

Linear Regression Questions (Make Sure to watch the videos 1st)

Quiz Question

Which of the following are related to bias in machine learning

- ☐ Images, text, video, and speech are all types of data that can contain bias
- ☐ Since the data we used to train the model can contain bias, that could be reflected in the model.
- ☐ Statistical validation can offset the bias reflected in the real-world validation of your models.

Submit

Write A, B, or C here (you can choose more than one answer if there is): _ _ _ A,B_ _ _ _

Let's say that we have a line whose equation is $y = -0.6x + 4$. For the point $(x,y) = (-5, 3)$, apply the **absolute trick** to get the new equation for the line, using a learning rate of $\alpha = 0.1$.

Report your answer in the form $y = w_1x + w_2$, substituting appropriate values for w_1 and w_2 .

Enter your response here: $-0.6(-5) + 4 = 7$

$$Y^{\wedge} - Y = 7 - 3 = 4$$

$$-0.6 - (0.1)(1)(-5) = -0.1$$

$$4 - (0.1)(1) = 3.9$$

$$-0.1x + 3.9$$

Let's say that we have a line whose equation is $y = -0.6x + 4$. For the point $(x,y) = (-5, 3)$, apply the **square trick** to get the new equation for the line, using a learning rate of $\alpha = 0.01$.

Report your answer in the form $y = w_1x + w_2$, substituting appropriate values for w_1 and w_2 .

Enter your response here: $-0.6 - (0.01)(4)(-5) = -0.4$

$$4 - (0.01)(4) = 3.96 \rightarrow -0.4x + 3.96$$

Quiz Question

Which of the following are true about gradient descent?

- ☐ Gradient descent is a strategy that helps minimize the error between points of the actual data and the "best-fit line"
- ☐ Gradient descent is a strategy that helps isolate outliers in the data.
- ☐ We use gradient descent to update the parameters of our model as we train

Submit

Write A, B, or C here (you can choose more than one answer if there is): _ _ _ A,C_ _ _ _

Quiz Question

Which of the following are accurate statements about 'mean absolute error'?

- ☐ It is the sum of all the errors divided by m
- ☐ It is the average of all points above the line
- ☐ It is the average error of all points

Submit

Write A, B, or C here (you can choose more than one answer if there is): _ _ _ C_ _ _ _

Compute the **mean absolute error** for the following line and points:

- line: $y = 1.2x + 2$
- points: (2, -2), (5, 6), (-4, -4), (-7, 1), (8, 14)

Enter your response here: $m = 5$

$Y1 = -2 - 4.4 = 6.4$, $Y2 = 6 - 8 = 2$, $Y3 = -4 - (-2.8) = 1.2$, $Y4 = 1 - (-7.4) = 8.4$, $Y5 = 14 - (-11.6) = 2.4$

MAE: $6.4 + 2 + 1.2 + 8.4 + 2.4 / 5 = 4.08$

Quiz for Mean Squared Error

Compute the **mean squared error** for the following line and points:

- line: $y = 1.2x + 2$
- points: (2, -2), (5, 6), (-4, -4), (-7, 1), (8, 14)

Enter your response here: $m=5$

$$y_1 = (-2 - 4.4) = 6.4 = 40.96$$

$$y_2 = (6 - 8) = 2 = 4$$

$$y_3 = (-4 - (-2.8)) = 1.2 = 1.44$$

$$y_4 = (1 - (-7.4)) = 8.4 = 70.56$$

$$y_5 = (14 - (-11.6)) = 2.4 = 5.76$$

$$\text{MSE} = (40.96 + 4 + 1.44 + 70.56 + 5.76) / 5 = 24.544$$

Quiz Question

There are 2 major ways to fit a line in machine learning. Which of the following are ways to fit a line?

- ☐ Minimize the error function using mean-squared or mean-absolute
- ☐ Try every possible position and slope of the line by hand until you get it right
- ☐ Using any of the tricks such as the absolute and the square trick

Write A, B, or C here (you can choose more than one answer if there is): _ _ _ _ A, _ _ _

Quiz Question

Which of the following could be possible new dimensions for the house price dataset

- ☐ Number of bedrooms
- ☐ Age of house
- ☐ Opinions of neighbors
- ☐ Distance to shopping
- ☐ Types of restaurants nearby

Submit

Write A, B, C, D, or E here (you can choose more than one answer if there is): _ _ _ A,B _ _ _ _