Autonomous driving (case study)
Graded Quiz • 45 min

**Due** Sep 14, 7:59 AM BST

**Congratulations! You passed!**

**TO PASS** 80% or higher

Keep Learning

**GRADE**
96.66%

# Autonomous driving (case study)

**LATEST SUBMISSION GRADE**

96.66%

1. To help you practice strategies for machine learning, in 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

$$y^{(i)} = \left[\begin{array}{c}0\\1\\0\\1\\0\end{array}\right]\begin{array}{l}\text{“stop sign”}\\\text{“pedestrian crossing sign”}\\\text{“construction ahead sign”}\\\text{“red traffic light”}\\\text{“green traffic light”}\end{array}$$

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 getting the internet data, so that you understand better what data is available.
- ○ 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.

✓ **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.

**1 / 1 point**

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?

- ○ True
- ◉ False

✓ **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 randomly chosen images
- ○ 500 randomly chosen images
- ◉ 500 images on which the algorithm made a mistake
- ○ 10,000 images on which the algorithm made a mistake

✓ **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:

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

- ○ 10,000 randomly chosen images
- ○ 500 randomly chosen images
- ◉ 500 images on which the algorithm made a mistake
- ○ 10,000 images on which the algorithm made a mistake

**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:

**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


- 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
- (•) False

**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.

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, $y^{(i)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ 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
- (•) False

**Correct**

As seen in the lecture on multi-task learning, you can compute the

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, $y^{(i)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

$\begin{bmatrix} 0 \end{bmatrix}$

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
- ◉ False

✓ **Correct**

- ○ 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.
- ○ 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.

✓ **Correct**

Yes. As seen in 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 of the data: **1 / 1 point**

| **Dataset:** | **Contains:** | **Error of the algorithm:** |
|---|---|---|
| | 940,000 images randomly picked from | |

| | | |
|---|---|---|
| Training | (900,000 internet images + 60,000 car’s front-facing camera images) | 8.8% |
| **Dataset:** | **Contains:** | **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).

- [ ] You have a large variance problem because your model is not generalizing well to data from the same training distribution but that it has never seen before.

- [ ] Your algorithm overfits the dev set because the error of the dev and test sets are very close.

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 error as this will make the best use of the team's time.
- ○ True because it is greater than the other error categories added together (8.0 > 4.1+2.2+1.0).
- ◉ 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.
- ○ First start with the sources of error that are least costly to fix.

✓ **Correct**

Correct. feedback: 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 help wipe off some of the raindrops on the front-facing camera. Based on the table from the previous question, which of the following statements do you agree with? **1 / 1 point**

- ◉ 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.
- ○ 2.2% would be a reasonable estimate of how much this windshield wiper will improve performance.
- ○ 2.2% would be a reasonable estimate of how much this windshield wiper could worsen performance in the worst case.

✓ **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

Correct

correct: feedback: 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 help wipe off some of the raindrops on the front-facing camera. Based on the table from the previous question, which of the following statements do you agree with?

1 / 1 point

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

Correct

Yes. You will probably not improve performance by more than

- (●) 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 (or a subset of it), since human vision is very accurate for the problem you're solving.

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 a foggy weather. I will very likely help.

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).

**0.5 / 1 point**

- [x] 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

**Correct**

Yes because you want to make sure that your dev and test data come from the same distribution for your algorithm to make your 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 help wipe off some of the raindrops on the front-facing camera. Based on the table from the previous question, which of the following statements do you agree with?

**1 / 1 point**

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

**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

- (•) 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 (or a subset of it), since human vision is very accurate for the problem you're solving.

**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 a foggy weather. I will very likely help.

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). **0.5 / 1 point**

- [x] 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

**Correct**

Yes because you want to make sure that your dev and test data come from the same distribution for your algorithm to make your team's iterative development process is efficient.

- [ ] You do not necessarily need to fix the incorrectly labeled data in the training set, because it's okay for the training set distribution to differ from the dev and test sets. Note that it is important that the dev set and test set have the same distribution.

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. **1 / 1 point**

What do you tell your colleague?

- (selected) She should try using weights pre-trained on your dataset, and fine-tuning further with the yellow-light dataset.
- If she has (say) 10,000 images of yellow lights, randomly sample 10,000 images from your dataset and put your and her data together. This prevents your dataset from “swamping” the yellow lights dataset.

- **(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?

- True
- (selected) False

**Correct**

Yes. (A) is an end-to-end approach as it maps directly the input (x) to the output (y).

15. Approach A (in the question above) tends to be more promising than approach B if you have a ________ (fill in the blank).

**1 / 1 point**

- (selected) Large training set
- Multi-task learning problem.
- Large bias problem.