

Autonomous Driving (Case Study)

1.

Question 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 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.

You are getting started with 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 point

Expand

Correct

Applied ML is highly iterative. Having a basic model to do an error analysis can point you in the most promising directions with a lot of certainties.

2.

Question 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

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.

Question 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?

1 / 1 point

Expand

Correct

Correct. This set should be too large to manually check all the images. It is better to focus on the images that the algorithm got wrong from the dev set. Also, choose a large enough subset that we can manually check.

4.

Question 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, $y^{(i)} = [10010] \text{ [stop sign] } [10010] \text{ [red traffic light]}$ 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 $[0?11?] \text{ [stop sign] } [0?11?] \text{ [red traffic light]}$ then the learning algorithm will not be able to use that example. True/False?

1 / 1 point

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.

Question 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. Which of the following are true about the train/dev/test split?

1 / 1 point

Expand

Correct

Great, you got all the right answers.

6.

Question 6

Assume you've finally chosen the following split between of 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)	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).

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

7.

Question 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 point

Expand

Correct

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.

Question 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	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?

1 / 1 point

Expand

Correct

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

9.

Question 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?

1 / 1 point

Expand

Correct

Yes. Remember that this 7.2% gives us an estimate for the ceiling of how much the error can be reduced when the cause is fixed.

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:



Which of the following statements do you agree with?

1 / 1 point

Expand

Correct

Yes. If the synthesized images look realistic, then the model will just see them as if you had added useful data to identify road signs and traffic signals in foggy weather. I will very likely help.

11.

Question 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. Which of the following is a necessary step to take?

0 / 1 point

Expand

Incorrect

No, once the labels have been corrected we use those labels.

12.

Question 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

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.

Question 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.

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?

1 / 1 point

Expand

Correct

Correct. To get good results when using an end-to-end approach, it is necessary to have a big dataset.

15.

Question 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.

1 / 1 point

Expand

Correct

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

