# 1 Problem 1

## 1.1 (i)

If we let $SSE = \frac{1}{N}\sum\limits_{t=1}^{N}(r^t - (w_1x^t + w_0))^2$, then we want to minimize the SSE by setting the derivative with respect to $w_0$ and $w_1$ to zero:

$$\frac{\partial SSE}{\partial w_0} = \frac{1}{N}\sum_{t=1}^{N} 2(r^t - w_1x^t - w_0)(-1) = 0 \tag{1}$$

We can get rid of the $\frac{2}{N}$ by multiplying each side by $\frac{N}{2}$ and carry the summations through to arrive at:

$$\sum_{t=1}^{N} w_0 + \sum_{t=1}^{N} w_1x^t = \sum_{t=1}^{N} r^t \tag{2}$$

$$Nw_0 + w_1\sum_{t=1}^{N} x^t = \sum_{t=1}^{N} r^t \tag{3}$$

and similarly for $w_1$:

$$\frac{\partial SSE}{\partial w_1} = \frac{1}{N}\sum_{t=1}^{N} 2(r^t - w_1x^t - w_0)(-x^t) = 0 \tag{4}$$

we can get rid of $\frac{2}{N}$ by multiplying each side by $\frac{N}{2}$ and carry the summations through to arrive at:

$$w_0\sum_{t=1}^{N} x^t + w_1\sum_{t=1}^{N}(x^t)^2 = \sum_{t=1}^{N} x^tr^t \tag{5}$$

Now divide both sides of (3) by N:

$$w_0 + w_1\bar{x^t} = \bar{r^t} \tag{6}$$

We'll use (3) and (5) to solve a system of equations. Multiply (3) by $\sum\limits_{t=1}^{N} x^t$ and multiply (5) by $N$

$$Nw_0\sum_{t=1}^{N} x^t + w_1(\sum_{t=1}^{N} x^t)^2 = \sum_{t=1}^{N} x^t \sum_{t=1}^{N} r^t \tag{7}$$

$$Nw_0\sum_{t=1}^{N} x^t + Nw_1\sum_{t=1}^{N}(x^t)^2 = N\sum_{t=1}^{N} x^tr^t \tag{8}$$

subtract (7) and (8)

$$Nw_1\sum_{t=1}^{N}(x^t)^2 - w_1(\sum_{t=1}^{N} x^t)^2 = N\sum_{t=1}^{N} x^tr^t - \sum_{t=1}^{N} x^t\sum_{t=1}^{N} r^t \tag{9}$$

Factor out $w_1$ from the L.H.S of (9) and divide and we have:

$$w_1 = \frac{N\sum\limits_{t=1}^{N} x^tr^t - \sum\limits_{t=1}^{N} x^t\sum\limits_{t=1}^{N} r^t}{N\sum\limits_{t=1}^{N}(x^t)^2 - (\sum\limits_{t=1}^{N} x^t)^2} \tag{10}$$

And rearranging (6) to solve for $w_0$:

$$w_0 = \bar{r^t} - w_1\bar{x^t} \tag{11}$$

The solutions above for $w_0$ and $w_1$ are the optimal values.

## 1.2 (ii)

This time we let $SSE = \frac{1}{N}\sum_{t=1}^{N}(r^t - (v_2(x^t)^2 + v_1 x^t + v_0))^2$ and we want to minimize the SSE by setting the derivative with respect to $v_0$ and $v_1$ and $v_2$ to zero:

$$\frac{\partial SSE}{\partial v_0} = \frac{1}{N}\sum_{t=1}^{N} 2(r^t - v_2(x^t)^2 - v_1 x^t - v_0)(-1) = 0 \tag{12}$$

$$\frac{\partial SSE}{\partial v_1} = \frac{1}{N}\sum_{t=1}^{N} 2(r^t - v_2(x^t)^2 - v_1 x^t - v_0)(-x) = 0 \tag{13}$$

$$\frac{\partial SSE}{\partial v_1} = \frac{1}{N}\sum_{t=1}^{N} 2(r^t - v_2(x^t)^2 - v_1 x^t - v_0)(-x^2) = 0 \tag{14}$$

We apply the same techniques as for the previous problem, pull the 2 out, multiple by $\frac{N}{2}$ and distribute the summations:

$$N v_0 + v_1 \sum_{t=1}^{N} x^t + v_2 \sum_{t=1}^{N}(x^t)^2 = \sum_{t=1}^{N} r^t \tag{15}$$

$$v_0 \sum_{t=1}^{N} x^t + v_1 \sum_{t=1}^{N}(x^t)^2 + v_2 \sum_{t=1}^{N}(x^t)^3 = \sum_{t=1}^{N} x^t r^t \tag{16}$$

$$v_0 \sum_{t=1}^{N}(x^t)^2 + v_1 \sum_{t=1}^{N}(x^t)^3 + v_2 \sum_{t=1}^{N}(x^t)^4 = \sum_{t=1}^{N}(x^t)^2 r^t \tag{17}$$

We get a linear system of three equations and three unknowns. We build a matrix from the linear system:

$$\begin{bmatrix} N & \sum_{t=1}^{N} x^t & \sum_{t=1}^{N}(x^t)^2 \\ \sum_{t=1}^{N} x^t & \sum_{t=1}^{N}(x^t)^2 & \sum_{t=1}^{N}(x^t)^3 \\ \sum_{t=1}^{N}(x^t)^2 & \sum_{t=1}^{N}(x^t)^3 & \sum_{t=1}^{N}(x^t)^4 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{N} r^t \\ \sum_{t=1}^{N} x^t r^t \\ \sum_{t=1}^{N}(x^t)^2 r^t \end{bmatrix}$$

At this point, directly solving for $v_0, v_1, v_2$ is best left to a computer. Given the data, the sums can be calculated, fed into the matrix and the parameters calculated by solving the system for $\vec{v}$

## 1.3 (iii)

Q: Professor Gopher claims that $E(v_2^*, v_1^*, v_0^* | Z_{train}) \leq E(w_1^*, w_0^* | Z_{train})$ is true for any $Z_{train}$, is Professor Gopher's claim correct?
A: Yes, Professor Gopher's claim is correct. For any given $Z_{train}$, adding more polynomial terms will give a better fit for the data and, hence, a lower error for the model. The difference in error will be exactly equal if the true model for $Z_{train}$ is linear and less if the true model for $Z_{train}$ is non-linear.

## 1.4 (iv)

Q: Professor Gopher claims that $E(v_2^*, v_1^*, v_0^* | Z_{test}) \leq E(w_1^*, w_0^* | Z_{test})$ is true for any $Z_{train}$, is Professor Gopher's claim correct?
A: Strictly speaking, Professor Gopher's claim is not true. The model is fit on $Z_{train}$ and it is possible that the true models could differ between $Z_{train}$ and $Z_{test}$. For instance, if $Z_{train}$'s data had a non-linear relationship and $Z_{test}$'s model had a linear relationship, the quadratic model could be overfit on $Z_{train}$ and not generalize well to $Z_{test}$. Generally speaking this should only happen when there isn't a lot of data or there is some disproportionate representative values in one set and not the other. These issues can be rectified by gathering more data and ensuring the data is randomly split between $Z_{train}$ and $Z_{test}$. Ensuring more data and randomly splitting data sets makes Professor Gopher's claims more believable.

# 2 Problem 2

## 2.1 (i)

tr($A$): 701
tr($A^T$): 701
tr($A^TA$): 484533
tr($AA^T$): 484533

## 2.2 (ii)

The absolute value of $|A|$ represents the "volume" of the 5 dimensional "parallelogram" formed by the rows of A.

## 2.3 (iii)

Because the columns of A are clearly not linearly dependent, $detA \neq 0$. The columns are not linearly dependent because there is no way to form a linear combination of one column from the other columns. I know this because observing the relationship between the columns I note that any column can be created based on element-wise multiplication of two other columns. For instance, column 4 can be created by multiplying the elements of columns 2 and 3. Because of this, there is no way to multiply a column by a scalar to create any other column. Note that the exception to "every column is an element-wise multiplication of two other columns" is column one, but no other column is a linear combination of that and vice-versa.

## 2.4 (iv)

As stated in the previous problem, because the columns of A are clearly not linearly dependent, $A$ is full rank, with rank 5.

# 3 Problem 3

## 3.1 (i) my_cross_val()

method: LinearSVC
dataset: Boston50

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1765 | 0.3333 | 0.1765 | 0.3137 | 0.4314 | 0.1961 | 0.2200 | 0.2800 | 0.1400 | 0.1600 | 0.2427 | 0.0889 |

method: LinearSVC
dataset: Boston75

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.2157 | 0.1373 | 0.1176 | 0.2549 | 0.1765 | 0.6863 | 0.0600 | 0.1600 | 0.1200 | 0.1000 | 0.2028 | 0.1699 |

method: LinearSVC
dataset: Digits

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0611 | 0.0667 | 0.0444 | 0.0778 | 0.0222 | 0.0444 | 0.0556 | 0.0670 | 0.0559 | 0.0559 | 0.0551 | 0.0146 |

method: SVC
dataset: Boston50

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.2353 | 0.4118 | 0.3725 | 0.2745 | 0.3333 | 0.2549 | 0.3400 | 0.3400 | 0.2400 | 0.4400 | 0.3242 | 0.0680 |

method: SVC
dataset: Boston75

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.2157 | 0.1373 | 0.3137 | 0.3333 | 0.2549 | 0.1961 | 0.2200 | 0.2200 | 0.3000 | 0.2600 | 0.2451 | 0.0566 |

method: SVC
dataset: Digits

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.4611 | 0.6222 | 0.3833 | 0.4222 | 0.4944 | 0.4278 | 0.5389 | 0.4637 | 0.5196 | 0.5140 | 0.4847 | 0.0651 |

method: LogisticRegression
dataset: Boston50

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0784 | 0.0784 | 0.0784 | 0.1176 | 0.1373 | 0.1961 | 0.1400 | 0.2000 | 0.1400 | 0.1200 | 0.1286 | 0.0421 |

method: LogisticRegression
dataset: Boston75

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0588 | 0.0588 | 0.0588 | 0.1569 | 0.0980 | 0.0980 | 0.1200 | 0.1000 | 0.0800 | 0.1200 | 0.0949 | 0.0304 |

method: LogisticRegression
dataset: Digits

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0278 | 0.0444 | 0.0222 | 0.0333 | 0.0389 | 0.0333 | 0.0556 | 0.0503 | 0.0279 | 0.0447 | 0.0378 | 0.0102 |

### 3.2 (ii) my_train_test()

method: LinearSVC
dataset: Boston50

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.2540 | 0.1349 | 0.2460 | 0.1429 | 0.2540 | 0.1508 | 0.2937 | 0.2222 | 0.1032 | 0.3492 | 0.2151 | 0.0752 |

method: LinearSVC
dataset: Boston75

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0556 | 0.0714 | 0.5794 | 0.1825 | 0.1984 | 0.2222 | 0.5079 | 0.1349 | 0.3571 | 0.5079 | 0.2817 | 0.1827 |

method: LinearSVC
dataset: Digits

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0334 | 0.0468 | 0.0445 | 0.0713 | 0.0334 | 0.0601 | 0.0624 | 0.0490 | 0.0646 | 0.0423 | 0.0508 | 0.0125 |

method: SVC
dataset: Boston50

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.3492 | 0.2698 | 0.3651 | 0.4286 | 0.3730 | 0.3730 | 0.4524 | 0.3889 | 0.3413 | 0.3175 | 0.3659 | 0.0495 |

method: SVC
dataset: Boston75

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.2222 | 0.2381 | 0.2778 | 0.2302 | 0.2222 | 0.2381 | 0.2302 | 0.2778 | 0.2619 | 0.1667 | 0.2365 | 0.0307 |

method: SVC
dataset: Digits

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.7082 | 0.4365 | 0.6125 | 0.5301 | 0.6481 | 0.6882 | 0.5991 | 0.5657 | 0.7082 | 0.4811 | 0.5978 | 0.0898 |

method: LogisticRegression
dataset: Boston50

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1667 | 0.1667 | 0.1746 | 0.1746 | 0.1032 | 0.1984 | 0.1667 | 0.1349 | 0.1429 | 0.1508 | 0.1579 | 0.0250 |

method: LogisticRegression
dataset: Boston75

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.1270 | 0.1111 | 0.0873 | 0.0714 | 0.1032 | 0.1746 | 0.1032 | 0.0873 | 0.1032 | 0.1270 | 0.1095 | 0.0272 |

method: LogisticRegression
dataset: Digits

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0423 | 0.0356 | 0.0535 | 0.0401 | 0.0512 | 0.0223 | 0.0267 | 0.0423 | 0.0223 | 0.0535 | 0.0390 | 0.0115 |

# 4 Problem 4

method: LinearSVC
dataset: X1

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0611 | 0.0778 | 0.0778 | 0.0889 | 0.1056 | 0.0722 | 0.1056 | 0.1006 | 0.1173 | 0.0615 | 0.0868 | 0.0187 |

method: LinearSVC
dataset: X2

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0056 | 0.0222 | 0.0111 | 0.0167 | 0.0167 | 0.0111 | 0.0056 | 0.0112 | 0.0000 | 0.0056 | 0.0106 | 0.0063 |

method: LinearSVC
dataset: X3

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0111 | 0.0667 | 0.0444 | 0.0389 | 0.0500 | 0.0444 | 0.0722 | 0.0168 | 0.0503 | 0.0838 | 0.0479 | 0.0216 |

method: SVC
dataset: X1

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.9500 | 0.9389 | 0.9222 | 0.9278 | 0.9222 | 0.9111 | 0.9167 | 0.9218 | 0.9106 | 0.9162 | 0.9237 | 0.0118 |

method: SVC
dataset: X2

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.9222 | 0.9278 | 0.9167 | 0.9333 | 0.9500 | 0.9556 | 0.9278 | 0.9330 | 0.9274 | 0.9385 | 0.9332 | 0.0114 |

method: SVC
dataset: X3

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.9611 | 0.9111 | 0.9167 | 0.9222 | 0.9222 | 0.9222 | 0.9278 | 0.9330 | 0.9218 | 0.9274 | 0.9265 | 0.0129 |

method: LogisticRegression
dataset: X1

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0500 | 0.0389 | 0.0556 | 0.0556 | 0.0556 | 0.0444 | 0.0444 | 0.0503 | 0.0726 | 0.0279 | 0.0495 | 0.0113 |

method: LogisticRegression
dataset: X2

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0111 | 0.0056 | 0.0056 | 0.0167 | 0.0056 | 0.0278 | 0.0222 | 0.0000 | 0.0056 | 0.0000 | 0.0100 | 0.0089 |

method: LogisticRegression
dataset: X3

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0778 | 0.0333 | 0.0500 | 0.0444 | 0.0667 | 0.0500 | 0.0389 | 0.1061 | 0.0559 | 0.0559 | 0.0579 | 0.0202 |

===================== ORIGINAL DATA =========================

method: LinearSVC
dataset: X1

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0615 | 0.0950 | 0.0838 | 0.0950 | 0.0670 | 0.1006 | 0.0782 | 0.0726 | 0.0615 | 0.0726 | 0.0788 | 0.0136 |

method: LinearSVC
dataset: X2

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0168 | 0.0168 | 0.0112 | 0.0112 | 0.0056 | 0.0112 | 0.0056 | 0.0056 | 0.0000 | 0.0112 | 0.0095 | 0.0050 |

method: LinearSVC
dataset: X3

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0615 | 0.0559 | 0.0503 | 0.0559 | 0.0503 | 0.0503 | 0.0670 | 0.0335 | 0.0559 | 0.0559 | 0.0536 | 0.0084 |

method: SVC
dataset: X1

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.8994 | 0.9106 | 0.9385 | 0.9218 | 0.9497 | 0.9553 | 0.9274 | 0.9274 | 0.9497 | 0.9050 | 0.9285 | 0.0187 |

method: SVC
dataset: X2

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.9330 | 0.9274 | 0.9274 | 0.9274 | 0.9441 | 0.9274 | 0.9218 | 0.9050 | 0.9330 | 0.9330 | 0.9279 | 0.0095 |

method: SVC
dataset: X3

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.9274 | 0.9385 | 0.9162 | 0.9050 | 0.9106 | 0.9162 | 0.9385 | 0.9441 | 0.9218 | 0.9274 | 0.9246 | 0.0123 |

method: LogisticRegression
dataset: X1

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0894 | 0.0391 | 0.0615 | 0.0503 | 0.1061 | 0.0782 | 0.0559 | 0.0670 | 0.0726 | 0.0670 | 0.0687 | 0.0184 |

method: LogisticRegression
dataset: X2

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0056 | 0.0056 | 0.0223 | 0.0056 | 0.0112 | 0.0000 | 0.0112 | 0.0335 | 0.0112 | 0.0056 | 0.0112 | 0.0093 |

method: LogisticRegression
dataset: X3

| Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 | mean | std dev |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.0615 | 0.0559 | 0.0279 | 0.0335 | 0.0335 | 0.0670 | 0.0391 | 0.0782 | 0.0447 | 0.0447 | 0.0486 | 0.0156 |