Congratulations! You passed!

Grade received 86.66% | Latest Submission Grade 86.67% | To pass 80% or higher

Go to next item

1. Problem Statement 1 / 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?

- (•) False
- ( ) True:

Expand

Correct
Yes. 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?

- ( ) Accuracy is an optimizing metric, therefore the most accurate model is the best choice.
- (•) Find the subset of models that meet the runtime and memory criteria. Then, choose the highest accuracy.
- ( ) Create one metric by combining the three metrics and choose the best performing model.

- Take the model with the smallest runtime because that will provide the most overhead to increase accuracy.

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**

- 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, running time and memory size are all optimizing metrics because you want to do well on all three.
- (•) Accuracy is an optimizing metric; running time and memory size are satisfying metrics.
- Accuracy is a satisfying metric; running time and memory size are an optimizing metric.

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
9,500,000	250,000	250,000

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

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: 0 / 1 point

- (●) 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.
- ( ) If we add the images to the test set then it won't reflect the distribution of data expected in production.
- ( ) The additional data would significantly slow down training time.

Expand

Incorrect
No. The important issue is mixing distributions.

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

- (●) Train a bigger network to drive down the >4.0% training error.
- ( ) Validate the human data set with a sample of your data to ensure the images are of sufficient quality.
- ( ) Try an ensemble model to reduce bias and variance.
- ( ) Get more data or apply regularization to reduce variance.

Expand

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

8. You want to define what human-level performance is to the city council. Which of the following is the best answer? 1 / 1 point

- ( ) The average of regular citizens of Peacetopia (1.2%).
- ( ) The average of all the numbers above (0.66%).
- (●) The performance of their best ornithologist (0.3%).
- ( ) The average performance of all their ornithologists (0.5%).

Expand

Correct
Yes. The best human performance is closest to Bayes' error.

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

- (●) Human-level performance -> the learning algorithm's performance -> Bayes error.
- ( ) The learning algorithm's performance -> Bayes error -> human-level performance.
- ( ) Human-level performance -> Bayes error -> the learning algorithm's performance.
- ( ) The learning algorithm's performance -> human-level performance -> Bayes error.

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.)

- [ ] Get a bigger training set to reduce variance.
- [x] Try decreasing regularization.

  ✓ Correct

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

  ✓ Correct

- [ ] Try increasing regularization.

Expand

✓ Correct
Great, you got all the right answers.

11. You've now also run your model on the test set and find that it is a 7.0% error compared to a 2.1% error for the dev set. What should you do? (Choose all that apply) **1 / 1 point**

- [ ] Try decreasing regularization for better generalization with the dev set.
- [ ] Get a bigger test set to increase its accuracy.
- [x] Try increasing regularization to reduce overfitting to the dev set.

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

- [x] Increase the size of the dev set.

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

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**

- [ ] With only 0.05% further progress to make, you should quickly be able to close the remaining gap to 0%
- [ ] 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] 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. It turns out Peacetopia has hired one of your competitors to build a system as well. You and your competitor both deliver systems with about the same running time and memory size. However, your system has higher accuracy! Still, when Peacetopia tries out both systems, they conclude they like your competitor's system better because, even though you have higher overall accuracy, you have more false negatives (failing to raise an alarm when a bird is in the air). What should you do? **1 / 1 point**

- Apply regularization to minimize the false negative rate.
- Pick false negative rate as the new metric, and use this new metric to drive all further development.
- Ask your team to take into account both accuracy and false negative rate during development.
- Brainstorm with your team to refine the optimizing metric to include false negatives as they further develop the model.

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? **1 / 1 point**

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

Expand

Correct
Yes. A sufficient number of images is necessary to account for the new species.

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 data takes about two weeks. Which of the statements do you agree with? (Check all that agree.) **0 / 1 point**

- [x] You could consider a tradeoff where you use a subset of the cat data to find reasonable performance with reasonable iteration pacing.

Correct
Yes. This is similar to satisficing metrics where "good enough" determines the size of the data.

- [ ] Accuracy should exceed the City Council's requirements but the project may take as long as the bird detector because of the two week training/iteration time.
- [x] Given a significant budget for cloud GPUs, you could mitigate the training time.

Correct
Yes. More resources will allow you to iterate faster.

- [ ] With the experience gained from the Bird detector you are confident to build a good Cat detector on the first try.

Expand

Incorrect
You didn't select all the correct answers