 | 0.983696 |
| test | 1.000000 | 0.954545 | 0.937500 | 0.936111 | 0.979167 | 0.963109 | 1.000000 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.978015 | 0.982423 | 0.983922 | 0.981582 | 0.004152 |
| test | 0.937500 | 0.963871 | 0.973684 | 0.964549 | 0.024003 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.982342 | 0.988151 | 0.985009 | 0.990826 | 0.981210 | 0.993902 | 0.991045 |
| test | 1.000000 | 0.972222 | 0.976190 | 0.946657 | 0.984848 | 0.963109 | 1.000000 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.985009 | 0.988189 | 0.986997 | 0.987268 | 0.003997 |
| test | 0.976190 | 0.963871 | 0.986842 | 0.976993 | 0.016755 |




ROC-AUC curve for Breast Cancer, RBF kernel, optimize AUC

**Diabetes:**

**Linear kernel:**

the best C for fold 1 is 0.25
the best C for fold 2 is 64.0
the best C for fold 3 is 8.0
the best C for fold 4 is 1.0
the best C for fold 5 is 1024.0
the best C for fold 6 is 4.0
the best C for fold 7 is 0.03125
the best C for fold 8 is 2.0
the best C for fold 9 is 32.0
the best C for fold 10 is 1024.0

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.767341 | 0.784682 | 0.789017 | 0.767341 | 0.776012 | 0.776012 | 0.777457 |
| test | 0.802632 | 0.697368 | 0.671053 | 0.802632 | 0.763158 | 0.776316 | 0.802632 |

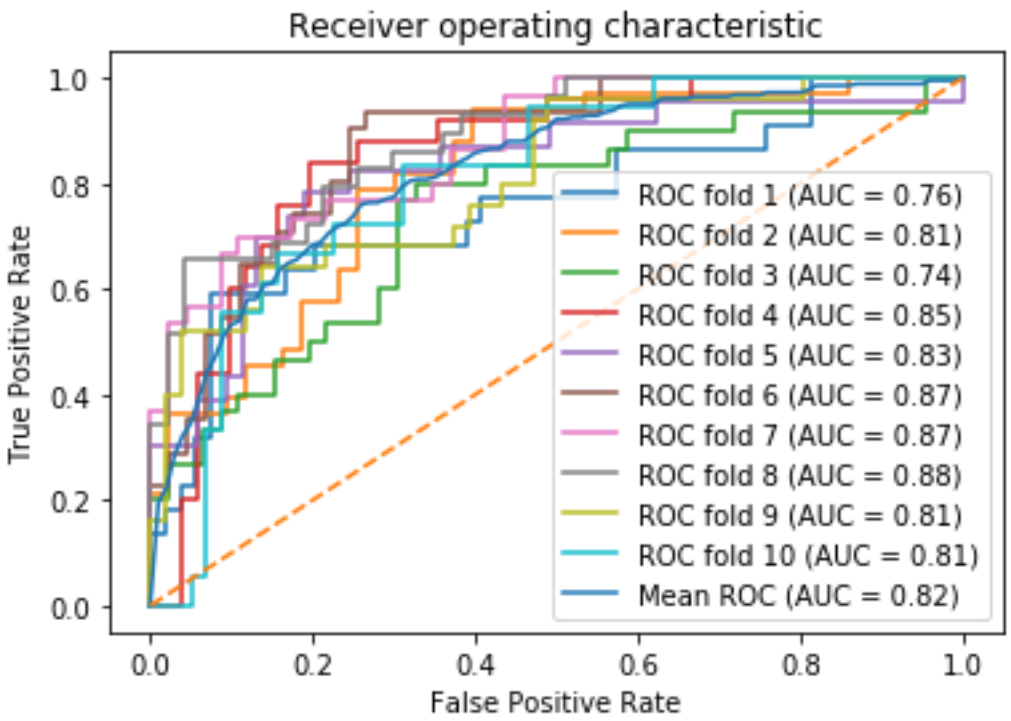
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.778902 | 0.780347 | 0.773121 | 0.777023 | 0.006848 |
| test | 0.815789 | 0.776316 | 0.763158 | 0.767105 | 0.047664 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.723805 | 0.732590 | 0.742254 | 0.718758 | 0.729786 | 0.721514 | 0.719450 |
| test | 0.739899 | 0.683228 | 0.635507 | 0.750980 | 0.670221 | 0.750896 | 0.755797 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.730319 | 0.736332 | 0.732914 | 0.728772 | 0.007736 |
| test | 0.785033 | 0.710980 | 0.729885 | 0.721243 | 0.045891 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.751807 | 0.769112 | 0.774439 | 0.751900 | 0.763839 | 0.761083 | 0.767208 |
| test | 0.763158 | 0.693452 | 0.653404 | 0.785714 | 0.726776 | 0.780449 | 0.851944 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.763171 | 0.766063 | 0.760784 | 0.762941 | 0.007095 |
| test | 0.818704 | 0.757663 | 0.692308 | 0.752357 | 0.061362 |




ROC-AUC curve for Diabetes, linear kernel, optimize AUC

**RBF kernel:**

the best C for fold 1 is 256.0, and the best gamma for fold 1 is 0.00390625
the best C for fold 2 is 4.0, and the best gamma for fold 2 is 0.03125
the best C for fold 3 is 16.0, and the best gamma for fold 3 is 0.00048828125
the best C for fold 4 is 64.0, and the best gamma for fold 4 is 0.0078125
the best C for fold 5 is 1.0, and the best gamma for fold 5 is 0.03125
the best C for fold 6 is 128.0, and the best gamma for fold 6 is 0.0001220703125
the best C for fold 7 is 1.0, and the best gamma for fold 7 is 0.0625
the best C for fold 8 is 4.0, and the best gamma for fold 8 is 0.015625
the best C for fold 9 is 128.0, and the best gamma for fold 9 is 6.103515625e-05
the best C for fold 10 is 16.0, and the best gamma for fold 10 is 0.00390625

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.793353 | 0.807803 | 0.767341 | 0.810694 | 0.789017 | 0.776012 | 0.802023 |
| test | 0.776316 | 0.736842 | 0.868421 | 0.697368 | 0.763158 | 0.763158 | 0.815789 |

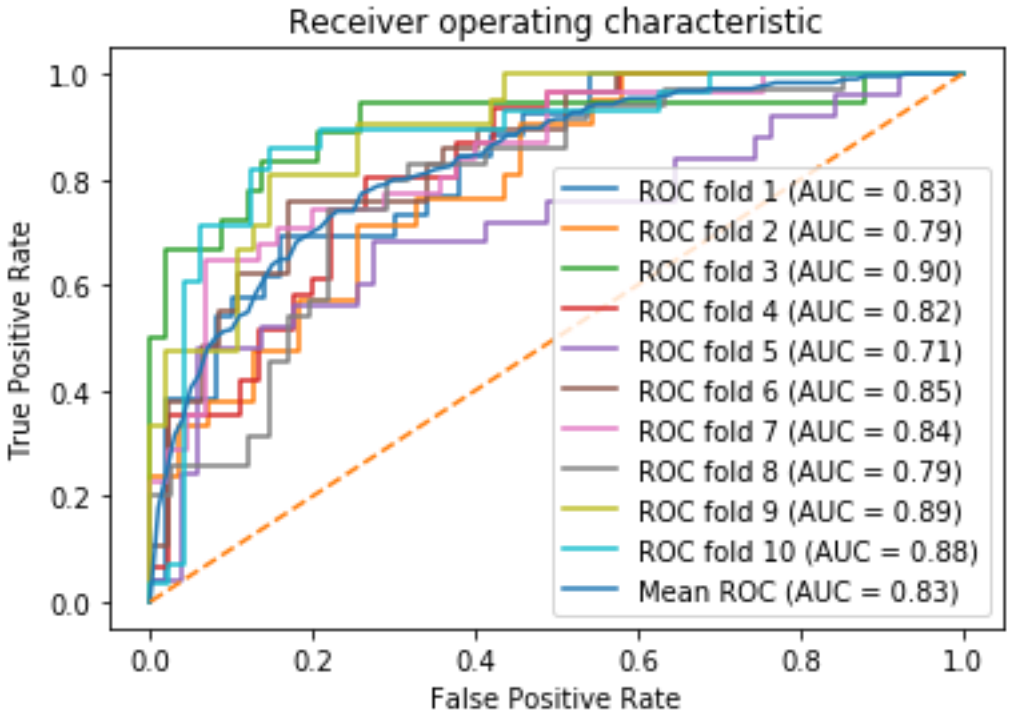
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.802023 | 0.777457 | 0.780347 | 0.790607 | 0.014916 |
| test | 0.671053 | 0.802632 | 0.828947 | 0.772368 | 0.060153 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.743701 | 0.764095 | 0.717964 | 0.756985 | 0.743958 | 0.720206 | 0.753424 |
| test | 0.756154 | 0.656277 | 0.798851 | 0.659140 | 0.690980 | 0.716067 | 0.789247 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.745103 | 0.726088 | 0.725350 | 0.739687 | 0.016293 |
| test | 0.655401 | 0.701732 | 0.790179 | 0.721403 | 0.058075 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.786936 | 0.806358 | 0.760404 | 0.809839 | 0.778161 | 0.764512 | 0.792366 |
| test | 0.751701 | 0.666667 | 0.825000 | 0.698214 | 0.742772 | 0.771930 | 0.831009 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.795391 | 0.772618 | 0.771224 | 0.783781 | 0.017123 |
| test | 0.687710 | 0.768433 | 0.837662 | 0.758110 | 0.061098 |




ROC-AUC curve for Diabetes, RBF kernel, optimize AUC

## 4. SVMs vs Multiclass Problems

1. Report the training and test perfomance of both kernels on the Wine dataset in a table (include mean and standard deviation of recall, precision, and accuracy per class). List the best hyper-parameters that you found by optimizing the accuracy over your parameter grid, for each fold per class.

2. Provide a ROC-AUC curve for each class, across all k-folds.

**Wine dataset:**

**Class 1 linear kernel:**

the best C for fold 1 is 0.125
the best C for fold 2 is 0.125
the best C for fold 3 is 0.125
the best C for fold 4 is 32.0
the best C for fold 5 is 0.25
the best C for fold 6 is 2.0
the best C for fold 7 is 0.0625
the best C for fold 8 is 0.125
the best C for fold 9 is 128.0
the best C for fold 10 is 2.0

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 \ |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.000000 |
| test | 1.0 | 1.0 | 1.0 | 0.941176 | 1.0 | 1.0 | 1.0 | 0.882353 | 1.0 | 0.941176 |

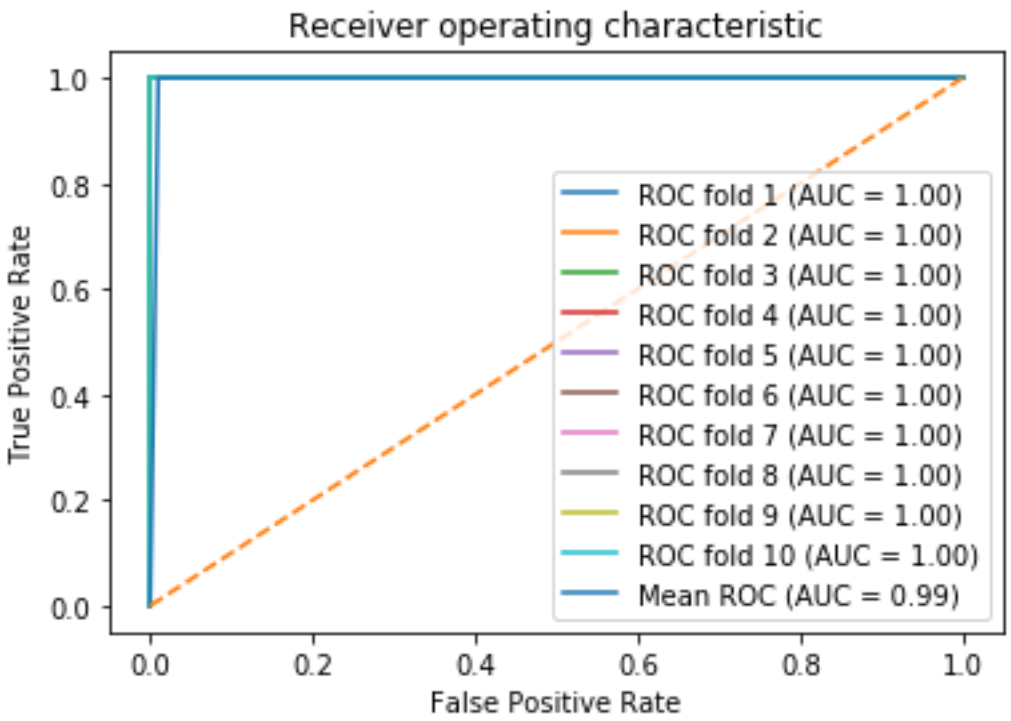
| | mean accuracy | std accuracy |
|---|---|---|
| train | 1.000000 | 0.00000 |
| test | 0.976471 | 0.04113 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 \ |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.000000 |
| test | 1.0 | 1.0 | 1.0 | 0.958333 | 1.0 | 1.0 | 1.0 | 0.916667 | 1.0 | 0.961538 |

| | mean recall | std recall |
|---|---|---|
| train | 1.000000 | 0.000000 |
| test | 0.983654 | 0.028844 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 \ |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 |
| test | 1.0 | 1.0 | 1.0 | 0.916667 | 1.0 | 1.0 | 1.0 | 0.857143 | 1.0 | 0.9 |

| | mean precision | std precision |
|---|---|---|
| train | 1.000000 | 0.00000 |
| test | 0.967381 | 0.05448 |




ROC-AUC curve for class 1, linear kernel

**Class 1 RBF kernel:**

the best C for fold 1 is 1024.0, and the best gamma for fold 1 is 0.000244140625
the best C for fold 2 is 512.0, and the best gamma for fold 2 is 0.000244140625
the best C for fold 3 is 512.0, and the best gamma for fold 3 is 0.0001220703125
the best C for fold 4 is 256.0, and the best gamma for fold 4 is 0.0001220703125
the best C for fold 5 is 256.0, and the best gamma for fold 5 is 0.0001220703125
the best C for fold 6 is 1024.0, and the best gamma for fold 6 is 3.0517578125e-05
the best C for fold 7 is 256.0, and the best gamma for fold 7 is 0.0001220703125
the best C for fold 8 is 1024.0, and the best gamma for fold 8 is 3.0517578125e-05
the best C for fold 9 is 256.0, and the best gamma for fold 9 is 0.000244140625
the best C for fold 10 is 1024.0, and the best gamma for fold 10 is 3.0517578125e-05

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 | 0.993789 | 1.0 |
| test | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.941176 | 1.0 | 1.0 | 1.000000 | 1.0 |

| | mean accuracy | std accuracy |
|---|---|---|
| train | 0.999379 | 0.001964 |
| test | 0.994118 | 0.018602 |

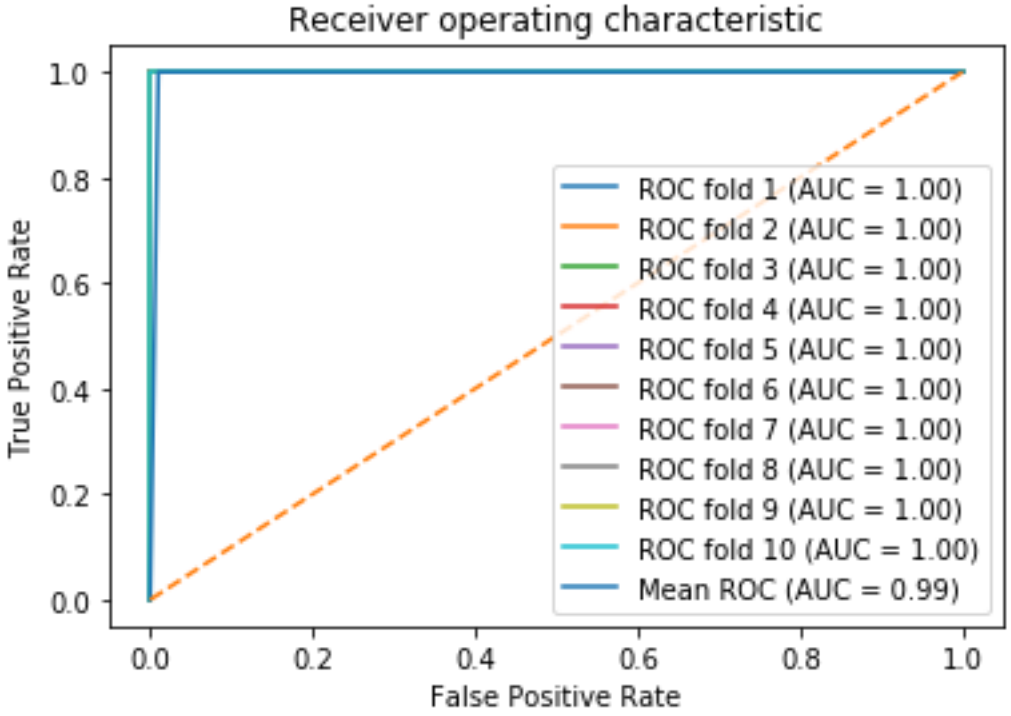
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 | 0.991071 | 1.0 |
| test | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.958333 | 1.0 | 1.0 | 1.000000 | 1.0 |

| | mean recall | std recall |
|---|---|---|
| train | 0.999107 | 0.002823 |

test 0.995833 0.013176

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 \ |
|---|---|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.000000 | 1.0 | 1.0 | 0.995283 | 1.0 |
| test | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 0.916667 | 1.0 | 1.0 | 1.000000 | 1.0 |

| | mean precision | std precision |
|---|---|---|
| train | 0.999528 | 0.001492 |
| test | 0.991667 | 0.026352 |




ROC-AUC curve for class 1, RBF kernel

**Class 2 linear kernel:**

the best C for fold 1 is 0.0625
the best C for fold 2 is 0.25
the best C for fold 3 is 1.0
the best C for fold 4 is 4.0
the best C for fold 5 is 0.03125
the best C for fold 6 is 0.125
the best C for fold 7 is 0.0625
the best C for fold 8 is 0.0625
the best C for fold 9 is 0.0625
the best C for fold 10 is 0.5

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.987578 | 0.987578 | 0.993789 | 1.000000 | 0.987578 | 0.993789 | 0.987578 |
| test | 0.941176 | 1.000000 | 1.000000 | 0.941176 | 1.000000 | 1.000000 | 1.000000 |

| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.987578 | 0.987578 | 0.993789 | 0.990683 | 0.004392 |

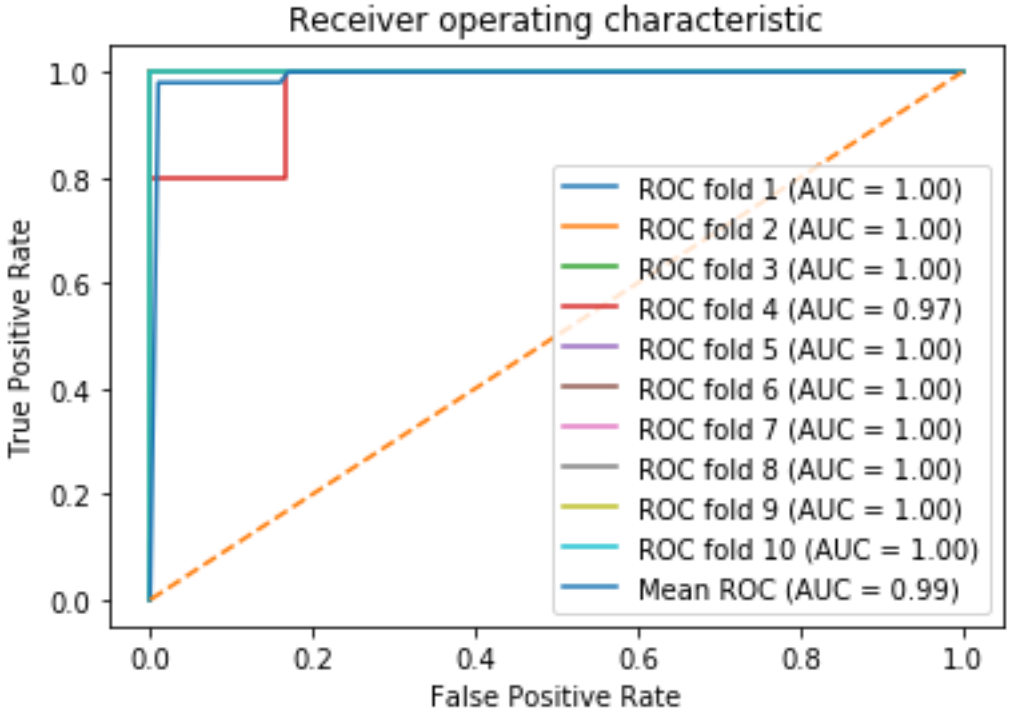
test 1.000000 0.941176 1.000000 0.982353 0.028414

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.983333 | 0.984848 | 0.992188 | 1.0 | 0.984375 | 0.992424 | 0.984848 |
| test | 0.954545 | 1.000000 | 1.000000 | 0.9 | 1.000000 | 1.000000 | 1.000000 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.984375 | 0.984375 | 0.992063 | 0.988283 | 0.005557 |
| test | 1.000000 | 0.928571 | 1.000000 | 0.978312 | 0.037215 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.990291 | 0.989691 | 0.994898 | 1.000000 | 0.989899 | 0.994792 | 0.989691 |
| test | 0.928571 | 1.000000 | 1.000000 | 0.961538 | 1.000000 | 1.000000 | 1.000000 |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.989899 | 0.989899 | 0.994949 | 0.992401 | 0.003561 |
| test | 1.000000 | 0.954545 | 1.000000 | 0.984466 | 0.026319 |




ROC-AUC curve for class 2, linear kernel

**Class 2 RBF kernel:**

the best C for fold 1 is 1024.0, and the best gamma for fold 1 is 6.103515625e-05
the best C for fold 2 is 2.0, and the best gamma for fold 2 is 0.0625
the best C for fold 3 is 1024.0, and the best gamma for fold 3 is 0.0001220703125
the best C for fold 4 is 1.0, and the best gamma for fold 4 is 0.0625
the best C for fold 5 is 64.0, and the best gamma for fold 5 is 0.00048828125
the best C for fold 6 is 4.0, and the best gamma for fold 6 is 0.015625
the best C for fold 7 is 1024.0, and the best gamma for fold 7 is 0.00048828125
the best C for fold 8 is 64.0, and the best gamma for fold 8 is 0.25
the best C for fold 9 is 1.0, and the best gamma for fold 9 is 0.125

the best C for fold 10 is 512.0, and the best gamma for fold 10 is 0.0001220703125

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.981366 | 1.000000 | 1.000000 | 0.993789 | 0.987578 | 1.000000 | 0.993789 |
| test | 1.000000 | 0.941176 | 0.941176 | 1.000000 | 1.000000 | 0.941176 | 1.000000 |

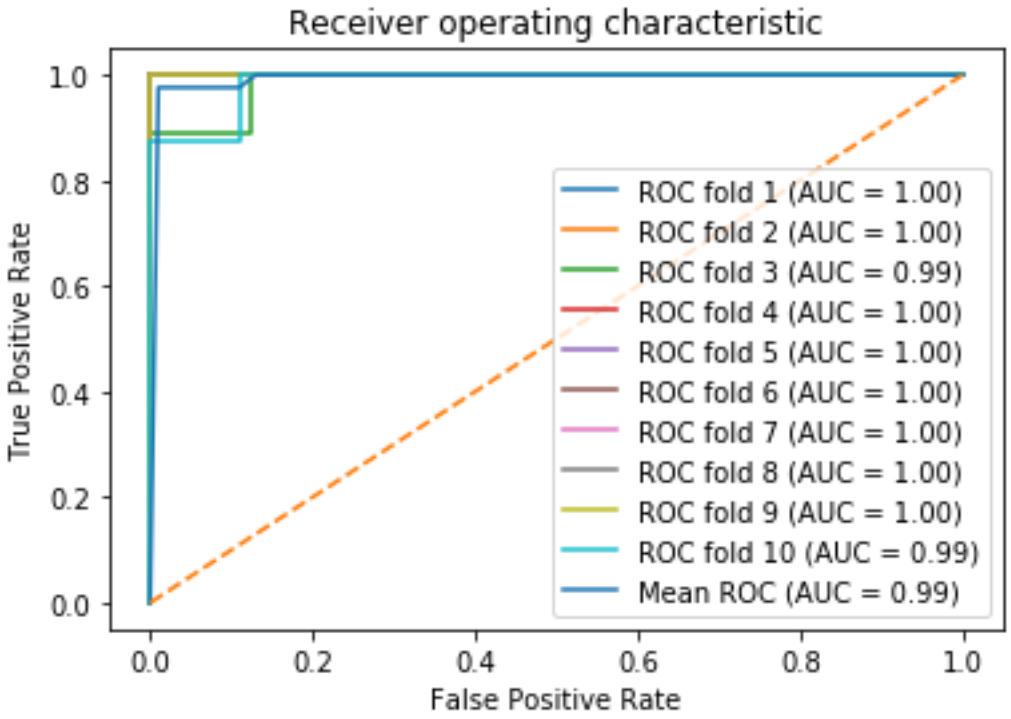
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 0.993789 | 0.995031 | 0.006415 |
| test | 1.0 | 1.0 | 0.882353 | 0.970588 | 0.041595 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.979585 | 1.000000 | 1.000000 | 0.992424 | 0.984375 | 1.000000 | 0.992063 |
| test | 1.000000 | 0.944444 | 0.944444 | 1.000000 | 1.000000 | 0.964286 | 1.000000 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 1.0 | 1.0 | 0.992063 | 0.994051 | 0.007382 |
| test | 1.0 | 1.0 | 0.881944 | 0.973512 | 0.039957 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 \ |
|---|---|---|---|---|---|---|---|---|
| train | 0.981891 | 1.000000 | 1.000000 | 0.994792 | 0.989899 | 1.000 | 0.994949 | 1.0 |
| test | 1.000000 | 0.944444 | 0.944444 | 1.000000 | 1.000000 | 0.875 | 1.000000 | 1.0 |

| | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|
| train | 1.0 | 0.994949 | 0.995648 | 0.005951 |
| test | 1.0 | 0.881944 | 0.964583 | 0.050763 |




ROC-AUC curve for class 2, RBF kernel

**Class 3 linear kernel:**

the best C for fold 1 is 128.0

the best C for fold 2 is 64.0

the best C for fold 3 is 0.0625
the best C for fold 4 is 128.0
the best C for fold 5 is 0.03125
the best C for fold 6 is 0.03125
the best C for fold 7 is 0.0625
the best C for fold 8 is 0.03125
the best C for fold 9 is 16.0
the best C for fold 10 is 128.0

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.000000 | 0.993789 | 1.000000 | 0.987578 | 0.993789 | 0.993789 |
| test | 1.0 | 0.941176 | 1.000000 | 0.941176 | 1.000000 | 1.000000 | 1.000000 |

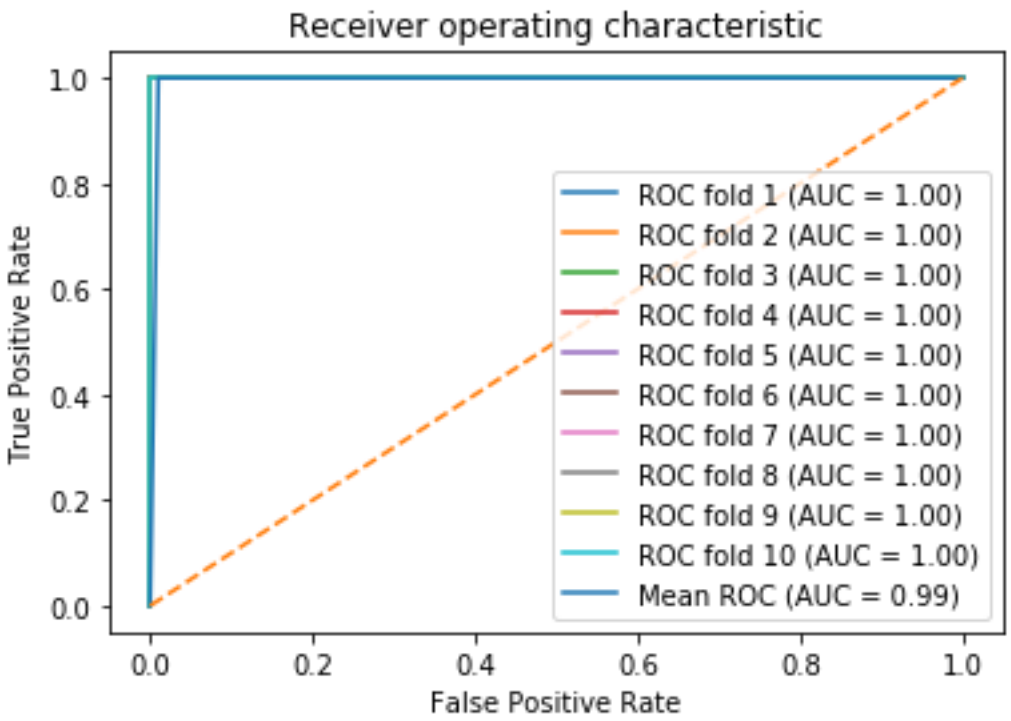
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 1.000000 | 1.000000 | 1.0 | 0.996894 | 0.004392 |
| test | 0.882353 | 0.941176 | 1.0 | 0.970588 | 0.041595 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 \ |
|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.00 | 0.995763 | 1.000000 | 0.991525 | 0.995763 | 0.99569 | 1.000000 |
| test | 1.0 | 0.75 | 1.000000 | 0.966667 | 1.000000 | 1.000000 | 1.00000 | 0.888889 |

| | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|
| train | 1.000000 | 1.0 | 0.997874 | 0.003002 |
| test | 0.961538 | 1.0 | 0.956709 | 0.080757 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 \ |
|---|---|---|---|---|---|---|---|---|
| train | 1.0 | 1.00000 | 0.988636 | 1.000000 | 0.977778 | 0.988636 | 0.98913 | 1.0 |
| test | 1.0 | 0.96875 | 1.000000 | 0.833333 | 1.000000 | 1.000000 | 1.00000 | 0.9 |

| | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|
| train | 1.0 | 1.0 | 0.994418 | 0.007879 |
| test | 0.9 | 1.0 | 0.960208 | 0.060480 |




ROC-AUC curve for class 3, linear kernel

**Class 3 RBF kernel:**

the best C for fold 1 is 128.0, and the best gamma for fold 1 is 3.0517578125e-05
the best C for fold 2 is 1.0, and the best gamma for fold 2 is 0.0078125
the best C for fold 3 is 512.0, and the best gamma for fold 3 is 3.0517578125e-05
the best C for fold 4 is 256.0, and the best gamma for fold 4 is 3.0517578125e-05
the best C for fold 5 is 2.0, and the best gamma for fold 5 is 0.001953125
the best C for fold 6 is 512.0, and the best gamma for fold 6 is 3.0517578125e-05
the best C for fold 7 is 128.0, and the best gamma for fold 7 is 6.103515625e-05
the best C for fold 8 is 256.0, and the best gamma for fold 8 is 3.0517578125e-05
the best C for fold 9 is 128.0, and the best gamma for fold 9 is 6.103515625e-05
the best C for fold 10 is 16.0, and the best gamma for fold 10 is 0.00048828125

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.993789 | 0.993789 | 1.0 | 0.993789 | 0.993789 | 0.987578 | 1.000000 |
| test | 1.000000 | 1.000000 | 1.0 | 1.000000 | 1.000000 | 1.000000 | 0.941176 |

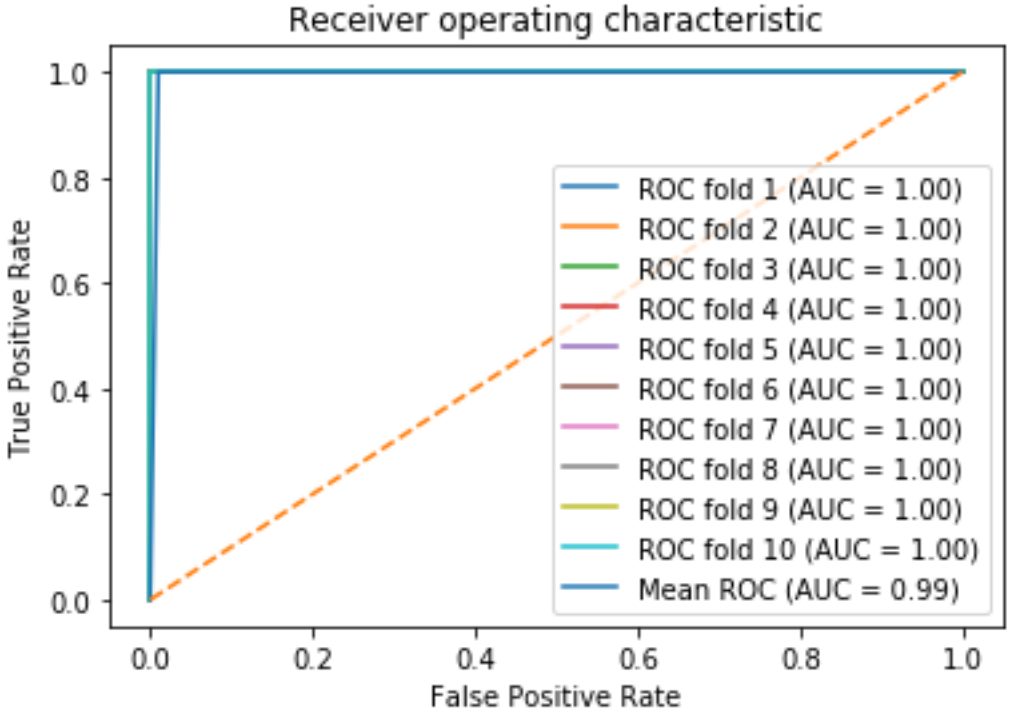
| | 8 | 9 | 10 | mean accuracy | std accuracy |
|---|---|---|---|---|---|
| train | 0.993789 | 0.993789 | 0.987578 | 0.993789 | 0.004141 |
| test | 1.000000 | 1.000000 | 0.941176 | 0.988235 | 0.024802 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 \ |
|---|---|---|---|---|---|---|---|
| train | 0.995868 | 0.995868 | 1.0 | 0.99569 | 0.99569 | 0.991304 | 1.000000 |
| test | 1.000000 | 1.000000 | 1.0 | 1.00000 | 1.00000 | 1.000000 | 0.958333 |

| | 8 | 9 | 10 | mean recall | std recall |
|---|---|---|---|---|---|
| train | 0.995763 | 0.995763 | 0.991304 | 0.995725 | 0.002900 |
| test | 1.000000 | 1.000000 | 0.966667 | 0.992500 | 0.015933 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | \ |
|---|---|---|---|---|---|---|---|---|
| train | 0.987805 | 0.987805 | 1.0 | 0.98913 | 0.98913 | 0.979167 | 1.000000 | |
| test | 1.000000 | 1.000000 | 1.0 | 1.00000 | 1.00000 | 1.000000 | 0.916667 | |

| | 8 | 9 | 10 | mean precision | std precision |
|---|---|---|---|---|---|
| train | 0.988636 | 0.988636 | 0.979167 | 0.988948 | 0.006980 |
| test | 1.000000 | 1.000000 | 0.833333 | 0.975000 | 0.056246 |




ROC-AUC curve for class 3, RBF kernel