

1 point

1. Consider four classifiers, whose classification performance is given by the following table:

	Classification error on training set	Classification error on validation set
Classifier 1	0.2	0.6
Classifier 2	0.8	0.6
Classifier 3	0.2	0.2
Classifier 4	0.5	0.4

Which of the four classifiers is most likely overfit?

- ☒ Classifier 1
- ☐ Classifier 2
- ☐ Classifier 3
- ☐ Classifier 4

1 point

2. Suppose a classifier classifies 23100 examples correctly and 1900 examples incorrectly. Compute accuracy by hand. Round your answer to 3 decimal places.

0.924

1 point

3. (True/False) Accuracy and error measured on the same dataset always sum to 1.

- ☒ True
- ☐ False

1 point

4. Which of the following is NOT a correct description of complex models?

- ☐ Complex models accommodate many features.
- ☐ Complex models tend to produce lower training error than simple models.
- ☒ Complex models tend to generalize better than simple models.
- ☐ Complex models tend to exhibit high variance in response to perturbation in the training data.
- ☐ Complex models tend to exhibit low bias, capturing many patterns in the training data that simple models may have missed.

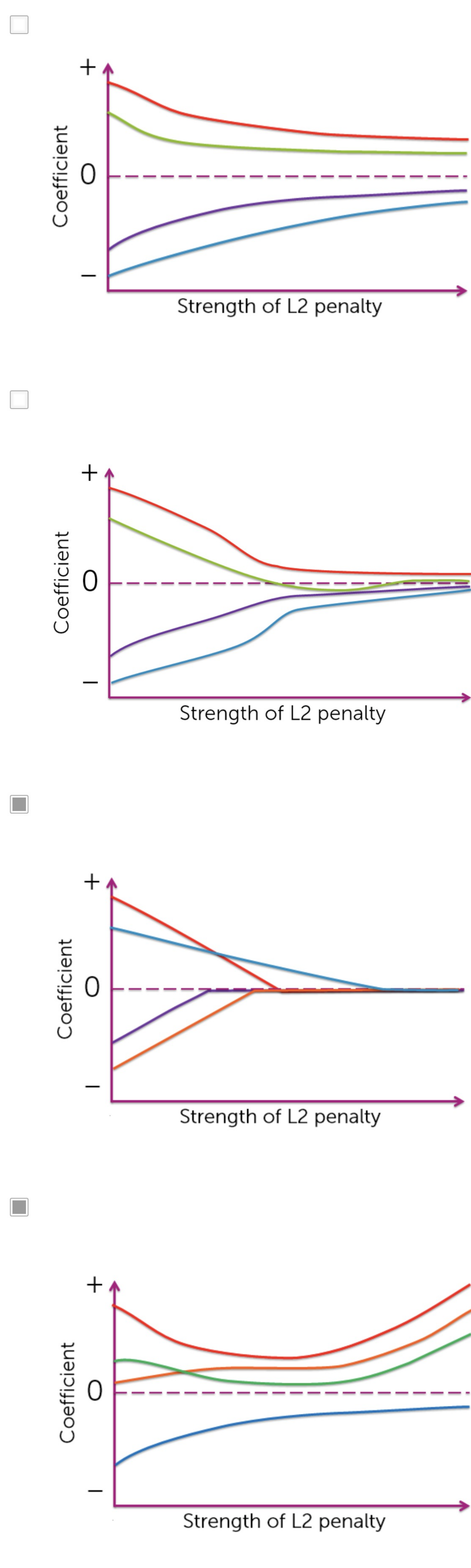
1 point

5. Which of the following is a symptom of overfitting in the context of logistic regression? Select all that apply.

- ☒ Large estimated coefficients
- ☐ Good generalization to previously unseen data
- ☐ Simple decision boundary
- ☒ Complex decision boundary
- ☒ Overconfident predictions of class probabilities

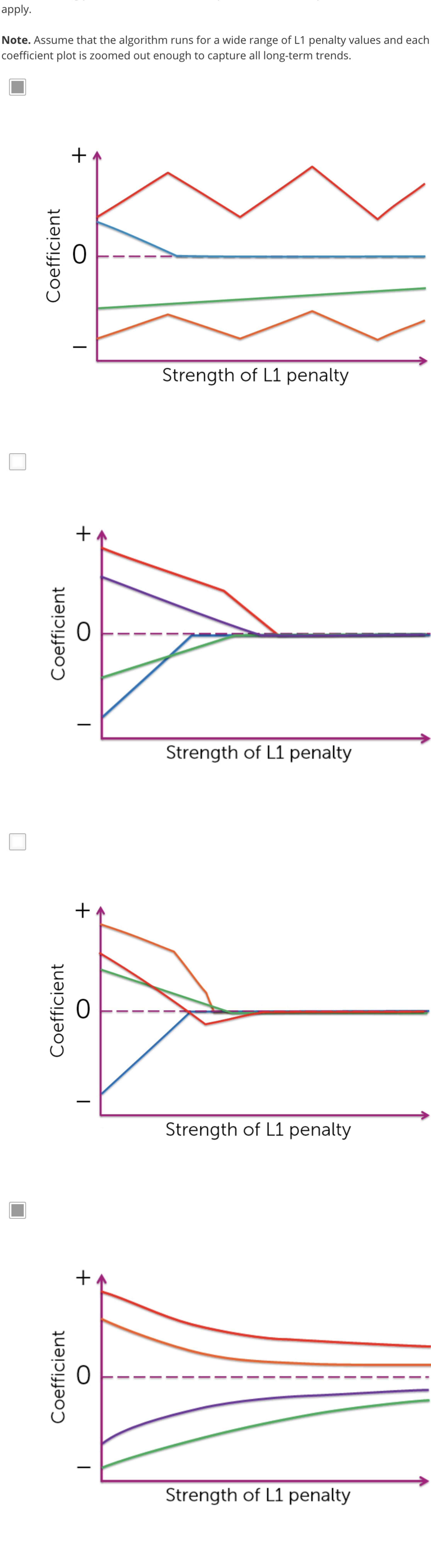
1 point

6. Suppose we perform L2 regularized logistic regression to fit a sentiment classifier. Which of the following plots does NOT describe a possible coefficient path? Choose all that apply.



1 point

7. Suppose we perform L1 regularized logistic regression to fit a sentiment classifier. Which of the following plots does NOT describe a possible coefficient path? Choose all that apply.



1 point

8. In the context of L2 regularized logistic regression, which of the following occurs as we increase the L2 penalty λ ? Choose all that apply.

- ☒ The L2 norm of the set of coefficients gets smaller
- ☐ Region of uncertainty becomes narrower, i.e., the classifier makes predictions with higher confidence.
- ☒ Decision boundary becomes less complex
- ☐ Training error decreases
- ☒ The classifier has lower variance
- ☐ Some features are excluded from the classifier