# Congratulations! You passed!

**Grade received** 93.33%

**Latest Submission Grade** 93.33%

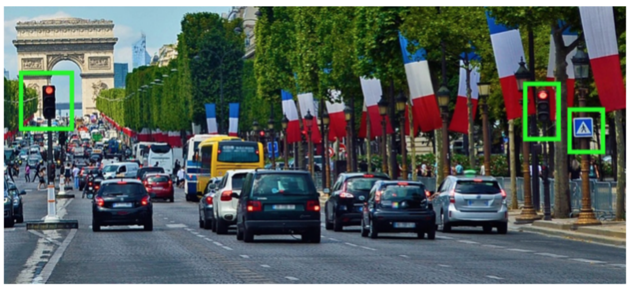
**To pass** 80% or higher

**Go to next item**

**1.** To help you practice strategies for machine learning, this week we'll present another scenario and ask how you would act. We think this "simulator" of working in a machine learning project will give a task of what leading a machine learning project could be like!

**1 / 1 point**

You are employed by a startup building self-driving cars. You are in charge of detecting road signs (stop sign, pedestrian crossing sign, construction ahead sign) and traffic signals (red and green lights) in images. The goal is to recognize which of these objects appear in each image. As an example, the above image contains a pedestrian crossing sign and red traffic lights



$$y^{(i)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{array}{l} \text{“stop sign”} \\ \text{“pedestrian crossing sign”} \\ \text{“construction ahead sign”} \\ \text{“red traffic light”} \\ \text{“green traffic light”} \end{array}$$

Your 100,000 labeled images are taken using the front-facing camera of your car. This is also the distribution of data you care most about doing well on. You think you might be able to get a much larger dataset off the internet, that could be helpful for training even if the distribution of internet data is not the same.

You are just getting started on this project. What is the first thing you do? Assume each of the steps below would take about an equal amount of time (a few days).

**Expand**

**Correct**

As discussed in lecture, applied ML is a highly iterative process. If you train a basic model and carry out error analysis (see what mistakes it makes) it will help point you in more promising directions.

**2.** Your goal is to detect road signs (stop sign, pedestrian crossing sign, construction ahead sign) and traffic signals (red and green lights) in images. The goal is to recognize which of these objects appear in each image. You plan to use a deep neural network with ReLU units in the hidden layers. **1 / 1 point**

Suppose that you use a sigmoid function for the output layer, and the output $\hat{y}$ has shape (5, 1). Which of the following best describes the cost function?

**Expand**

**Correct**

Correct. Here we compare each component of the prediction $\hat{y}$ with the respective component of the label $y$, and sum over the individual losses.

**3.** You are carrying out error analysis and counting up what errors the algorithm makes. Which of these datasets do you think you should manually go through and carefully examine, one image at a time?

**1 / 1 point**

**Expand**

**Correct**

Focus on images that the algorithm got wrong. Also, 500 is enough to give you a good initial sense of the error statistics. There's probably no need to look at 10,000, which will take a long time.

**4.** After working on the data for several weeks, your team ends up with the following data:

**1 / 1 point**

- 100,000 labeled images taken using the front-facing camera of your car.
- 900,000 labeled images of roads downloaded from the internet.
- Each image's labels precisely indicate the presence of any specific road signs and traffic signals or combinations of them. For example, $y^{(i)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ means the image contains a stop sign and a red traffic light.

Because this is a multi-task learning problem, when an image is not fully labeled (for example: $\begin{pmatrix} 0 \\ ? \\ ? \\ 1 \\ 0 \end{pmatrix}$) we can use it if we ignore those entries when calculating the loss function. True/False?

**Expand**

**Correct**

Correct. We can't use the components of the labels that are missing but we can use the ones we have to train the model.

**5.** The distribution of data you care about contains images from your car's front-facing camera; which comes from a different distribution than the images you were able to find and download off the internet. How should you split the dataset into train/dev/test sets?

**1 / 1 point**

**Expand**

**Correct**
Yes. As seen in the lecture, it is important that your dev and test set have the closest possible distribution to "real" data. It is also important for the training set to contain enough "real" data to avoid having a data-mismatch problem.

**6.** Assume you've finally chosen the following split between the data:

**1 / 1 point**

| **Dataset:** | **Contains:** | **Error of the algorithm:** |
|---|---|---|

| | | |
|---|---|---|
| Training | 940,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images) | 12% |
| Training-Dev | 20,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images) | 15.1% |
| Dev | 20,000 images from your car's front-facing camera | 12.6% |
| Test | 20,000 images from the car's front-facing camera | 15.8% |

You also know that human-level error on the road sign and traffic signals classification task is around 0.5%. Which of the following is True?

**Expand**

**Correct**

Correct. The avoidable bias is significantly high since the training error is a lot higher than the human-level error.

**7.** Assume you've finally chosen the following split between the data: **1 / 1 point**

| Dataset: | Contains: | Error of the algorithm: |
|---|---|---|

| | | |
|---|---|---|
| Training | 940,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images) | 8.8% |
| Training-Dev | 20,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images) | 9.1% |
| Dev | 20,000 images from your car's front-facing camera | 14.3% |
| Test | 20,000 images from the car's front-facing camera | 14.8% |

You also know that human-level error on the road sign and traffic signals classification task is around 0.5%. Based on the information given, a friend thinks that the training data distribution is much easier than the dev/test distribution. What do you think?

Expand

**Correct**
The algorithm does better on the distribution of data it trained on. But you don't know if it's because it trained on that distribution or if it really is easier. To get a better sense, measure human-level error separately on both distributions.

**8.** You decide to focus on the dev set and check by hand what the errors are due to. Here is a table summarizing your discoveries: **1 / 1 point**

| | |
|---|---|
| Overall dev set error | 15.3% |

| | |
|---|---|
| Errors due to incorrectly labeled data | 4.1% |
| Errors due to foggy pictures | 2.0% |
| Errors due to partially occluded elements. | 8.2% |
| Errors due to other causes | 1.0% |

In this table, 4.1%, 8.2%, etc. are a fraction of the total dev set (not just examples of your algorithm mislabeled). For example, about 8.2/15.3 = 54% of your errors are due to partially occluded elements in the image.

Which of the following is the correct analysis to determine what to prioritize next?

**Expand**

**Correct**
Correct. You should consider the tradeoff between the data accessibility and potential improvement of your model trained on this additional data.

**9.** You decide to focus on the dev set and check by hand what the errors are due to. Here is a table summarizing your discoveries: **1 / 1 point**

| | |
|---|---|
| Overall dev set error | 15.3% |
| Errors due to incorrectly labeled data | 4.1% |
| Errors due to foggy pictures | 3.0% |

| | |
|---|---|
| Errors due to partially occluded elements. | 7.2% |
| Errors due to other causes | 1.0% |

In this table, 4.1%, 7.2%, etc. are a fraction of the total dev set (not just examples of your algorithm mislabeled). For example, about 7.2/15.3 = 47% of your errors are due to partially occluded elements in the image.

From this table, we can conclude that if we fix the incorrectly labeled data we will reduce the overall dev set error to 11.2%. True/False?

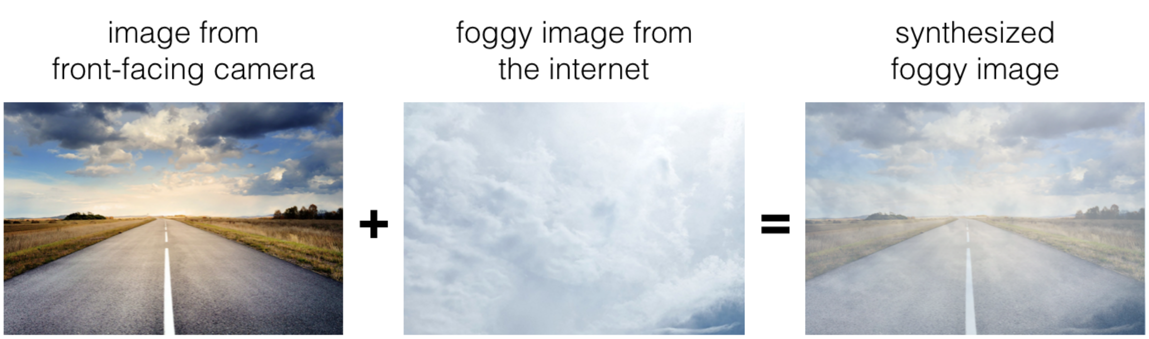
**Expand**

**Correct**

Correct. The 4.1 only gives you an estimate of the ceiling of how much the error can be improved by fixing the labels.

**10.** You decide to use data augmentation to address foggy images. You find 1,000 pictures of fog off the internet, and "add" them to clean images to synthesize foggy days, like this:

**1 / 1 point**



We can't use this data since they have a different distribution from the ones we used (internet and front-facing camera). True/False?

Expand

**Correct**

Correct. The new synthesized images are added to the training set and as long as they look realistic to the human eye this will be useful data to train the model.

**11.** After working further on the problem, you've decided to correct the incorrectly labeled data. Your team corrects the labels of the wrongly predicted images on the dev set. **1 / 1 point**

You have to correct the labels of the test so test and dev sets have the same distribution, but you won't change the labels on the train set because most models are robust enough they don't get severely affected by the difference in distributions. True/False?

**Expand**

**Correct**

Correct! To successfully train a model, the dev set and test set should come from the same distribution. Also, the deep learning models are robust enough to handle a small change in distributions, but if the errors are systematic they can significantly affect the training of the model.

**12.** One of your colleagues at the startup is starting a project to classify road signs as stop, dangerous curve, construction ahead, dead-end, and speed limit signs. Given how specific the signs are, he has only a small dataset and hasn't been able to create a good model. You offer your help providing the trained weights (parameters) of your model to transfer knowledge.

**1 / 1 point**

But your colleague points out that his problem has more specific items than the ones you used to train your model. This makes the transfer of knowledge impossible. True/False?

**Expand**

**Correct**

Correct. The model can benefit from the pre-trained model since there are many features learned by your model that can be used in the new problem.

**13.** Another colleague wants to use microphones placed outside the car to better hear if there are other vehicles around you. For example, if there is a police vehicle behind you, you would be able to hear their siren. However, they don't have much to train this audio system. How can you help? **1 / 1 point**

**Expand**

**Correct**
Yes. The problem he is trying to solve is quite different from yours. The different dataset structures make it probably impossible to use transfer learning or multi-task learning.

**14. To recognize a stop sign you use the following approach:** First, you use a neural network to predict bounding box co-ordinates around all traffic signs (if any) within an input image. You then pass the results to a different neural network to determine if the predicted traffic signs (if any) are a stop sign or not. We are using multi-task learning. True/False? **1 / 1 point**

**Expand**

**Correct**

Correct. Multi-task learning is about joining several tasks that can benefit from each other. Since there are 2 different neural networks being used here that do not share weights (i.e. structure), this problem has 2 single task learning neural networks and not a multi-task learning setup.

**15.** Consider the following two approaches, A and B: **0 / 1 point**

· **(A)** Input an image (x) to a neural network and have it directly learn a mapping to make a prediction as to whether there's a red light and/or green light (y).

· **(B)** In this two-step approach, you would first (i) detect the traffic light in the image (if any), then (ii) determine the color of the illuminated lamp in the traffic light.

Approach A tends to be more promising than approach B if you have a ________ (fill in the blank).

**Expand**

**Incorrect**