Autonomous Driving (Case Study)

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! 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

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

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).

- ○ 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.
- ◉ Spend a few days training a basic model and see what mistakes it makes.
- ○ 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.

Expand

✓ 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. 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

- ○ True
- ◉ False

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.

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%

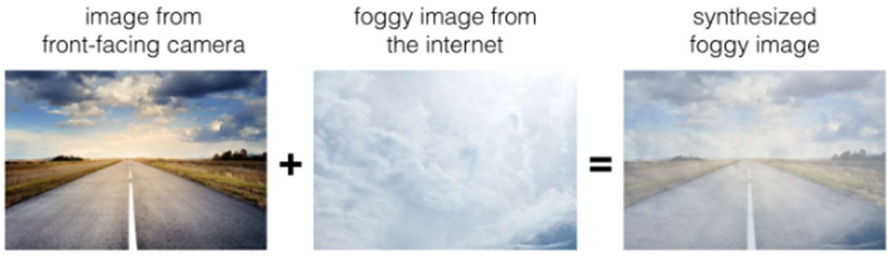
You also know that human-level error on the road sign and traffic signals classification task is around 0.5%. Based on the information given, a friend thinks that the training data distribution is much harder than the dev/test distribution. What do you think?

- ○ Your friend is wrong. (i.e., Bayes error for the dev/test distribution is probably higher than for the train distribution.)
- ○ There's insufficient information to tell if your friend is right or wrong.
- ◉ Your friend is probably right. (i.e., Bayes error for the dev/test distribution is probably lower than for the train distribution.)

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.

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



We can't use this data since they have a different distribution from the ones we used (internet and front-facing camera). True/False?

- ○ True
- ◉ False

Expand

✓ Correct
Correct. The new synthesized images are added to the training set and as long as they look realistic to the human eye this will be useful data to train 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 point

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

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?

- (●) False
- ( ) True

Expand

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

15. Consider the following two approaches, A and B: 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).

· **(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.

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

- ( ) Problem with a high Bayes error.
- (●) Large training set
- ( ) Large bias problem.
- ( ) Multi-task learning problem.

Expand

Correct
Yes. In many fields, it has been observed that end-to-end learning works better in practice, but requires a large amount of data.

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 point

- ○ Softmax
- ○ ReLU
- ○ Linear
- ◉ Sigmoid

Expand

✓ Correct
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. 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 point

- ○ 500 images of the training-dev set, on which the algorithm made a mistake.
- ○ 500 images of the train set, on which the algorithm made a mistake.
- ◉ 500 images of the dev set, on which the algorithm made a mistake.
- ○ 500 images of the test set, on which the algorithm made a mistake.

Expand

✓ Correct
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. 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.

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

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. 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 point

- False
- True

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.

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	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.

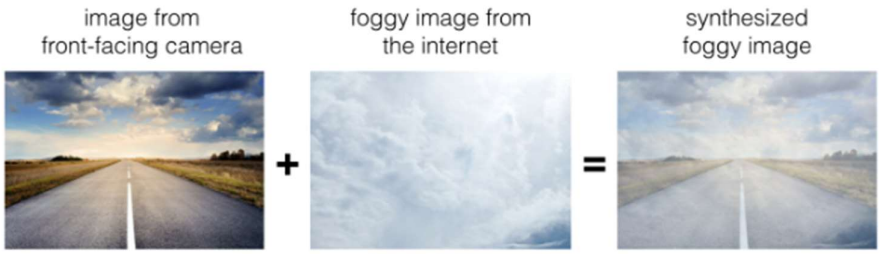
You shouldn't invest all your efforts to get more images with partially occluded elements since 4.1 + 3.0 + 1.0 = 8.1 > 7.2. True/False?

- (•) False
- ( ) True

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.

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?

- (●) 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.
- ( ) 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.
- ( ) 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.

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. 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 poi

- [ ] You should correct incorrectly labeled data in the training set as well so as to avoid your training set now being even more different from your dev set.
- [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 efficient.

- [x] 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.

  ✓ **Correct**
  True, deep learning algorithms are quite robust to having slightly different train and dev distributions.

- [ ] You should not correct the incorrectly labeled data in the test set, so that the dev and test sets continue to come from the same distribution.

Expand

✓ **Correct**
Great, you got all the right answers.

12. 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. Given how specific the signs are, he has only a small dataset and hasn't been able to create a good model. You offer your help providing the trained weights (parameters) of your model to transfer knowledge. 1 / 1 point

But your colleague points out that his problem has more specific items than the ones you used to train your model. This makes the transfer of knowledge impossible. True/False?

- ○ True
- ◉ False

Expand

✓ **Correct**
Correct. The model can benefit from the pre-trained model since there are many features learned by your model that can be used in the new problem.

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

Expand

✓ 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. 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?

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

15. An end-to-end approach doesn't require that we hand-design useful features, it only requires a large enough model. True/False? 1 / 1 point

- ◉ True
- ○ False

Expand

✓ Correct
Correct. This is one of the major characteristics of deep learning models, that we don't need to hand-design the features.

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?

- It is not possible to use non fully labeled images if we train as a multi-task learning problem.
- Make the missing entries equal to 1.
- 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 0.

Expand

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. Which of the following are true about the train/dev/test split? 1 / 1 point

- [x] The dev and test set must come from the front-facing camera.

  ✓ **Correct**
  Correct. This is the distribution we care about most, thus we should use this as a target.

- [ ] The dev and test sets must contain some images from the internet.

- [x] The dev and test sets must come from the same distribution.

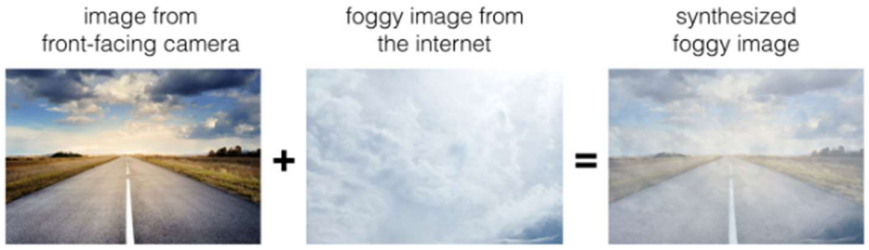
  ✓ **Correct**
  Correct. This is required to aim the target where we want to be.

- [ ] The train, dev, and test must come from the same distribution.

Expand

Correct
Great, you got all the right answers.

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 do you agree with?

- ○ If used, the synthetic data should be added to the training/dev/test sets in equal proportions.
- ○ With this technique, we duplicate the size of the training set by synthesizing a new foggy image for each image in the training set.
- ○ It is irrelevant how the resulting foggy images are perceived by the human eye, the most important thing is that they are correctly synthesized.
- ◉ If used, the synthetic data should be added to the training set.

Expand

✓ Correct
Yes. The synthetic data can help to train the model to get better performance at the dev set, but shouldn't be added to the dev or test sets because they don't represent our target in a completely accurate way.

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?

- ◉ 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.
- ○ False, the test set shouldn't be changed since we want to know how the model performs in real data.

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.

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

- ○ Transfer learning from your vision dataset could help your colleague get going faster. Multi-task learning seems significantly less promising.
- ○ Multi-task learning from your vision dataset could help your colleague get going faster. Transfer learning seems significantly less promising.
- ◉ Neither transfer learning nor multi-task learning seems promising.
- ○ Either transfer learning or multi-task learning could help our colleague get going faster.

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.

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

Expand

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

9. You can buy a specially designed windshield wiper that helps wipe off some of the raindrops on the front-facing camera. 1 / 1 point

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?

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

Expand

✓ 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 windshield wiper that removes the raindrops.

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

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

Expand

✓ Correct
Correct. Recall that the dev set and the test set must come from the same distribution.