Congratulations! You passed!

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

Go to next item

1. Problem Statement 1 / 1 points

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

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 points

- "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 10Kb of memory."

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

- Create one metric by combining the three metrics and choose the best performing model.
- Find the subset of models that meet the runtime and memory criteria. Then, choose the highest accuracy.
- Accuracy is an optimizing metric. Therefore the most accurate model is the best choice.
- 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. Which of the following best answers why it is important to identify optimizing and satisficing metrics? 1 / 1 points

- It isn't. All metrics must be met for the model to be acceptable.
- Knowing the metrics provides input for efficient project planning.
- Identifying the optimizing metric informs the team which models they should try first.
- Identifying the metric types sets thresholds for satisficing metrics. This provides explicit evaluation criteria.

Expand

Correct
Yes. Thresholds are essential for evaluation of key use case constraints.

4. Structuring your data 1 / 1 points

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
	3,333,334	3,333,334	3,333,334

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

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

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

Expand

Correct
Yes.

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. Which of the following is the best use of that additional data? 0 / 1 points

- Split it among train/dev/test equally.
- Add it to the training set.
- Do not use the data. It will change the distribution of any set it is added to.
- Add it to the dev set to evaluate how well the model generalizes across a broader set.

Expand

Incorrect
No. Different dev and test distributions would be an issue in evaluating variance because of the different 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 proportionally to the train/dev/test sets. You object because 0 / 1 points

- The training set will not be an accurate because of the different distributions.
- The additional data would significantly slow down training time.
- 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.

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? 0 / 1 points

- Validate the human data set with a sample of your data to ensure the images are of sufficient quality.
- Get more data or apply regularization to reduce variance.
- Try an ensemble model to reduce bias and variance.
- Train a bigger network to drive down the +5.0% training error.

Expand

Incorrect
No. We do not have test set accuracy.

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 points

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

Expand

Correct

9. A learning algorithm's performance can be better than human-level performance but it can never be better than Bayes error. True/False? 1 / 1 points

- False.
- True.

Expand

Correct
Yes. By definition, human level error is worse than Bayes error.

10. After working on your algorithm you have to decide the next steps. Currently, human-level performance is 0.2%, training is at 2.0% and the dev set is at 2.1%. Which statement below best describes your thought process? 1 / 1 points

- [x] Decrease regularization to boost smaller signals.
  - Correct
    Yes. Bias is higher than variance.
- [ ] Decrease variance via regularization so training and dev sets have similar performance.
- [ ] Get a bigger training set to reduce variance.
- [x] Address bias first through a larger model to get closest to human level error.
  - Correct
    Yes. Selecting the largest difference from (train set error - human level error) and (dev set error - train set error) and reducing bias or variance accordingly is the most productive step.

Expand

Correct
Great, you got all the right answers.

11. After running your model with the best 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 points

- [x] You have overfitted to the dev set.
  - Correct
    Yes. The dev set performance versus the test set indicates it is overfitting.
- [x] You should try to get a bigger dev set.
  - Correct
    Yes. The dev set performance versus the test set indicates it is overfitting.
- [ ] Try decreasing regularization for better generalization with the dev set.
- [ ] You have underfitted to the dev set.

Expand

Correct
Great, you got all the right answers.

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

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
    With only 0.05% further progress to make, you should quickly be able to close the remaining gap to 0%
- [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
- [ ] This is a statistical anomaly (or must be the result of statistical noise) since it should not be possible to surpass human level performance.

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! Bad news, 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 points

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

Expand

Correct
Yes. The target has shifted as an updated metric is required.

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 model 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 points

- Augment your data to increase the images of the new bird.
- Put them into the dev set to evaluate the bias and in tune.
- Add the new images and split them among train/dev/test.
- Add hidden layers to further refine feature development.

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 20,000,000 cat images. Training on this data takes about two weeks. Which of the statements do you agree with? (Check all that agree.) 1 / 1 points

- [x] Lowering the number of images will reduce training time and likely allow for an acceptable trade-off between iteration speed and accuracy.
  - Correct
    Yes. There is a sweet spot that allows development at a reasonable rate without significant accuracy loss.
- [ ] Reducing the model complexity will allow the use of the larger data set but preserve accuracy.
- [x] This significantly impacts iteration speed.
  - Correct
    Yes. This training time is an absolute constraint on iteration.

Expand

Correct
Great, you got all the right answers.