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 you an idea 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, this 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, which 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).

- ○ Leave aside the pedestrian detection, to move faster and then later solve the pedestrian problem alone.
- ◉ Spend a few days collecting more data to determine how hard it will be to include more pedestrians in your dataset.
- ○ Start by solving pedestrian detection, since you already have the experience to do this.
- ○ Train a basic model and proceed with error analysis.

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
- ( ) 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 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**

- ( ) 10,000 images on which the algorithm made a mistake
- ( ) 10,000 randomly chosen images
- (•) 500 images on which the algorithm made a mistake
- ( ) 500 randomly chosen images

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: **0 / 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?

- (●) False
- ( ) True

Expand

Incorrect

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. 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?

- ○ True
- ◉ False

Expand

Correct
Correct. 100,000 images are too many to use in dev and test. A better distribution would be to use 80,000 of those images to train, and split the rest between dev and test.

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?

- ◉ You have a high bias.
- ○ You have a high variance problem.
- ○ You have a large data-mismatch problem.
- ○ You have a too low avoidable bias.

Expand

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

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

1 / 1 p

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

- (•) False
- ( ) True

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?

- ( ) True
- (•) False

Expand

Correct

Correct. These kinds of arguments don't help us to decide on the strategy to follow. Other factors should be used, such as the tradeoff between the cost of getting new images and the improvement of the system performance.

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:

0 / 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%

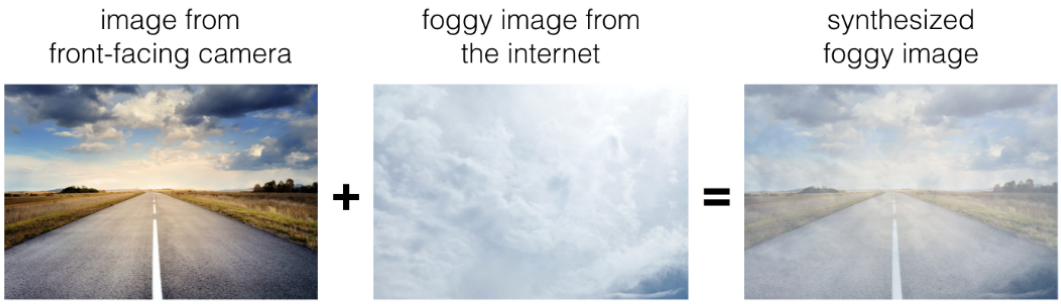
In this table, 4.1%, 7.2%, etc. are a fraction of the total dev set (not just examples of your algorithm mislabeled). For example, about 7.2/15.3 = 47% of your errors are due to partially occluded elements.

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 overall test set error will be reduced by at most 7.2%.
- ( ) The film will reduce the dev set error with 7.2% at the most.
- ( ) The film will reduce at least 7.2% of the dev set error.

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:

0 / 1 point



Which of the following do you agree with?

- (●) It is irrelevant how the resulting foggy images are perceived by the human eye, the most important thing is that they are correctly synthesized.

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 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.
- ( ) False, the test set should be changed, but also the train set to keep the same distribution between the train, dev, and test sets.

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 po

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

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 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? 0 / 1 point

- ◉ True
- ○ False

Expand

✗ **Incorrect**
When using transfer learning we can remove the last layer, which is one of the aspects that is different from a binary classification problem.

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? 0 / 1 point

- ○ That is the default approach on computer vision tasks.
- ◉ This approach will make use of useful hand-designed components.
- ○ It requires less computational resources.
- ○ There is a large dataset available.

Expand

✗ **Incorrect**
An end-to-end approach avoids the use of hand-designed components.

15. To recognize a stop sign you use the following approach: 1 / 1 point

First, we localize any traffic sign in an image. After that, we determine if the sign is a stop sign or not.

This is a better approach than an end-to-end model for which of the following cases? Choose the best answer.

- ○ There is a large amount of data.
- ○ The problem has a high Bayes error.
- ○ There are available models which we can use to transfer knowledge.
- ◉ There is not enough data to train a big neural network.

Correct

Correct. This might be the most important factor when deciding whether to use an end-to-end approach.