Congratulations! You passed!

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

Go to next item Retake the assignment in **23h 49m**

1. This example is adapted from a real production application, but with details disguised to protect confidentiality. **0 / 1 point**

You are a famous researcher in the City of Peacetopia. The people of Peacetopia have a common characteristic: they are afraid of birds. To save them, you have **to build an algorithm that will detect any bird flying over Peacetopia** and alert the population.

The City Council gives you a dataset of 10,000,000 images of the sky above Peacetopia, taken from the city's security cameras. They are labeled:

- y = 0: There is no bird on the image
- y = 1: There is a bird on the image

Your goal is to build an algorithm able to classify new images taken by security cameras from Peacetopia.

There are a lot of decisions to make:

- What is the evaluation metric?
- How do you structure your data into train/dev/test sets?

Metric of success

The City Council tells you the following that they want an algorithm that

1. Has high accuracy.
2. Runs quickly and takes only a short time to classify a new image.
3. Can fit in a small amount of memory, so that it can run in a small processor that the city will attach to many different security cameras.

You meet with them and ask for just one evaluation metric. True/False?

↙↗ **Expand**

⊗ **Incorrect**
No. More than one metric expands the choices and tradeoffs you have to decide for each with unknown effects on the other two.

2. The city revises its criteria to: **1 / 1 point**

- "We **need** an algorithm that can let us know a bird is flying over Peacetopia as accurately as possible."

- "We *want* the trained model to take no more than 10 sec to classify a new image."
- "We *want* the model to fit in 10MB of memory."

Given models with different accuracies, runtimes, and memory sizes, how would you choose one?

↙↗ **Expand**

✓ **Correct**
Yes. Once you meet the runtime and memory thresholds, accuracy should be maximized.

3. Based on the city's requests, which of the following would you say is true? **1 / 1 point**

Expand

Correct

4. With 10,000,000 data points, what is the best option for train/dev/test splits? **1 / 1 point**

Expand

Correct
Yes. The size of the data set allows for bias and variance evaluation with smaller data sets.

5. Now that you've set up your train/dev/test sets, the City Council comes across another 1,000,000 images from social media and offers them to you. These images are different from the distribution of images the City Council had originally given you, but you think it could help your algorithm. You should add the citizens' data to the training set. True/False?

0 / 1 point

Expand

⊗ **Incorrect**

No. Adding this data to the training set will change the training set distribution. However, it is not a problem to have different training and dev distributions. In contrast, it would be very problematic to have different dev and test set distributions.

6. One member of the City Council knows a little about machine learning, and thinks you should add the 1,000,000 citizens' data images to the test set. You object because: **1 / 1 point**

Expand

Correct
Great, you got all the right answers.

7. Human performance for identifying birds is < 1%, training set error is 5.2% and dev set error is 7.3%. Which of the options below is the best next step? **1 / 1 point**

Expand

Correct
Yes. Avoidable bias is >4.2% which is larger than the 2.1% variance.

8. If your goal is to have “human-level performance” be a proxy (or estimate) for Bayes error, how would you define “human-level performance”? **1 / 1 point**

Expand

Correct

Yes. This is the peak of human performance in this task.

9. Which of the below shows the optimal order of accuracy from worst to best? **1 / 1 point**

Expand

Correct

Yes. A learning algorithm's performance can be better than human-level performance but it can never be better than Bayes error.

10. You find that a team of ornithologists debating and discussing an image gets an even better 0.1% performance, so you define that as “human-level performance.” After working further on your algorithm, you end up with the following:

1 / 1 point

Human-level performance	0.1%
Training set error	2.0%
Dev set error	2.1%

Based on the evidence you have, which two of the following four options seem the most promising to try? (Check two options.)

Expand

Correct
Great, you got all the right answers.

11. After running your model with the test set you find it is a 7.0% error compared to a 2.1% error for the dev set and 2.0% for the training set. What can you conclude? (Choose all that apply)

1 / 1 point

Expand

Correct
Great, you got all the right answers.

12. After working on this project for a year, you finally achieve: Human-level performance, 0.10%, Training set error, 0.05%, Dev set error, 0.05%. Which of the following are true? (Check all that apply.)

1 / 1 point

Expand

Correct
Great, you got all the right answers.

13. Your system is now very accurate but has a higher false negative rate than the City Council of Peacetopia would like. What is your best next step? **0 / 1 point**

↙↗ **Expand**

⊗ **Incorrect**
No. You must maintain accuracy and include false negatives.

14. You've handily beaten your competitor, and your system is now deployed in Peacetopia and is protecting the citizens from birds! But over the last few months, a new species of bird has been slowly migrating into the area, so the performance of your system slowly degrades because your data is being tested on a new type of data. **1 / 1 point**

You have only 1,000 images of the new species of bird. The city expects a better system from you within the next 3 months. Which of these should you do first?

Expand

Correct

15. The City Council thinks that having more Cats in the city would help scare off birds. They are so happy with your work on the Bird detector that they also hire you to build a Cat detector. You have a huge dataset of 100,000,000 cat images. Training on this

1 / 1 point

data takes about two weeks. Which of the statements do you agree with? (Check all that agree.)

Expand

Correct

Great, you got all the right answers.