Quiz, 15 questions



Congratulations! You passed!

Next Item



1/1 points

1

Problem Statement

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 labelled:

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

Quiz, 15 questions

- 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.

<u>Note</u>: Having three evaluation metrics makes it harder for you to quickly choose between two different algorithms, and will slow down the speed with which your team can iterate. True/False?

0	True			
Corr	ect			
	False			



1/1 points

2.

After further discussions, the city narrows down its criteria to:

- "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 10sec to classify a new image."
- "We want the model to fit in 10MB of memory."

If you had the three following models, which one would you choose?

Test Accuracy	Runtime	Memory size
97%	1 sec	ЗМВ
Test Accuracy	Runtime	Memory size
99%	13 sec	9MB
Test Accuracy	Runtime	Memory size
97%	3 sec	2MB

Birdrecognition in the city of Peacetopia (case study)Memory size				
Quiz, 15 questior 9 8%	9 sec	9MB		

Correct

Correct! As soon as the runtime is less than 10 seconds you're good. So, you may simply maximize the test accuracy after you made sure the runtime is <10sec.



1/1 points

3.

Based on the city's requests, which of the following would you say is true?

Accuracy is an optimizing metric; running time and memory size are a satisficing metrics.

Correct

A	!		_ +
Accuracy is a satisficing metric; running time and memo	ory size are	e an optimizing m	etric.

- 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 satisficing metrics because you have to do sufficiently well on all three for your system to be acceptable.



1/1 points

4.

Structuring your data

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
9,500,000	250,000	250,000

Correct

Yes.

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

15 0110	stions	20.	
, 15 que	stions 3,333,334	3,333,333	3,333,333
	Train	Dev	Test
	6,000,000	1,000,000	3,000,000
citizens of the s he dist ou sho distribu	or data". Apparently the citizens of Pook yand label them, thus contributing ribution of images the City Council local not add the citizens' data to the	e City Council comes across another eacetopia are so scared of birds tha g these additional 1,000,000 images. nad originally given you, but you thin e training set, because this will cause ting dev and test set performance. T	t they volunteered to take pict These images are different fronk it could help your algorithm to the training and dev/test set
	False		
have	ct ng this data to the training set will cl	hange the training set distribution. Hon. On the contrary, it would be ver	-
Addir have	ct ng this data to the training set will cl different training and dev distributi	_	-
Addir have differ	ng this data to the training set will cl different training and dev distributi ent dev and test set distributions. 1 / 1 points	on. On the contrary, it would be ver	y problematic to have
Addir have differ	ng this data to the training set will cl different training and dev distributi ent dev and test set distributions. 1/1 points mber of the City Council knows a lit data images to the test set. You ob	on. On the contrary, it would be ver ttle about machine learning, and thin ject because: set distributions to become differen	y problematic to have
Addir have differ	ng this data to the training set will cledifferent training and dev distributions. 1/1 points This would cause the dev and test syou're not aiming where you want so the set will clear the content of the content of the content of the test set.	on. On the contrary, it would be ver ttle about machine learning, and thin ject because: set distributions to become differen	y problematic to have
Addir have differ	ng this data to the training set will cledifferent training and dev distributions. 1/1 points mber of the City Council knows a lite data images to the test set. You obtain the deviand test syou're not aiming where you want set.	on. On the contrary, it would be ver ttle about machine learning, and thin ject because: set distributions to become differen	y problematic to have nks you should add the 1,000 t. This is a bad idea because

Quiz, 15 questions **Correct**

The 1,000,000 citizens' data images do not have a consistent x>y mapping as the rest of the data
(similar to the New York City/Detroit housing prices example from lecture).

Un-selected is correct



1/1 points

7.

You train a system, and its errors are as follows (error = 100%-Accuracy):

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 having 4.0% training error shows you have high bias.
. //	165, Decause Having 4.0% Calling et lot Shows you have high blas.
	, , , , , , , , , , , , , , , , , , , ,

Yes, because this shows your bias is higher than your variance.

No, because this shows your variance is higher than your bias.

No, because there is insufficient information to tell.

Correct



1/1 points

8.

You ask a few people to label the dataset so as to find out what is human-level performance. You find the Birdoregognitical in the city of Peacetopia (case study)

Quiz, 15 qu	estions	
	d watching expert #1	0.3% error
Bir	d watching expert #2	0.5% error
No	rmal person #1 (not a bird watching expert)	1.0% error
No	rmal person #2 (not a bird watching expert)	1.2% error
-	goal is to have "human-level performance" be a proxy (or estimate) for Bayes err n-level performance"?	or, how would you define
	0.0% (because it is impossible to do better than this)	
0	0.3% (accuracy of expert #1)	
Corr	ect	
	0.4% (average of 0.3 and 0.5)	
	0.75% (average of all four numbers above)	
9.	1 / 1 points	
	of the following statements do you agree with?	
0	A learning algorithm's performance can be better than human-level performance better than Bayes error.	ce but it can never be
Corr	ect	
	A learning algorithm's performance can never be better than human-level performance than Bayes error.	rmance but it can be
	A learning algorithm's performance can never be better than human-level performance can never be better than human-level performance.	rmance nor better than
	A learning algorithm's performance can be better than human-level performance error.	ce and better than Bayes

/

1/1 points Bird recognition in the city of Peacetonia (case study) an even better 0.1% performance, so Quixous deciments as "human-level performance." After working further on your algorithm, you end up with the following:

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.	
Un-selected is correct	
Train a bigger model to try to do better on the training set.	
Correct	
Try increasing regularization. Un-selected is correct	
on-selected is correct	
Try decreasing regularization.	



Correct

1/1 points

11

You also evaluate your model on the test set, and find the following:

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 should get a bigger test set. Bird recognition in the city of Peacetopia (case study)

Quiz, ¹5 fixe rect is correct

You should try to get a bigger dev set.
Correct
You have underfit to the dev set.
Un-selected is correct
You have overfit to the dev set. Correct

1/1 points

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%

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

	With only 0.09% further progress to make, you should quickly be able to close the remaining gap to 0%
Un-s	elected is correct
	If the test set is big enough for the 0.05% error estimate to be accurate, this implies Bayes error is ≤ 0.05
Corr	ect
	It is now harder to measure avoidable bias, thus progress will be slower going forward.

Correct

Quiz, 15 questions

Un-selected is correct



1/1 points

13.

It turns out Peacetopia has hired one of your competitors to build a system as well. Your system and your competitor both deliver systems with about the same running time and memory size. However, your system has higher accuracy! However, when Peacetopia tries out your and your competitor's systems, they conclude they actually 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?

	Look at all the models you've developed during the development process and find the one with the lowest false negative error rate.
	Ask your team to take into account both accuracy and false negative rate during development.
0	Rethink the appropriate metric for this task, and ask your team to tune to the new metric.
Correct	

Pick false negative rate as the new metric, and use this new metric to drive all further development.



1/1 points

14.

You've handily beaten your competitor, and your system is now deployed in Peacetopia and is protecting the Biticle and I will be a solution of the Biticle and I will be a solut



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.

Correct

Put the 1,000 images into the training set so as to try to do better on these birds.
Try data augmentation/data synthesis to get more images of the new type of bird.
Add the 1,000 images into your dataset and reshuffle into a new train/dev/test split



1/1 points

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.)

Needing two weeks to train will limit the speed at which you can iterate.

Quiz, 15 questions			
	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.		
	Un-selected is correct		
	Buying faster computers could speed up your teams' iteration speed and thus your team's productivity.		
Corre	ect		
Corre	If 100,000,000 examples is enough to build a good enough Cat detector, you might be better of 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.		