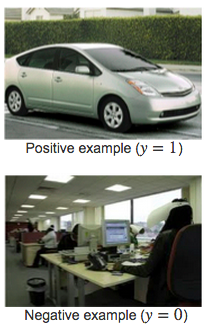
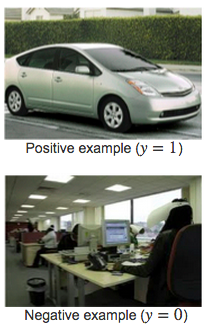
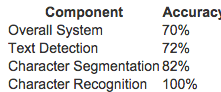
* Suppose you are running a sliding window detector to find text in images. Your input images are 1000x1000 pixels. You run sliding windows detector at 2 scales, 10x10 + 20x20 + will step your detector by 2 pixels each time. About how many times will you end up running your classifier on a single 1000x1000 test set image?
* **500,000** 🡺 ((1000-10)/2+1)^2+((1000-20)/2+1)^2
* Suppose you just joined a product team developing a ML application, using m = 1,000 training examples. You have the option of hiring additional personnel to help collect + label data. You estimate you have to pay each of the labelers $10/hour + that each labeler can label 4 examples per minute. About how much will it cost to hire labelers to label 10,000 new training examples?
* **$400** 🡺 10k examples / 4 per min = 2500 minutes = 41 hrs. = $410
* What are the benefits of performing a ceiling analysis? Check all that apply.
* **Can help indicate certain components of a system might not be worth a significant amount of work improving, b/c even w/ perfect performance its impact on overall system may be small.**
* **Gives info about which components, if improved, are most likely to have a significant impact on the performance of the final system.**
* Suppose you’re building an object classifier that takes as input an image + recognizes it as either containing a car (y = 1) or not (y = 0)

* After carefully analyzing the algorithm performance, you conclude you need more positive (y = 1) training examples. Which of the following might be a good way to get additional positive examples?
* **Mirror training images across the vertical axis (a left-facing car becomes a right-facing one).**
* **Apply translations, distortions, and rotations to the images already in your training set.**
* 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?
* **Potential benefit to having a significantly improved text detection system is small, + thus it may not be worth significant effort trying to improve it.**
* **If we conclude the character recognition's errors are mostly due to the character recognition system having high variance, it may be worth significant effort obtaining additional training data for character recognition.**
* **If the text detection system was trained using gradient descent, running gradient descent for more iterations is unlikely to help much.**