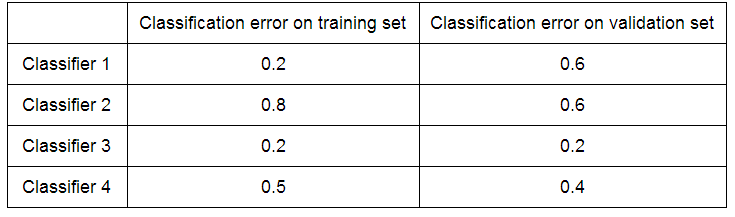
1. Consider four classifiers, whose classification performance is given by the following table:



Which of the four classifiers is most likely overfit?

**Classifier 1**

Classifier 2

Classifier 3

Classifier 4

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

0.076

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

**True**

False

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.

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**

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.

**Note.** Assume that the algorithm runs for a wide range of L2 penalty values and each coefficient plot is zoomed out enough to capture all long-term trends.

C,D

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.

**Note.** Assume that the algorithm runs for a wide range of L1 penalty values and each coefficient plot is zoomed out enough to capture all long-term trends.

D

A,D

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