


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



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

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).

1 / 1 poin
Luaskan

Benar
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.
Pertanyaan #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, which of the following gives you the most appropriate activation function?

1 / 1 poin
Luaskan

Benar
Correct. This works well since the output would be valued between 0 and 1 which represents the probability that one of the possibilities is present in an image.

3.
Pertanyaan #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 poin
Luaskan

Benar
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.
Pertanyaan #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, $\begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} y^{(i)} = \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}$ 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: (0??10) \ \ \ \ \ (0??10) \ \ \ \ \) we can use it if we ignore those entries when calculating the loss function. True/False?

1 / 1 poin

Luaskan

Benar

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.

Pertanyaan #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?

1 / 1 poin

Luaskan

Benar

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.

6.

Pertanyaan #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?

0 / 1 poin

Luaskan

Salah

The training-dev error and the dev error are not that different to come to this conclusion.

7.

Pertanyaan #7

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)	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?

1 / 1 poin

Luaskan

Benar
Correct. Since the training-dev error is higher than the dev and test errors, the dev/test distribution is probably "easier" than the training distribution.

8.
Pertanyaan #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.

1 / 1 poin
Luaskan

Benar
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.
Pertanyaan #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?

1 / 1 poin
Luaskan

Benar
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.
Pertanyaan #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:

image from
front-facing camera



+

foggy image from
the internet



=

synthesized
foggy image



Which of the following do you agree with?

0 / 1 poin

Luaskan

Salah

No. Our objective is to have images that look realistic to the human eye.

11.

Pertanyaan #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).

1 / 1 poin

Luaskan

Benar

Great, you got all the right answers.

12.

Pertanyaan #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?

1 / 1 poin

Luaskan

Benar

Yes. You have trained your model on a huge dataset, and she has a small dataset. Although your labels are different, the parameters of your model have been trained to recognize many characteristics of road and traffic images which will be useful for her problem. This is a perfect case for transfer learning, she can start with a model with the same architecture as yours, change what is after the last hidden layer and initialize it with your trained parameters.

13.

Pertanyaan #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 poin

Luaskan

Benar

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.

Pertanyaan #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 poin

Luaskan

Benar

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

15.

Pertanyaan #15

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.

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

0 / 1 poin

Luaskan

Salah

No, such aspects don't play a major role in deciding whether to use an end-to-end approach or not.