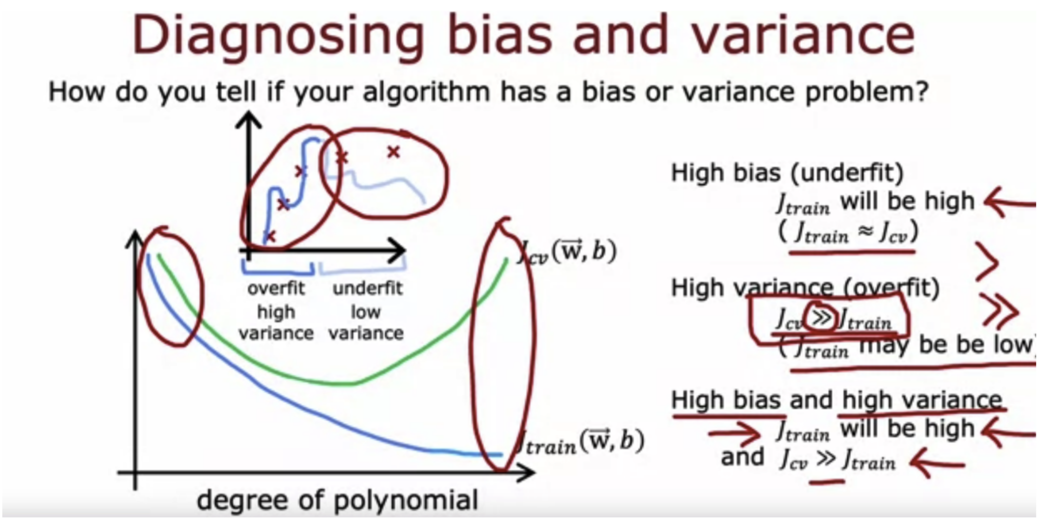
Bias and variance

✔ **Congratulations! You passed!**

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

Go to next item

1. 1 / 1 point



If the model's cross validation error J_{cv} is much higher than the training error J_{train}, this is an indication that the model has…

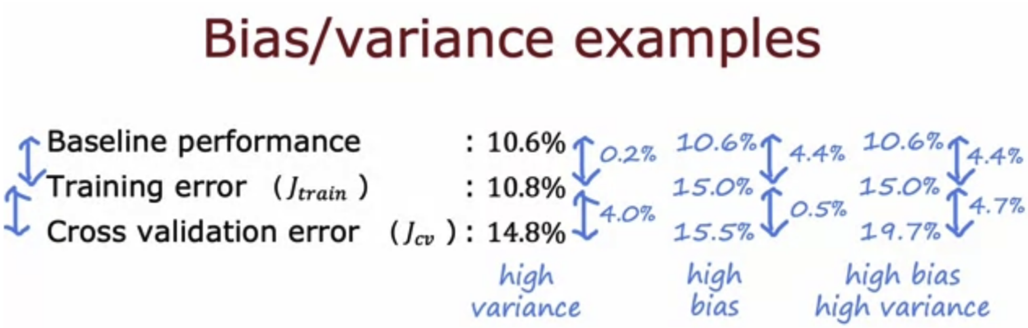
- ◯ Low variance
- ◉ high variance
- ◯ high bias
- ◯ Low bias

Correct

When $J_{cv} >> J_{train}$ (whether J_{train} is also high or not, this is a sign that the model is overfitting to the training data and performing much worse on new examples.

2.

1 / 1 point



Which of these is the best way to determine whether your model has high bias (has underfit the training data)?

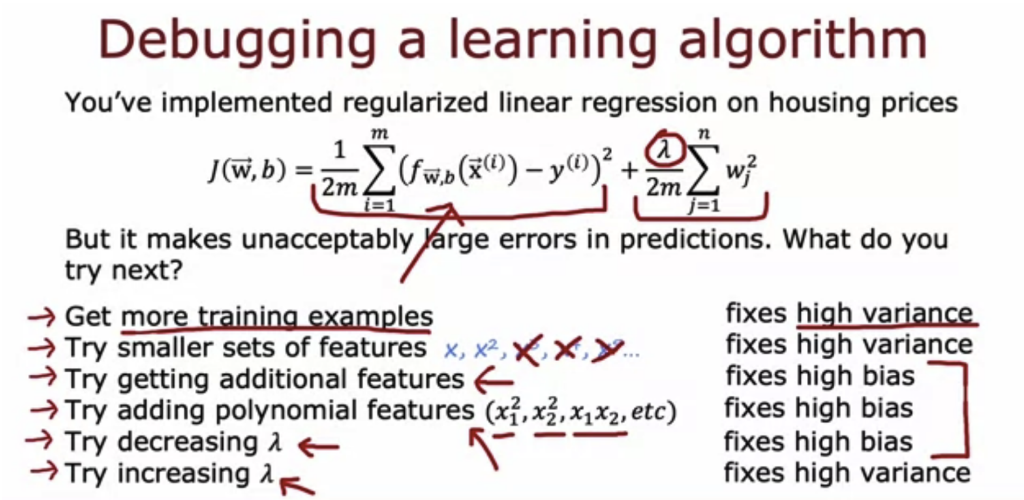
- (●) Compare the training error to the baseline level of performance
- ( ) See if the cross validation error is high compared to the baseline level of performance
- ( ) Compare the training error to the cross validation error.
- ( ) See if the training error is high (above 15% or so)

Correct

Correct. If comparing your model's training error to a baseline level of performance (such as human level performance, or performance of other well-established models), if your model's training error is much higher, then this is a sign that the model has high bias (has underfit).

3.

1 / 1 point



You find that your algorithm has high bias. Which of these seem like good options for improving the algorithm's performance? Hint: two of these are correct.

- [x] Decrease the regularization parameter λ (lambda)

Correct
Correct. Decreasing regularization can help the model better fit the training data.

- [ ] Remove examples from the training set
- [ ] Collect more training examples
- [x] Collect additional features or add polynomial features

Correct
Correct. More features could potentially help the model better fit the training examples.

4.

1 / 1 point

You find that your algorithm has a training error of 2%, and a cross validation error of 20% (much higher than the training error). Based on the conclusion you would draw about whether the algorithm has a high bias or high variance problem, which of these seem like good options for improving the algorithm's performance? Hint: two of these are correct.

- [x] Increase the regularization parameter λ

Correct
Yes, the model appears to have high variance (overfit), and increasing regularization would help reduce high variance.

- [ ] Decrease the regularization parameter λ
- [ ] Reduce the training set size
- [x] Collect more training data

Correct
Yes, the model appears to have high variance (overfit), and collecting more training examples would help reduce high variance.