## Congratulations! You passed!

Grade received 86.66%

Latest Submission Grade 86.67%

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

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

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).

- ○ Spend a few days checking what is human-level performance for these tasks so that you can get an accurate estimate of Bayes error.
- ◉ Spend a few days training a basic model and see what mistakes it makes.
- ○ Spend a few days collecting more data using the front-facing camera of your car, to better understand how much data per unit time you can collect.
- ○ Spend a few days getting the internet data, so that you understand better what data is available.

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. For the output layer, a softmax activation would be a good choice for the output layer because this is a multi-task learning problem. True/False?

1 / 1 point

- ◉ False
- ○ True

Expand

**Correct**

Softmax would be a good choice if one and only one of the possibilities (stop sign, speed bump, pedestrian crossing, green light and red light) was present in each image.

**3.** You are working out error analysis and counting up what errors the algorithm makes. Which of the following do you think you should manually go through and carefully examine, one image at a time? **1 / 1 point**

- 500 images of the train set, on which the algorithm made a mistake.
- 500 images of the test set, on which the algorithm made a mistake.
- (selected) 500 images of the dev set, on which the algorithm made a mistake.
- 500 images of the training-dev set, on which the algorithm made a mistake.

Expand

**Correct**
Correct. We focus on images that the algorithm got wrong from the dev set. That is the one we use to make choices between different iterations of the system.

**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, you need to have all your $y^{(i)}$ vectors fully labeled. If one example is equal to $\begin{bmatrix} 0 \\ ? \\ 1 \\ 1 \\ ? \end{bmatrix}$ then the learning algorithm will not be able to use that example. True/False?

- True
- (selected) False

Expand

**Correct**
As seen in the lecture on multi-task learning, you can compute the cost such that it is not influenced by the fact that some entries haven't been labeled.

**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**

- (selected) Choose the training set to be the 900,000 images from the internet along with 80,000 images from your car's front-facing camera. The 20,000 remaining images will be split equally in dev and test sets.
- Mix all the 100,000 images with the 900,000 images you found online. Shuffle everything. Split the 1,000,000 images dataset into 600,000 for the training set, 200,000 for the dev set and 200,000 for the test set.
- Mix all the 100,000 images with the 900,000 images you found online. Shuffle everything. Split the 1,000,000 images dataset into 980,000 for the training set, 10,000 for the dev set and 10,000 for the test set.
- Choose the training set to be the 900,000 images from the internet along with 20,000 images from your car's front-facing camera. The 80,000 remaining images will be split equally in dev and test sets.

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: 0 / 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) | 1% |
| Training-Dev | 20,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images) | 5.1% |
| Dev | 20,000 images from your car's front-facing camera | 5.6% |
| Test | 20,000 images from the car's front-facing camera | 6.8% |

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?

- You have a high variance problem.
- You have a large data-mismatch problem.
- You have a high bias.
- The size of the train-dev set is too high.

Expand

Incorrect
Although not at human-level, the training error is not that high to say we have a high bias.

7. Assume you've finally chosen the following split between the data: 0 / 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) | 2% |
| Training-Dev | 20,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images) | 2.3% |
| Dev | 20,000 images from your car's front-facing camera | 1.3% |
| Test | 20,000 images from the car's front-facing camera | 1.1% |

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 harder than the dev/test distribution. What do you think?

- There's insufficient information to tell if your friend is right or wrong.
- Your friend is probably right. (i.e., Bayes error for the dev/test distribution is probably lower than for the train distribution.)
- Your friend is wrong. (i.e., Bayes error for the dev/test distribution is probably higher than for the train distribution.)

Expand

Incorrect
Notice that the test and dev errors are lower than the train and train-dev errors.

**8.** You decide to focus on the dev set and check by hand what are the errors 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 | 8.0% |
| Errors due to rain drops stuck on your car's front-facing camera | 2.2% |
| Errors due to other causes | 1.0% |

In this table, 4.1%, 8.0%, etc. are a fraction of the total dev set (not just examples of your algorithm mislabeled). For example, about 8.0/15.3 = 52% of your errors are due to foggy pictures.

The results from this analysis implies that the team's highest priority should be to bring more foggy pictures into the training set so as to address the 8.0% of errors in that category. True/False?

Additional note: there are subtle concepts to consider with this question, and you may find arguments for why some answers are also correct or incorrect. We recommend that you spend time reading the feedback for this quiz, to understand what issues that you will want to consider when you are building your own machine learning project.

- True because it is the largest category of errors. We should always prioritize the largest category of errors as this will make the best use of the team's time.
- First start with the sources of error that are least costly to fix.
- True because it is greater than the other error categories added together $8.0 > 4.1 + 2.2 + 1.0$.
- **(selected)** False because it depends on how easy it is to add foggy data. If foggy data is very hard and costly to collect, it might not be worth the team's effort.

Expand

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

**9.** You can buy a specially designed windshield wiper that helps wipe off some of the raindrops on the front-facing camera. **1 / 1 point**

| | |
|---|---|
| Overall dev set error | 15.3% |
| Errors due to incorrectly labeled data | 4.1% |
| Errors due to foggy pictures | 8.0% |
| Errors due to rain drops stuck on your car's front-facing camera | 2.2% |
| Errors due to other causes | 1.0% |

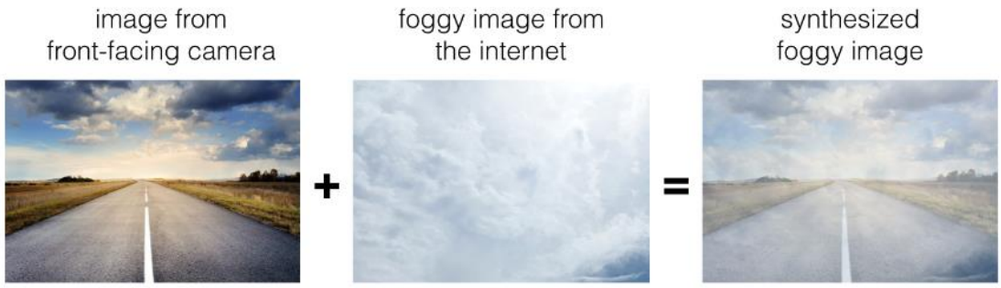
Which of the following statements do you agree with?

- 2.2% would be a reasonable estimate of how much this windshield wiper could worsen performance in the worst case.
- 2.2% would be a reasonable estimate of how much this windshield wiper will improve performance.
- **(selected)** 2.2% would be a reasonable estimate of the maximum amount this windshield wiper could improve performance.
- 2.2% would be a reasonable estimate of the minimum amount this windshield wiper could improve performance.

Expand

**Correct**
Yes. You will probably not improve performance by more than 2.2% by solving the raindrops problem. If your dataset was infinitely big, 2.2% would be a perfect estimate of the improvement you can achieve by purchasing a specially designed windshield wiper that removes the raindrops.

**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?

- ○ 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?

- ○ False, the test set should be changed, but also the train set to keep the same distribution between the train, dev, and test sets.
- ○ False, the test set shouldn't be changed since we want to know how the model performs in real data.
- ◉ True, as pointed out, we must keep dev and test with the same distribution. And the labels at training should be fixed only in case of a systematic error.

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.** Your client asks you to add the capability to detect dogs that may be crossing the road to the system. He can provide a relatively small set containing dogs. Which of the following do you agree most with? **1 / 1 point**

- ○ You should train a single new model for the dogs' task, and leave the previous model as it is.
- ○ Using pre-trained weights can severely hinder the ability of the model to detect dogs since they have too many learned features.
- ○ You will have to re-train the whole model now including the dogs' data.
- ◉ You can use weights pre-trained on the original data, and fine-tune with the data now including the dogs.

Expand

**Correct**
Correct. Since your model has learned useful low-level features to tackle the new task we can conserve those by using the pre-trained weights.

**13.** 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. He has approximately 30,000 examples of each image and 30,000 images without a sign. This case could benefit from using multi-task learning. True/False?

1 / 1 point

- ○ False
- ◉ True

Expand

**Correct**
Correct. There are a lot of high-level features that all the required signs share. This is a great scenario to make use of multi-task learning.

**14.** To recognize a stop sign you use the following approach: First, we localize any traffic sign in an image. After that, we determine if the sign is a stop sign or not. We are using multi-task learning. True/False?

1 / 1 point

- ◉ False
- ○ True

Expand

**Correct**
Correct. Multi-task learning is about joining several tasks that can benefit from each other.

**15.** Consider the following two approaches, A and B:

1 / 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).

- ○ Multi-task learning problem.
- ○ Large bias problem.
- ◉ Large training set
- ○ Problem with a high Bayes error.

Expand

**Correct**
Yes. In many fields, it has been observed that end-to-end learning works better in practice, but requires a large amount of data.