Congratulations! You passed!

Grade received 93.33% **Latest Submission Grade** 93.33% **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 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.

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

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.

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).

- (●) Train a basic model and do error analysis.
- ( ) Spend some time searching the internet for the data most similar to the conditions you expect on production.
- ( ) Invest a few days in thinking on potential difficulties, and then some more days brainstorming about possible solutions, before training any model.

- [ ] 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.

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. 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
- [x] 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.

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
- [x] 500 images on which the algorithm made a mistake
- [ ] 10,000 images on which the algorithm made a mistake

Expand

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?

- (selected) Calculate the loss as $\sum \mathcal{L}(\hat{y}_j^{(i)}, y_j^{(i)})$ where the sum goes over all the know components of $y^{(i)}$.
- 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 0.
- Make the missing entries equal to 1.

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.

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

- [x] You have a large data-mismatch problem because your model does a lot better on the training-dev set than on the dev set

Correct

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

- [ ] 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.

- [x] You have a large avoidable-bias problem because your training error is quite a bit higher than the human-level error.

✓ Correct

- [ ] You have a large variance problem because your training error is quite higher than the human-level error.

Expand

Correct
Great, you got all the right answers.

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 probably higher than for the train distribution. True/False?

- ○ True
- ◉ False

Expand

Correct

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?

- ( ) Since 8.2 > 4.1 + 2.0 + 1.0, the priority should be to get more images with partially occluded elements.
- ( ) 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.

Expand

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?

- (selected) 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.

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. 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:

0 / 1 point

Which of the following do you agree with?

- With this technique, we duplicate the size of the training set by synthesizing a new foggy image for each image in the training set.

- [ ] If used, the synthetic data should be added to 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.
- [x] If used, the synthetic data should be added to the training/dev/test sets in equal proportions.

Expand

Incorrect

No. The synthetic data 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 on the dev set. Which of these statements do you agree with? (Check all that apply).

1 / 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 efficient.

- [ ] 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.

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.

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?

1 / 1 point

- (•) False
- ( ) True

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. 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

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

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. 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

- (selected) There is a large dataset available.
- It requires less computational resources.
- 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. 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).

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