

Machine Learning System Design

LATEST SUBMISSION GRADE

100%

1. You are working on a spam classification system using regularized logistic regression. "Spam" is a positive class ($y = 1$) and "not spam" is the negative class ($y = 0$). You have trained your classifier and there are $m = 1000$ examples in the cross-validation set. The chart of predicted class vs. actual class is:

1 / 1 point

	Actual Class: 1	Actual Class: 0
Predicted Class: 1	85	890
Predicted Class: 0	15	10

For reference:

- Accuracy = (true positives + true negatives) / (total examples)
- Precision = (true positives) / (true positives + false positives)
- Recall = (true positives) / (true positives + false negatives)
- F_1 score = $(2 * \text{precision} * \text{recall}) / (\text{precision} + \text{recall})$

What is the classifier's recall (as a value from 0 to 1)?

Enter your answer in the box below. If necessary, provide at least two values after the decimal point.

can correctly predict y when given only x .

☐ The classifier is likely to have unchanged precision and recall, but

lower accuracy.

of 99%.

precision and recall.

✓ Correct

A sufficiently large training set will not be overfit, as the model cannot overfit some of the examples without doing poorly on the others.