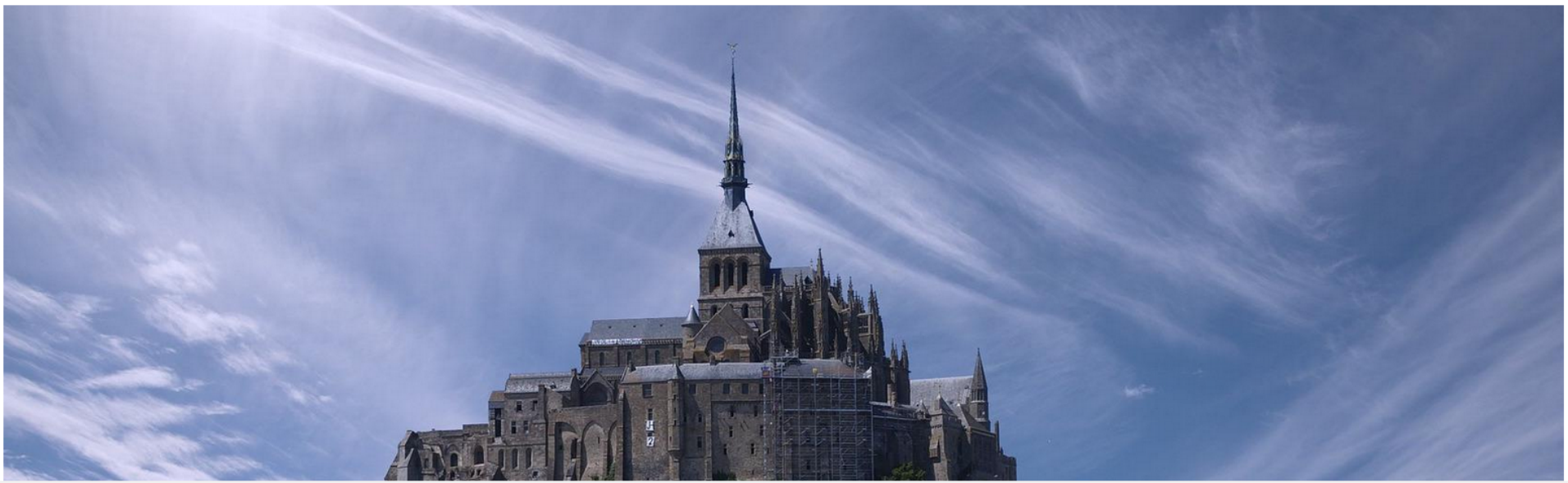
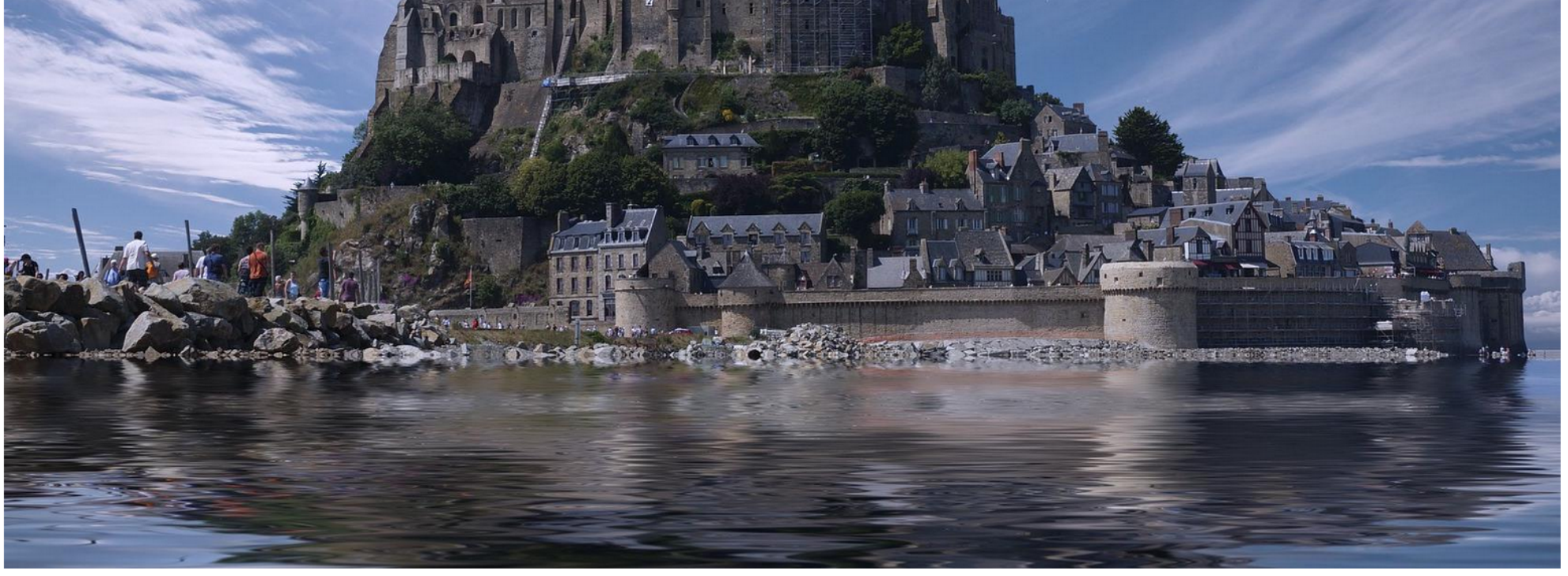
Congratulations! You passed!

Grade received 93.33% **Latest Submission Grade** 93.33% **To pass** 80% or higher

Go to next item

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

- (●) True:
- ( ) False

Expand

Correct

Yes. The goal is to have one metric that focuses the development effort and increases iteration velocity.

2. The city asks for your help in further defining the criteria for accuracy, runtime, and memory. How would you suggest they identify the criteria?

1 / 1 point

- Suggest that they purchase more infrastructure to ensure the model runs quickly and accurately.
- (●) Suggest to them that they define which criterion is most important. Then, set thresholds for the other two.
- Suggest to them that they focus on whichever criterion is important and then eliminate the other two.

Expand

Correct

Yes. The thresholds provide a way to evaluate models head to head.

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

- Accuracy, running time and memory size are all optimizing metrics because you want to do well on all three.
- Accuracy, running time and memory size are all satisfying metrics because you have to do sufficiently well on all three for your system to be acceptable.
- Accuracy is a satisfying metric; running time and memory size are an optimizing metric.
- (selected) Accuracy is an optimizing metric; running time and memory size are satisfying metrics.

Expand

Correct

4. Structuring your data **1 / 1 point**

Before implementing your algorithm, you need to split your data into train/dev/test sets. Which of these do you think is the best choice?

Train	Dev	Test
6,000,000	1,000,000	3,000,000

Train	Dev	Test
6,000,000	3,000,000	1,000,000

Train	Dev	Test
3,333,334	3,333,334	3,333,334

Train	Dev	Test
9,500,000	250,000	250,000

Expand

Correct
Yes.

5. After setting up your train/dev/test sets, the City Council comes across another 1,000,000 images, called the “citizens’ data”. Apparently the citizens of Peacetopia are so scared of birds that they volunteered to take pictures of the sky and label them, thus contributing these additional 1,000,000 images. These images are different from the distribution of images the City Council had originally given you, but you think it could help your algorithm. **1 / 1 point**

Notice that adding this additional data to the training set will make the distribution of the training set different from the distributions of the dev and test sets.

Is the following statement true or false?

"You should not add the citizens' data to the training set, because if the training distribution is different from the dev and test sets, then this will not allow the model to perform well on the test set."

- ○ True
- ◉ False

Expand

Correct

False is correct: Sometimes we'll need to train the model on the data that is available, and its distribution may not be the same as the data that will occur in production. Also, adding training data that differs from the dev set may still help the model improve performance on the dev set. What matters is that the dev and test set have the same distribution.

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 proportionately to the train/dev/test sets. You object because: **1 / 1 point**

- (●) If we add the images to the test set then it won't reflect the distribution of data expected in production.
- ( ) The 1,000,000 citizens' data images do not have a consistent x-->y mapping as the rest of the data.
- ( ) The training set will not be as accurate because of the different distributions.
- ( ) The additional data would significantly slow down training time.

Expand

Correct

Yes. Using the data in the training set could be beneficial, but you wouldn't want to include such images in your test set as they are not from the expected distribution of data you'll see in production.

7. You train a system, and its errors are as follows (error = 100%-Accuracy): **1 / 1 point**

Training set error	4.0%
Dev set error	4.5%

This suggests that one good avenue for improving performance is to train a bigger network so as to drive down the 4.0% training error. Do you agree?

- No, because this shows your variance is higher than your bias.
- No, because there is insufficient information to tell.
- Yes, because this shows your bias is higher than your variance.
- Yes, because having a 4.0% training error shows you have a high bias.

Expand

Correct

8. You ask a few people to label the dataset so as to find out what is human-level performance. You find the following levels of accuracy: **1 / 1 point**

Bird watching expert #1	0.3% error
Bird watching expert #2	0.5% error
Normal person #1 (not a bird watching expert)	1.0% error
Normal person #2 (not a bird watching expert)	1.2% error

If your goal is to have "human-level performance" be a proxy (or estimate) for Bayes error, how would you define "human-level performance"?

- (●) 0.3% (accuracy of expert #1)
- ( ) 0.4% (average of 0.3 and 0.5)
- ( ) 0.0% (because it is impossible to do better than this)
- ( ) 0.75% (average of all four numbers above)

Expand

Correct

9. Which of the following statements do you agree with? 1 / 1 point

- A learning algorithm's performance can never be better than human-level performance nor better than Bayes error.
- A learning algorithm's performance can never be better than human-level performance but it can be better than Bayes error.
- A learning algorithm's performance can be better than human-level performance and better than Bayes error.
- A learning algorithm's performance can be better than human-level performance but it can never be better than Bayes error.

Expand

Correct

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.)

- [x] Train a bigger model to try to do better on the training set.

  ✓ **Correct**

- [ ] Get a bigger training set to reduce variance.

- [x] Try decreasing regularization.

  ✓ **Correct**

- [ ] Try increasing regularization.

Expand

Correct
Great, you got all the right answers.

11. You also evaluate your model on the test set, and find the following: **1 / 1 point**

Human-level performance	0.1%
Training set error	2.0%
Dev set error	2.1%
Test set error	7.0%

What does this mean? (Check the two best options.)

- [ ] You have underfitted to the dev set.
- [x] You have overfit to the dev set.

Correct

- [ ] You should get a bigger test set.
- [x] You should try to get a bigger dev set.

Correct

Expand

Correct

Great, you got all the right answers.

12. After working on this project for a year, you finally achieve: **1 / 1 point**

Human-level performance	0.10%
Training set error	0.05%
Dev set error	0.05%

What can you conclude? (Check all that apply.)

- [x] It is now harder to measure avoidable bias, thus progress will be slower going forward.

  ✓ **Correct**

- [ ] This is a statistical anomaly (or must be the result of statistical noise) since it should not be possible to surpass human-level performance.
- [x] If the test set is big enough for the 0.05% error estimate to be accurate, this implies Bayes error is ≤ 0.05

  ✓ **Correct**

- [ ] With only 0.05% further progress to make, you should quickly be able to close the remaining gap to 0%

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?

1 / 1 point

- ○ Pick false negative rate as the new metric, and use this new metric to drive all further development.
- ◉ Reset your "target" (metric) for the team and tune to it.
- ○ Expand your model size to account for more corner cases.
- ○ Look at all the models you've developed during the development process and find the one with the lowest false negative error rate.

Expand

Correct

Yes. The target has shifted so an updated metric is required.

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

0 / 1 point

- ○ Add pooling layers to downsample features to accommodate the new species.
- ○ Augment your data to increase the images of the new bird.
- ○ Put the new species' images in training data to learn their features.
- ◉ Split them between dev and test and re-tune.

Expand

⊗ **Incorrect**

No. The First you'll need more data so augmenting the existing data to create more training examples would be the next step

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. (Wow Cat detectors are just incredibly useful, aren't they?) Because of years of working on Cat detectors, you have such a huge dataset of 100,000,000 cat images that training on this data takes about two weeks. Which of the statements do you agree with? (Check all that agree.)

1 / 1 point

- [x] If 100,000,000 examples is enough to build a good enough Cat detector, you might be better off training with just 10,000,000 examples to gain a $\approx$10x improvement in how quickly you can run experiments, even if each model performs a bit worse because it's trained on less data.

  ✓ **Correct**

- [x] Buying faster computers could speed up your teams' iteration speed and thus your team's productivity.

  ✓ **Correct**

- [x] Needing two weeks to train will limit the speed at which you can iterate.

  ✓ **Correct**

- [ ] Having built a good Bird detector, you should be able to take the same model and hyperparameters and just apply it to the Cat dataset, so there is no need to iterate.

Expand

Correct

Great, you got all the right answers.