



✓ Congratulations! You passed!

TO PASS 80% or higher



GRADE 100%

Bird recognition in the city of Peacetopia (case study)

LATEST SUBMISSION GRADE 100%

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 labelled:

- 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 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.

Note: 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?

True

○ False

✓ Correct

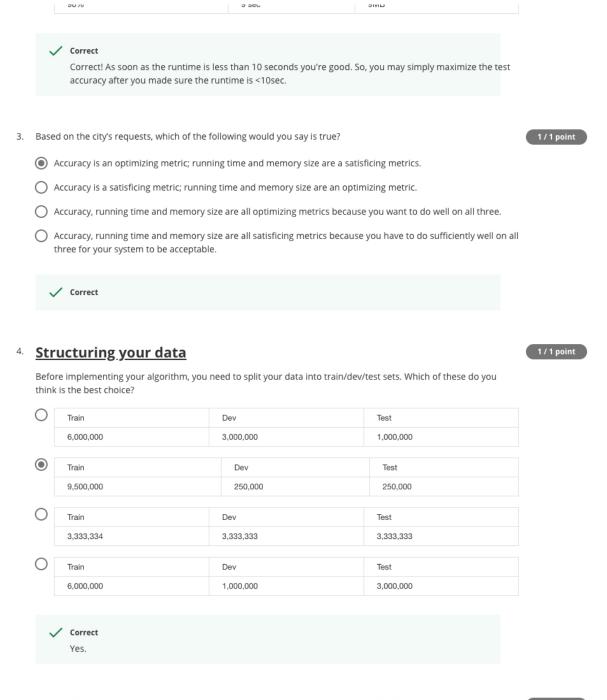
- 2. After further discussions, the city narrows down its criteria to:
 - "We need an algorithm that can let us know a bird is flying over Peacetopia as accurately as possible."
 - "We want the trained model to take no more than 10sec to classify a new image."
 - "We want the model to fit in 10MB of memory."

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

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

$\overline{}$	lest Accuracy	nultille	Welliory Size	
	97%	3 sec	2MB	
•	Test Accuracy	Runtime	Memory size	
	2000	9 000	OMP	

1 / 1 point



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

1/1 point

anterent from the distribution of images the city council had originally given you, but you think it could help your algorithm. 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 ✓ 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 to the test set. You object because: 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). A bigger test set will slow down the speed of iterating because of the computational expense of evaluating models on the test set. This would cause the dev and test set distributions to become different. This is a bad idea because you're not aiming where you want to hit. Correct The test set no longer reflects the distribution of data (security cameras) you most care about. Correct 7. You train a system, and its errors are as follows (error = 100%-Accuracy): 1 / 1 point

4.0% Training set error 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.		
	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		
8.	You ask a few people to label the dataset so as to find out what is human-level perfor following levels of accuracy:	rmance. You find the	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 edefine "human-level performance"?	error, how would you	
	0.0% (because it is impossible to do better than this)		
	0.3% (accuracy of expert #1)		
	0.4% (average of 0.3 and 0.5)		
	0.75% (average of all four numbers above)		
	✓ 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 Bayes error.	e but it can never be better t	han
	A learning algorithm's performance can never be better than human-level perfor Bayes error.	mance but it can be better t	han
	A learning algorithm's performance can never be better than human-level perfor error.	mance nor better than Baye	25
	A learning algorithm's performance can be better than human-level performance	e and better than Bayes erro	or.

Correc

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.

Try decreasing regularization.

./	Correc

Try increasing regularization.

✓ Train a bigger model to try to do better on the training set.

/	Correc

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%

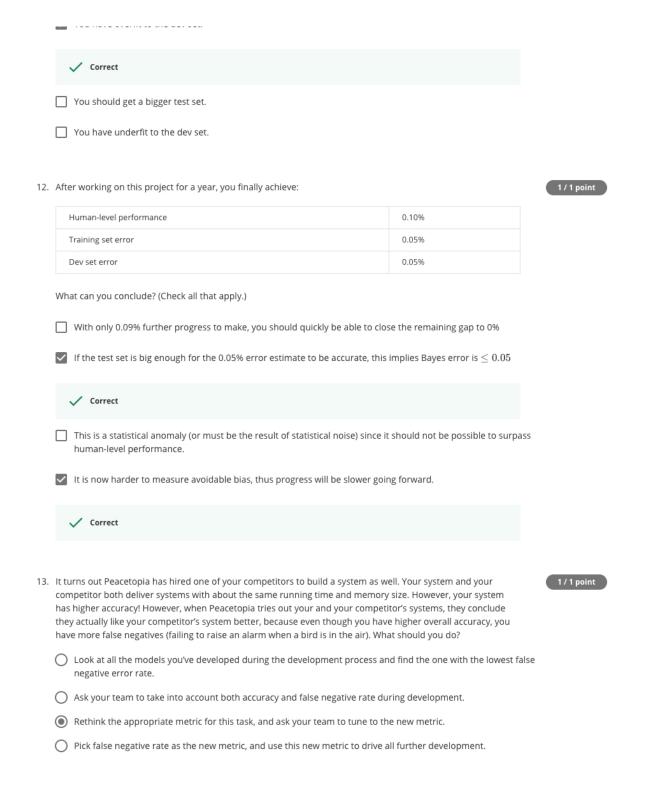
What does this mean? (Check the two best options.)

You should try to get a bigger dev set.



You have overfit to the deviset.

1 / 1 point



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.
- 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.

✓ 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.)

1/1 point

✓	If 100,000,000 examples is enough to build a good enough Cat detector, you might be better of training with 10,000,000 examples to gain a \approx 10x improvement in how quickly you can run experiments, even if each more 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.
	✓ Correct
~	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.