Your grade: 100%

Your latest: **100%** • Your highest: **100%** • To pass you need at least 80%. We keep your highest score.

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 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.

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

   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.

   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).

   - ◯ 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.
   - ◉ Train a basic model and proceed with error analysis.

   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.

   1 / 1 point

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, which of the following gives you the most appropriate activation function?

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

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

$\begin{pmatrix}0\\?\end{pmatrix}$

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
- [x] True

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

6. Assume you've finally chosen the following split between the data:

1 / 1 point

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

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)	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 bias.

Correct
Correct. The avoidable bias is significantly high since the training error is a lot higher than the human-level error.

7. Assume you've finally chosen the following split between the data:

1 / 1 point

- ○ False
- ◉ True
- ○

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?

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

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%

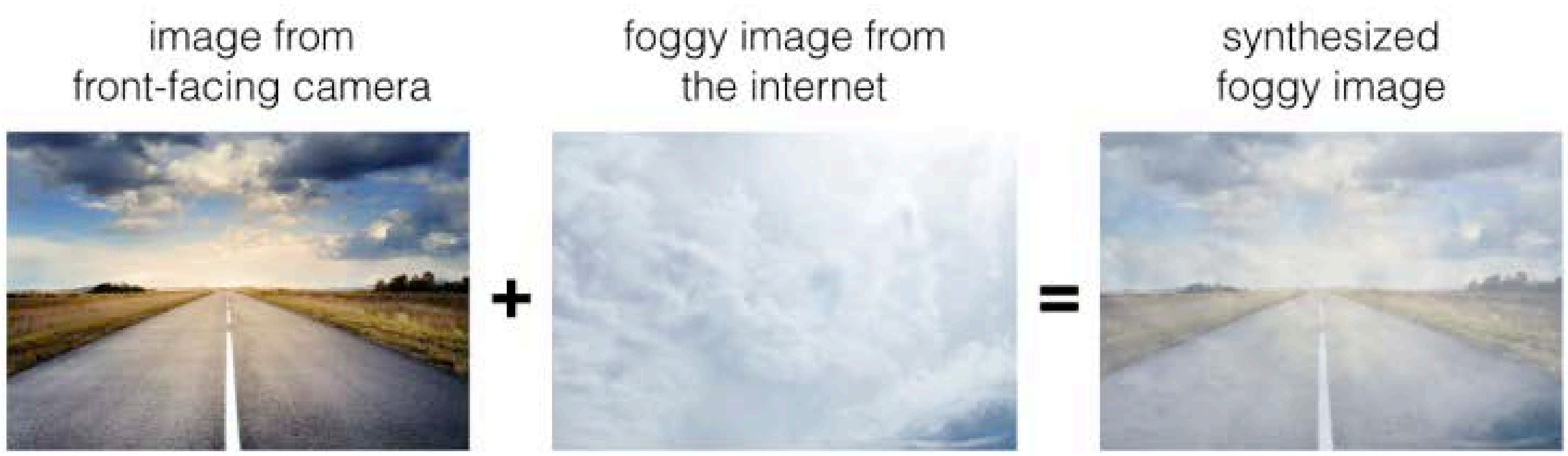
In this table, 4.1%, 7.2%, etc. are a fraction of the total dev set (not just examples of your algorithm mislabeled). For example, about 7.2/15.3 = 47% of your errors are due to partially occluded elements.

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 at least 7.2% of the dev set error.
- ◉ 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%.

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**



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

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. Your team corrects the labels of the wrongly predicted images on the dev set. **1 / 1 point**

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?

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

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.

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?

- ○ 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.
- ○ Recommend that she try multi-task learning instead of transfer learning using all the data.
- ◉ She should try using weights pre-trained on your dataset, and fine-tuning further with the yellow-light dataset.
- ○ You cannot help her because the distribution of data you have is different from hers, and is also lacking the yellow label.

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 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?

1 / 1 point

- ( ) True
- (•) False

Correct
Correct. When using transfer learning we can remove the last layer. That 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

- ( ) 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.

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).

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

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.