**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!

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

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.

**Suppose that you came from working with a project for human detection in city parks, so you know that detecting humans in diverse environments can be a difficult problem. What is the first thing you do? Assume each of the steps below would take about an equal amount of time (a few days).**



Train a basic model and proceed with error analysis.

**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 training a basic model and see what mistakes it makes.

### 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.

**Suppose that you use a sigmoid function for the output layer, and the output �^*y*^​ has shape (5, 1). Which of the following best describes the cost function?**

∑i=15∑j=1L(^y(i)j,yij)

**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?**

FALSE

### 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?

500 images on which the algorithm made a mistake

### 3.When trying to determine what strategy to implement to improve the performance of a model, we manually check all images of the training set where the algorithm was successful. True/False?

### FALSE

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

* 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*)=⎣​1 0 0 1 0​⎦​ means** the image contains a stop sign and a red traffic light.

When using a non fully labeled image such **as *y*(*i*)=⎣​0 ? 1 ? 1​⎦​, which** of the following strategies is most appropriate to calculate the loss function to train as a multi-task learning problem?

Calculate the loss as ∑L(^y(i)j,y(i)jwhere the sum goes over all the know components of y(i)

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

* 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) = [1 0 0 1 0]** 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 **[0 ? 1 1 ?]** then the learning algorithm will not be able to use that example. True/False?

### FALSE

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

* 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*)=⎣1 0 0 1 0​⎦​** 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: (0 ? ? 1 0​ ) we** can use it if we ignore those entries when calculating the loss function. True/False?

TRUE

### 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. The best way to split the data is using the 900,000 internet images to train, and divide the 100,000 images from your car's front-facing camera between dev and test sets. True/False?

### FALSE

**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?**



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.

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

* + **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%. Which of the following are True? (Check all that apply).**

* + You have a large avoidable-bias problem because your training error is quite a bit higher than the human-level error.
  + You have a large data-mismatch problem because your model does a lot better on the training-dev set than on the dev set.

### 6.Assume you’ve finally chosen the following split between the data:

|  |  |  |
| --- | --- | --- |
| **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.

### 7.

Question 7

|  |  |  |
| --- | --- | --- |
| **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% |

Assume you’ve finally chosen the following split between the data:

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 you conclude that the Bayes error for the dev/test distribution is probably higher than for the train distribution. True/False?

### FALSE

**7.Based on table from the previous question, a friend thinks that the training data distribution is much easier than the dev/test distribution. What do you think?**

* There’s insufficient information to tell if your friend is right or wrong.

### 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:

|  |  |
| --- | --- |
| 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.

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.

**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:**

* Overall dev set error 14.3%
* 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 your algorithm mislabeled). I.e. about 8.0/14.3 = 56% 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?

*  False because this would depend on how easy it is to add this data and how much you think your team thinks it’ll help.

### 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:

|  |  |
| --- | --- |
| 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.

You shouldn't invest all your efforts to get more images with partially occluded elements since 4.1 + 3.0 + 1.0 = 8.1 > 7.2. True/False?

### FALSE

**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:

|  |  |
| --- | --- |
| 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.

**You find out that there is an anti-reflective film guarantee to eliminate the sun reflection, but it is quite costly. Which of the following gives the best description of what the investment in the film can do to the model?**



The film will reduce the dev set error with 7.2% at the most.

### 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:

|  |  |
| --- | --- |
| 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?

FALSE

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

|  |  |
| --- | --- |
| 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 the maximum amount this windshield wiper could improve performance.

### 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:

Which of the following do you agree with?

If used, the synthetic data should be added to the training set.

**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:**

Which of the following statements do you agree with? (Check all that apply.)

* So long as the synthesized fog looks realistic to the human eye, you can be confident that the synthesized data is accurately capturing the distribution of real foggy images, since human vision is very accurate for the problem you’re solving.

### 10.

Question 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:

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

FALSE

**11. After working further on the problem, you’ve decided to correct the incorrectly labeled data on the dev set. Which of these statements do you agree with? (Check all that apply).**

* You should not correct incorrectly labeled data in the training set as well so as to avoid your training set now being even more different from your dev set.
* You should also correct the incorrectly labeled data in the test set, so that the dev and test sets continue to come from the same distribution

### 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.

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?



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.

### 12.So far your algorithm only recognizes red and green traffic lights. One of your colleagues in the startup is starting to work on recognizing a yellow traffic light. (Some countries call it an orange light rather than a yellow light; we’ll use the US convention of calling it yellow.) Images containing yellow lights are quite rare, and she doesn’t have enough data to build a good model. She hopes you can help her out using transfer learning.

What do you tell your colleague?



She should try using weights pre-trained on your dataset, and fine-tuning further with the yellow-light dataset.

### 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?



You can use weights pre-trained on the original data, and fine-tune with the data now including the dogs.

### 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.

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?

FALSE

### 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?

TRUE

**13.Another colleague wants to use microphones placed outside the car to better hear if there’re 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?**

* Neither transfer learning nor multi-task learning seems promising.

### 13.One of your colleagues at the startup is starting a project to classify stop signs in the road as speed limit signs or not. He has approximately 30,000 examples of each image and 30,000 images without a sign. He thought of using your model and applying transfer learning but then he noticed that you use multi-task learning, hence he can't use your model. True/False?

FALSE

### 14.

Question 14

When building a system to detect cattle crossing a road from images taken with the front-facing camera of a truck, the designers had a large dataset of images. Which of the following might be a reason to use an end-to-end approach?  


There is a large dataset available.

**14. To recognize red and green lights, you have been using this approach:**

* **(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).**

**A teammate proposes a different, two-step approach:**

* **(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. Between these two, Approach B is more of an end-to-end approach because it has distinct steps for the input end and the output end. True/False?**

FALSE

### 14.

Question 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?

TRUE

### 15.An end-to-end approach doesn't require that we hand-design useful features, it only requires a large enough model. True/False?

TRUE

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

·       **(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).



Large training set