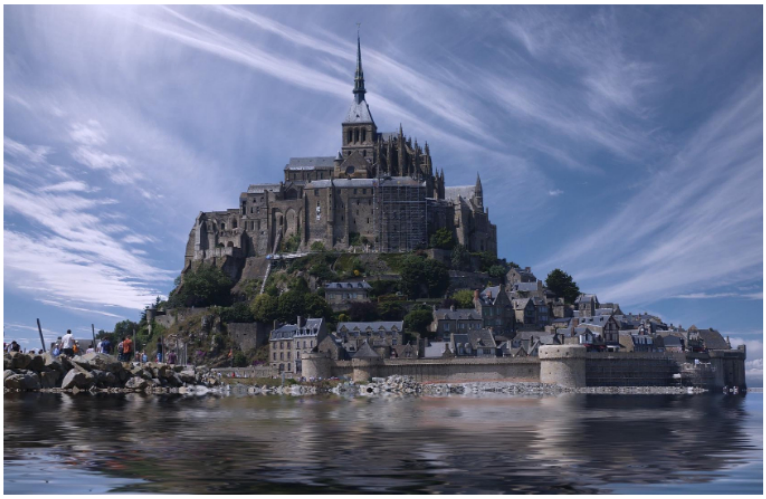
Try again once you are ready

Grade received 66.66%
Latest Submission Grade 66.67%
To pass 80% or higher

Try again

1. Problem Statement 0 / 1 point

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



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 are delighted because this list of criteria will speed development and provide guidance on how to evaluate two different algorithms. True/False?

- (●) True:
- ( ) False

Expand

Incorrect
No. 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. The essential difference between an optimizing metric and satisficing metrics is the priority assigned by the stakeholders. True/False? **1 / 1 point**

- True
- False

Expand

Correct
Yes. Satisficing metrics have thresholds for measurement and an optimizing metric is unbounded.

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

- train - 60%, dev - 30%, test - 10%
- train - 33.3%, dev - 33.3%, test - 33.3%
- train - 95%, dev - 2.5%, test - 2.5%
- train - 60%, dev - 10%, test - 30%

Expand

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

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. **0 / 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

Incorrect

True is incorrect: 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 training set will not be as accurate because of the different distributions.
- ( ) The 1,000,000 citizens' data images do not have a consistent x-->y mapping as the rest of the data.
- ( ) 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?

- ( ) 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.
- ( ) No, because this shows your variance is higher than your bias.
- (●) No, because there is insufficient information to tell.

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"?

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

Expand

Correct

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

- (•) A learning algorithm's performance can be better than human-level performance but it can never 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 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.

Expand

Correct

10. Which of the following best expresses how to evaluate the next steps in your project when your results for human-level performance, train, and dev set error are 0.1%, 2.0%, and 2.1% respectively? **1 / 1 point**

- ( ) Keep tuning until the train set accuracy is equal to human-level performance because it is the optimizing metric.
- ( ) Port the code to the target devices to evaluate if your model meets or exceeds the satisficing metrics.
- (•) Based on differences between the three levels of performance, prioritize actions to decrease bias and iterate.
- ( ) Evaluate the test set to determine the magnitude of the variance.

Expand

Correct
Yes. Always choose the area with the biggest opportunity for improvement.

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) **0 / 1 point**

- [ ] Try decreasing regularization for better generalization with the dev set.
- [x] You have underfitted to the dev set.

! **This should not be selected**
No. The dev set performance versus the test set indicates it is overfitting.

- [x] You have overfitted to the dev set.

✓ **Correct**
Yes. The dev set performance versus the test set indicates it is overfitting.

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

Expand

Incorrect
You didn't select all the correct 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**

- [ ] This is a statistical anomaly (or must be the result of statistical noise) since it should not be

possible to surpass human-level performance.

- [ ] With only 0.05% further progress to make, you should quickly be able to close the remaining gap to 0%
- [x] You are close to Bayes error and possible overfitting.

Correct
Yes. By definition, Bayes error cannot be exceeded except for overfitting.

- [x] All or almost all of the avoidable bias has been accounted for.

Correct
Yes. Exceeding human performance makes the identification of avoidable bias very challenging.

Expand

Correct
Great, you got all the right answers.

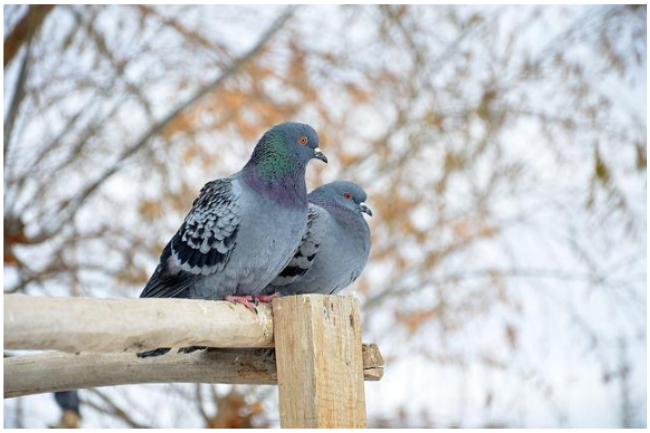
13. Your system is now very accurate but has a higher false negative rate than the City Council of Peacetopia would like. What is your best next step? **0 / 1 point**

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

Expand

Incorrect
No. You must select the correct target. More data will not change the optimizing metric.

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?

- (•) Use the data you have to define a new evaluation metric (using a new dev/test set) taking into account the new species, and use that to drive further progress for your team.
- ( ) Put the 1,000 images into the training set so as to try to do better on these birds.
- ( ) Add the 1,000 images into your dataset and reshuffle into a new train/dev/test split.
- ( ) Try data augmentation/data synthesis to get more images of the new type of bird.

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. (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.)

0 / 1 point

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

  Correct

- [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 ≈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

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

Incorrect
You didn't select all the correct answers