

## ✔ Congratulations! You passed!

Go to next item

Grade received 100% Latest Submission Grade 100% To pass 80% or higher

1. Which of the following can address overfitting?

1 / 1 point

☒ Apply regularization

✔ Correct

Regularization is used to reduce overfitting.

☒ Select a subset of the more relevant features.

✔ Correct

If the model trains on the more relevant features, and not on the less useful features, it may generalize better to new examples.

☐ Remove a random set of training examples

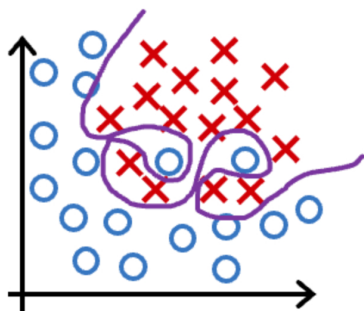
☒ Collect more training data

✔ Correct

If the model trains on more data, it may generalize better to new examples.

2. You fit logistic regression with polynomial features to a dataset, and your model looks like this.

1 / 1 point



What would you conclude? (Pick one)

- ☐ The model has high bias (underfit). Thus, adding data is likely to help
- ☒ The model has high variance (overfit). Thus, adding data is likely to help
- ☐ The model has high variance (overfit). Thus, adding data is, by itself, unlikely to help much.
- ☐ The model has high bias (underfit). Thus, adding data is, by itself, unlikely to help much.

✔ Correct

The model has high variance (it overfits the training data). Adding data (more training examples) can help.

3. Regularization

1 / 1 point

$$\min_{\vec{w}} J(\vec{w}; h) = \min_{\vec{w}} \left[ \underbrace{\frac{1}{m} \sum_{i=1}^m (f_{\vec{w}}(\vec{x}^{(i)}) - y^{(i)})^2}_{\text{mean squared error}} + \underbrace{\frac{\lambda}{2} \sum_{j=1}^n w_j^2}_{\text{regularization term}} \right]$$

$$\vec{w}, b = \underset{\text{argmin}}{\vec{w}, b} \left[ 2m \sum_{i=1}^n (w_0 x_i + b - y_i)^2 + \lambda \sum_{j=1}^n w_j^2 \right]$$

Suppose you have a regularized linear regression model. If you increase the regularization parameter  $\lambda$ , what do you expect to happen to the parameters  $w_1, w_2, \dots, w_n$ ?

- ☐ This will increase the size of the parameters  $w_1, w_2, \dots, w_n$
- ☒ This will reduce the size of the parameters  $w_1, w_2, \dots, w_n$

✓ **Correct**

Regularization reduces overfitting by reducing the size of the parameters  $w_1, w_2, \dots, w_n$ .