Congratulations! You passed!

Grade received 80% | Latest Submission Grade 80% | To pass 80% or higher

Go to next item

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 checking what is human-level performance for these tasks so that you can get an accurate estimate of Bayes error.
- ○ Spend a few days getting the internet data, so that you understand better what data is available.
- ◉ Spend a few days training a basic model and see what mistakes it makes.

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, which of the following gives you the most appropriate activation function? **1 / 1 point**

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

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. When trying to determine what strategy to implement to improve the performance of a model, we manually check **1 / 1 point**

all images of the training set where the algorithm was successful. True/False?

- ○ True
- ◉ False

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. 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?

- ◉ True
- ○ False

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? **0 / 1 point**

- ○ False
- ◉ True

Expand

Incorrect
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. Assume you've finally chosen the following split between the data: **0 / 1 point**

Dataset:	Contains:	Error of the algorithm:
	940,000 images randomly picked from (900,000 internet images +	

Training	60,000 car's front-facing camera images)	12%
Training-Dev	20,000 images randomly picked from (900,000 internet images + 60,000 car's front-facing camera images)	15.1%
Dev	20,000 images from your car's front-facing camera	12.6%
Test	20,000 images from the car's front-facing camera	15.8%

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 high variance problem.
- ( ) You have a high bias.
- ( ) You have a too low avoidable bias.
- ( ) You have a large data-mismatch problem.

Expand

Incorrect
The difference between the training error and the training-dev error is not high enough to conclude that.

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?

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

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?

- ( ) True

- (●) False

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.

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%

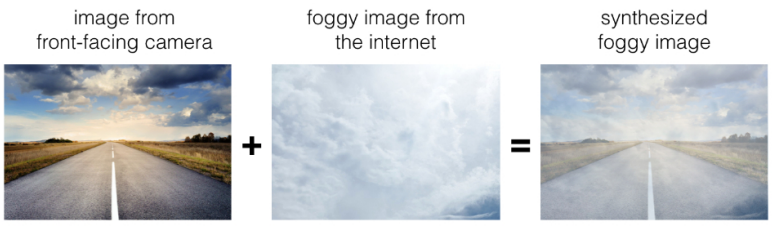
Which of the following statements do you agree with?

- (●) 2.2% would be a reasonable estimate of the maximum 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 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.

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.

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.

11. After working further on the problem, you've decided to correct the incorrectly labeled data on the dev set. Which

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 point**

- [ ] 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.
- [ ] 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 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.

- [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.

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 stop signs in the road as speed limit signs or not. He has approximately 30,000 examples of each image and 30,000 images without a sign. He thought of using your model and applying transfer learning but then he noticed that you use multi-task learning, hence he can't use your model. True/False? **0 / 1 point**

- ○ False
- ◉ True

Expand

✗ **Incorrect**
When using transfer learning we can remove the last layer, which is one of the aspects that is different from a binary classification problem.

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**

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

Expand

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

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.