Congratulations! You passed!

Grade received 100% **To pass** 80% or higher

Item Navigation

Go to next item Retake the assignment in **7h 43m**

Autonomous Driving (Case Study)

Quiz • 45 min

Autonomous Driving (Case Study)

Submit your assignment

Latest Submission Grade 100%

Due Apr 18, 2:59 PM CST **Attempts** 3 every 8 hours

Try again

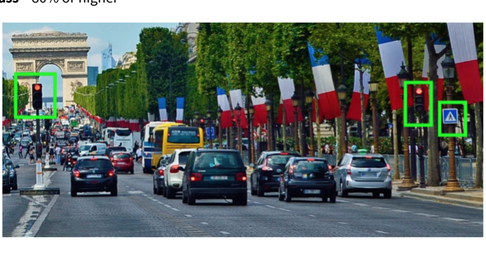
Retake the quiz in **7h 43m**

Receive grade

To Pass 80% or higher

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 you an idea 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, this image contains a pedestrian crossing sign and red traffic lights.



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

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, which could be helpful for training even if the distribution of internet data is not the same.

Like Dislike Report an issue

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.
- ( ) 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.
- ( ) Leave aside the pedestrian detection, to move faster and then later solve the pedestrian problem alone.

Correct
Correct. As discussed in the lecture, it is better to create your first system quickly and then iterate.

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?

- (●) False
- ( ) True

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**

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

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 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.

When using a non fully labeled image such as $y^{(i)} = \begin{bmatrix} 0 \\ ? \\ 1 \\ ? \\ 1 \end{bmatrix}$, which of the following strategies is most appropriate to calculate the loss function to train as a multi-task learning problem?

- ( ) Make the missing entries equal to 0.
- (●) Calculate the loss as $\sum \mathcal{L}(\hat{y}_j^{(i)}, y_j^{(i)})$ where the sum goes over all the know components of $y^{(i)}$.
- ( ) Make the missing entries equal to 1.
- ( ) It is not possible to use non fully labeled images if we train as a multi-task learning problem.

Correct
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. 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? **1 / 1 point**

- ( ) 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.
- (●) 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.
- ( ) Mix all the 100,000 images with the 900,000 images you found online. Shuffle everything. Split the 1,000,000 images dataset into 600,000 for the training set, 200,000 for the dev set and 200,000 for the test set.

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

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 large data-mismatch problem.
- ( ) The size of the train-dev set is too high.
- (●) You have a high variance problem.
- ( ) You have a high bias.

Correct
Correct. Since the difference between the training-dev error and the training error is high.

7. Assume you've finally chosen the following split between the data: **1 / 1 point**

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 higher than for the train distribution. True/False?

- (●) True
- ( ) False

Correct
Correct. Notice that the test and dev errors are lower than the train and train-dev errors.

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: **1 / 1 point**

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?

- ( ) You should prioritize getting more foggy pictures since that will be easier to solve.
- (●) You should weigh how costly it would be to get more images with partially occluded elements, to decide if the team should work on it or not.
- ( ) Since there is a high number of incorrectly labeled data in the dev set, you should prioritize fixing the labels on the whole training set.
- ( ) Since 8.2 > 4.1 + 2.0 + 1.0, the priority should be to get more images with partially occluded elements.

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

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: **1 / 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%

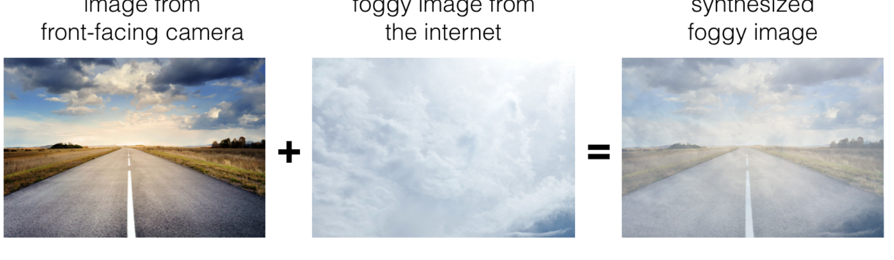
In this table, 4.1%, 7.2%, etc. are a fraction of the total dev set (not just examples of your algorithm mislabeled). For example, about 7.2/15.3 = 47% of your errors are due to partially occluded elements.

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

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. 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: **1 / 1 point**



Which of the following statements do you agree with?

- ( ) Adding synthesized images that look like real foggy pictures taken from the front-facing camera of your car to the training dataset won't help the model improve because it will introduce avoidable bias.
- (●) 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.
- ( ) There is little risk of overfitting to the 1,000 pictures of fog so long as you are combining it with a much larger (>>1,000) set of clean/non-foggy images.

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. 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? **1 / 1 point**

- ( ) Create a train-dev set to estimate how many incorrectly labeled examples are in the train set.
- ( ) Correct the labels of the train set.
- (●) Correct the labels of the test set.
- ( ) Use a correctly labeled version and an incorrectly labeled version to make the model more robust.

Correct
Correct. Recall that the dev set and the test set must come from 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?

- ( ) Recommend that she try multi-task learning instead of transfer learning using all the data.
- ( ) You cannot help her because the distribution of data you have is different from hers, and is also lacking the yellow label.
- (●) 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.

Correct
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. 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? **1 / 1 point**

- ( ) False
- (●) True

Correct
Correct. There are a lot of high-level features that all the required signs share. This is a great scenario to make use of multi-task learning.

14. To recognize red and green lights, you have been using this approach: **1 / 1 point**

- **(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?

- ( ) True
- (●) False

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

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

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