Quiz 6

Problem 1

1/1 point (graded)

You are given a binary 4-dimensional linear decision boundary with coefficient vector $\mathbf{w} = [2, 1, 4, 3]$ and $b = -12$. How would you classify the point $(2, 1, 1, 2)$?

- ○ -1
- ○ 0
- ◉ 1

✔

Submit

Problem 2

1/1 point (graded)

In which of the following situations has our linear classifier correctly labeled a data point? Select all that apply.

- ☑ $\mathbf{w} \cdot \mathbf{x} + b > 0$ and $y > 0$
- ☑ $y\,(\mathbf{w} \cdot \mathbf{x} + b) > 0$

Generating Speech Output

- [x] $\mathbf{w} \cdot \mathbf{x} + b < 0$ and $y < 0$
- [ ] $y > \mathbf{w} \cdot \mathbf{x} + b$

Submit

Problem 3

1/1 point (graded)

Let's say that we have a linear classifier given by $\mathbf{w} = [1, 1, -3, 0]$ and $b = -2$. Our loss function measures the amount by which our prediction is incorrect: $\text{loss} = -y\,(\mathbf{w} \cdot \mathbf{x} + b)$. If our prediction is correct, there is no loss.

What is the loss on the data point $(\mathbf{x}, y)$ where $\mathbf{x} = (3, 1, 1, 4)$ and $y = 1$?

- ( ) 0
- (x) 1
- ( ) 2
- ( ) 3

Submit

Problem 4

1/1 point (graded)

If the Perceptron algorithm does 9 updates before converging on a solution, what value of

Generating Speech Output ave?

- [ ] $b = 9$
- [ ] $b = -9$
- [x] $b \in [-9, 9]$
- [ ] $b \in [0, 9]$

✓

Submit

Problem 5

1/1 point (graded)

What is a support vector?

- [ ] A data point from the test set that is used to test the classifier
- [ ] A vector that we are trying to minimize
- [ ] A data point which is correctly classified by the optimal solution for $\mathbf{w}$
- [x] A data point from the training set that contributes to the optimal solution for $\mathbf{w}$

✓

Submit

Problem 6

1/1 point (graded)

Generating Speech Output between the perceptron algorithm and the support vector machine?

- ( ) The perceptron uses gradient descent while the SVM uses stochastic gradient descent
- ( ) The perceptron finds a linear separator that separates most of the data points in the training set, while a SVM finds a linear separator that separates all of the data in the training set
- (•) The perceptron finds any solution that perfectly separates the training set, while the SVM finds the solution that perfectly separates the training set with the greatest margin of separation
- ( ) The perceptron algorithm may not find a solution while the SVM is guaranteed to find a solution

✓

Submit

Problem 7

1/1 point (graded)

The optimal solution for a SVM is given by the coefficient vector $\mathbf{w}$ and the constant b. The width of the margin is given by γ. What is the value of γ?

- (•) $\gamma = \frac{1}{||\mathbf{w}||}$
- ( ) $\gamma = \frac{1}{||\mathbf{w}||}$
- ( ) $\gamma = ||\mathbf{w}||$
- ( ) $\gamma = b - \frac{1}{||\mathbf{w}||}$

✓

Generating Speech Output

Submit

Problem 8

1/1 point (graded)

True or false: A soft-margin SVM has fewer support vectors than a hard-margin SVM.

- ( ) True
- (x) False

✔

Submit

Problem 9

1/1 point (graded)

Decreasing the value of C in the soft-margin SVM results in which of the following:

- [ ] fewer number of support vectors
- [x] wider margin
- [ ] more data points being correctly classified
- [x] lower penalty for incorrectly classified data points

✔

Submit

Problem 10

1/1 point (graded)

True or false: All support vectors are contained between, or on, the margins of the two classes.

- ( ) True
- (x) False

✔

Submit

Problem 11

1/1 point (graded)

What does the slack variable represent?

- (x) It is a vector containing the amount of error each point $(x^{(i)}, y^{(i)})$ contributes to the optimization problem
- ( ) It is a coefficient that we must determine to optimize the problem
- ( ) It is a vector containing the number of times each w_i is updated
- ( ) it is a value that determines how much error the optimization problem is allowed to have

✔

Submit

Generating Speech Output

1/1 point (graded)

Using the dual form of the perceptron algorithm, which of the following values are updated during each pass over the training set?

- [ ] $\mathbf{w}$
- [x] α
- [x] b
- [ ] $\mathbf{y}$

✔

Submit

Problem 13

1/1 point (graded)

When optimizing the dual form of the hard-margin SVM, when are the values α_i non-zero?

- ◯ When the data point $(x^{(i)}, y^{(i)})$ is on the linear separator between the two classes
- ◉ When the data point $(x^{(i)}, y^{(i)})$ is right on the margin for its class
- ◯ When the data point $(x^{(i)}, y^{(i)})$ is in the interior of the region for its class
- ◯ When the data point $(x^{(i)}, y^{(i)})$ is on the wrong side of the linear separator

✔

Submit

Generating Speech Output

Problem 14

1/1 point (graded)

When using multiclass logistic regression on data with labels, $Y = \{1, 2, \ldots, k\}$, and a linear classifier specified by $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k \in \mathbb{R}^d$ and $b_1, b_2, \ldots, b_k \in \mathbb{R}$, and given a point $(\mathbf{x}, y)$, what is the probability that $y = j$, where $0 < j \leq k$?

- ( ) $Pr\left(y = j|\mathbf{x}\right) = e^{\mathbf{w}_j \cdot \mathbf{x} + b_j}$
- ( ) $Pr\left(y = j|\mathbf{x}\right) = \frac{e^{\mathbf{w}_j \cdot \mathbf{x} + b_j}}{e^{\mathbf{w}_k \cdot \mathbf{x} + b_k}}$
- (•) $Pr\left(y = j|\mathbf{x}\right) = \frac{e^{\mathbf{w}_j \cdot \mathbf{x} + b_j}}{e^{\mathbf{w}_1 \cdot \mathbf{x} + b_1} + e^{\mathbf{w}_2 \cdot \mathbf{x} + b_2} + \ldots + e^{\mathbf{w}_k \cdot \mathbf{x} + b_k}}$
- ( ) $Pr\left(y = j|\mathbf{x}\right) = \frac{e^{\mathbf{w}_j \cdot \mathbf{x} + b_j}}{1 + e^{\mathbf{w}_j \cdot \mathbf{x} + b_j}}$

✔

Submit

Problem 15

1/1 point (graded)

What does ξ_i represent in the soft-margin SVM?

- ( ) It is the number of times the i'th point was updated
- (•) It is the amount of slack the i'th point has
- ( ) It represents the i'th support vector
- ( ) It represents the width of the margin

✔

Generating Speech Output

Submit

Generating Speech Output