Application: Photo OCR

Latest Submission Grade
100%
1.
Question 1

Suppose you are running a sliding window detector to find text in images. Your input images are 1000x1000 pixels. You will run your sliding windows detector at two scales, 10x10 and 20x20 (i.e., you will run your classifier on lots of 10x10 patches to decide if they contain text or not; and also on lots of 20x20 patches), and you will "step" your detector by 2 pixels each time. About how many times will you end up running your classifier on a single 1000x1000 test set image?

1 / 1 point

250,000

1,000,000

100,000

500,000

Correct

With a stride of 2, you will run your classifier approximately 500 times for each dimension. Since you run the classifier twice (at two scales), you will run it 2 * 500 * 500 = 500,000 times.

2.
Question 2

Suppose that you just joined a product team that has been

developing a machine learning application, using m=1,000m = 1,000m=1,000

training examples. You discover that you have the option of

hiring additional personnel to help collect and label data.

You estimate that you would have to pay each of the labellers

$10 per hour, and that each labeller can label 4 examples per

minute. About how much will it cost to hire labellers to

label 10,000 new training examples?

1 / 1 point

$600

$10,000

$400

$250

Correct

On labeller can label 4×60=2404 \times 60 = 2404×60=240 examples in one hour. It will thus take him 10,000/240≈4010,000 / 240 \approx 4010,000/240≈40 hours to complete 10,000 examples. At $10 an hour, this is $400.

3.
Question 3

What are the benefits of performing a ceiling analysis? Check all that apply.

1 / 1 point

☐

It is a way of providing additional training data to the algorithm.

☐

If we have a low-performing component, the ceiling analysis can tell us if that component has a high bias problem or a high variance problem.

☑

It helps us decide on allocation of resources in terms of which component in a machine learning pipeline to spend more effort on.

Correct

The ceiling analysis reveals which parts of the pipeline have the most room to improve the performance of the overall system.

☑

It can help indicate that certain components of a system might not be worth a significant amount of work improving, because even if it had perfect performance its impact on the overall system may be small.

Correct

An unpromising component will have little effect on overall performance when it is replaced with ground truth.

4.
Question 4

Suppose you are building an object classifier, that takes as input an image, and recognizes that image as either containing a car (y=1y=1y=1) or not (y=0y=0y=0). For example, here are a positive example and a negative example:

After carefully analyzing the performance of your algorithm, you conclude that you need more positive (y=1y=1y=1) training examples. Which of the following might be a good way to get additional positive examples?

1 / 1 point

Mirror your training images across the vertical axis (so that a left-facing car now becomes a right-facing one).

Take a few images from your training set, and add random, gaussian noise to every pixel.

Take a training example and set a random subset of its pixel to 0 to generate a new example.

Select two car images and average them to make a third example.

Correct

A mirrored example is different from the original but equally likely to occur, so mirroring is a good way to generate new data.

5.
Question 5

Suppose you have a PhotoOCR system, where you have the following pipeline:

You have decided to perform a ceiling analysis on this system, and find the following:

Which of the following statements are true?

1 / 1 point

The potential benefit to having a significantly improved text detection system is small, and thus it may not be worth significant effort trying to improve it.

Correct

Plugging in ground truth text detection improved the overall system by only 2%, so it is not a good candidate for development effort.

☑

If we conclude that the character recognition's errors are mostly due to the character recognition system having high variance, then it may be worth significant effort obtaining additional training data for character recognition.

Correct

Since the biggest improvement comes from character recognition ground truth, we would like to improve the performance of that system. It the character recognition system has high variance, additional data will improve its performance.

☐

We should dedicate significant effort to collecting additional training data for the text detection system.

☐

The most promising component to work on is the text detection system, since it has the lowest performance (72%) and thus the biggest potential gain.