Your grade: 100%

Your latest: **100%** • Your highest: **100%** • To pass you need at least 80%. We keep your highest score.

Next item →

1. This scenario is adapted from a real-world application, with specifics altered for confidentiality. 1 / 1 point

As a distinguished researcher in the City of Peacetopia, you face a unique challenge. The citizens are universally afraid of birds, and your task is to develop an algorithm to detect birds flying over the city and alert the populace.

The City Council has provided you with 10,000,000 sky images from Peacetopia's security cameras, labeled as follows:

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

Your objective is to engineer an algorithm that can accurately classify new security camera images. Decisions regarding the evaluation metric and data structuring into train/dev/test sets are critical.

The City Council has outlined their desires for the algorithm:

1. High accuracy.
2. Quick response time for classifying new images.
3. Low memory requirement to function on small processors across various cameras.

True or False: You realize that having a single evaluation metric will expedite development and ease algorithm comparison.

- ○ False
- ◉ True

Correct
A singular focus on one evaluation metric can significantly speed up development and make it easier to compare and refine algorithms.

2. After further discussions, the city narrows down 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 seconds to classify a new image."
- "We want the model to fit in 10MB of memory."
- "We require a minimum of 98% test accuracy."

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

Test Accuracy	Runtime	Memory size
97%	1 sec	3MB

Test Accuracy	Runtime	Memory size
97%	3 sec	2MB

Test Accuracy	Runtime	Memory size
99%	13 sec	9MB

Test Accuracy	Runtime	Memory size
98%	9 sec	9MB

Correct
The runtime is less than 10 seconds, and the accuracy meets the minimum 98% requirement.

3. **True or False:** The essential difference between an optimizing metric and satisficing metrics is the priority assigned by the stakeholders. **1 / 1 point**

- False
- True

Correct
Satisficing metrics have specific thresholds that need to be met, while an optimizing metric is one we aim to maximize beyond any set threshold.

4. Before implementing your algorithm, you need to split your data into train/dev/test sets. You have a dataset of 10,000,000 examples. Which of these do you think is the best choice? **1 / 1 point**

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

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

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

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

Correct
With a large dataset, smaller dev and test sets are sufficient for evaluating bias and variance.

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 have a different distribution from the images the City Council originally provided, but you think they 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 distribution of the dev and test sets. 1 / 1 point

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
Sometimes we'll need to train the model on available data, even if its distribution differs from the data that will occur in production. Adding training data that differs from the dev set may still help the model improve performance. What matters is that the dev and test sets have the same distribution.

6. One member of the City Council wants to add 1,000,000 citizen data images evenly to the training, development (dev), and test sets. Your original data is from security cameras, and you object because: 1 / 1 point

- ○ 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 (security cameras) expected in production.
- ○ The training set will not be as accurate because of the different distributions.
- ○ The 1,000,000 citizen data images do not have a consistent input-output relationship as the security camera data.

Correct
The test set must accurately represent the real-world data distribution to properly evaluate model performance. Adding citizen data images to the test set would skew this distribution.

7. Human performance for identifying birds is < 1%, training set error is 5.2%, and dev set error is 7.3%. 1 / 1 point

Which of the options below is the best next step?

- ○ Try an ensemble model to reduce bias and variance.
- ○ Get more data or apply regularization to reduce variance.
- ○ Validate the human data set with a sample of your data to ensure the images are of sufficient quality.
- ◉ Train a bigger network to reduce the 5.2% training error.

Correct
Avoidable bias is 4.2%, which is larger than the 2.1% variance, so reducing bias is the priority.

8. You ask a few people to label a bird species dataset to determine human-level performance. The following error rates were recorded: 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 use "human-level performance" as an estimate for Bayes error, how would you define "human-level performance" in this scenario?

- ○ 0.0% (Perfect accuracy, representing an unattainable ideal)
- ○ 0.4% (Average of the two experts' error rates)
- ○ 0.75% (Average of all four error rates)
- ◉ 0.3% (The lowest error rate achieved by an expert)

Correct
The best performance of an expert is the closest practical estimate of Bayes error.

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 but it can be better than Bayes error.
- ○ A learning algorithm's performance can never be better than human-level performance nor better than Bayes error.

Correct
This statement accurately reflects the relationship between the learning algorithm, human-level performance, and Bayes error.

10. After evaluating your algorithm's performance, you need to determine the next steps. **1 / 1 point**

The metrics are:

- Human-level performance: 0.1%
- Training set error: 2.0%
- Dev set error: 2.1%.

Which **two** of the following statements best describe **your most effective thought process**?

- ☑ Address bias by increasing model complexity to reduce the gap between training error and human-level performance.

Correct
Yes, the largest difference (training error - human-level error) indicates high bias, necessitating a focus on increasing model complexity.

- ☑ Decrease regularization to reduce bias and allow the model to capture finer patterns.

Correct
Yes, given the high bias, reducing regularization can help the model fit the training data better.

- ☐ Increase the training dataset size to reduce variance.
- ☐ Decrease variance through increased regularization to align training and dev set performance.

11. You've now also run your model on the test set and find that the error rate is 7.0% compared to a 2.1% error rate 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.
- ☑ Try increasing regularization to reduce overfitting to the dev set.

Correct
Increasing regularization can help reduce overfitting to the dev set.

- ☐ Get a bigger test set to increase its accuracy.
- ☑ Increase the size of the dev set.

Correct
A larger dev set can help provide a more accurate estimate of model performance and reduce overfitting.

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 followingare 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%.
- ☑ All or almost all of the avoidable bias has been accounted for.

Correct
Exceeding human performance strongly suggests that most, if not all, significant sources of avoidable bias have been addressed.

- ☑ You are close to Bayes error and there is a possibility of overfitting.

Correct
By definition, Bayes error cannot be exceeded except for overfitting, and the model's performance indicates it's approaching this limit.

- ☐ It is highly unlikely this result is purely a statistical anomaly, but statistical noise may still contribute to the error.

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! **1 / 1 point**

Still, when Peacetopia tries out both your system and your competitor's system, 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?

- ○ Ask your team to take into account both accuracy and false negative rate during development.
- ○ 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.
- ◉ Rethink the appropriate metric for this task, and ask your team to tune to the new metric.

Correct
The target has shifted so 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 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?

- ○ Add the 1,000 images into your dataset and reshuffle into a new train/dev/test split
- ○ 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.
- ◉ Use the data you have to define a new evaluation metric (using a new dev/test set) that accounts for the new species, and use that metric to guide further improvements.

Correct
The true data distribution has changed. You need to adjust your evaluation to accurately reflect the new data. Because you evaluate your learning algorithm on dev and test sets, adding data only to the training set won't directly improve performance on 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. **1 / 1 point**

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.)

- ☐ Reducing the model complexity will allow the use of the larger dataset but preserve accuracy.
- ☑ This significantly impacts iteration speed.

Correct
The long training time is a significant constraint on iteration speed.

- ☑ Lowering the number of images will reduce training time and likely allow for an acceptable tradeoff between iteration speed and accuracy.

Correct
There is a balance that allows development at a reasonable rate without significant accuracy loss.