축하합니다! 통과하셨습니다!

받은 학점 80% **최신 제출물 학점** 80% **통과 점수:** 80% 이상

다음 항목으로 이동

1. Problem Statement 1 / 1점

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:

더 보기

맞습니다
Yes. More than one metric expands the choices and tradeoffs you have to decide for each with unknown effects on the other two.

2. After further discussions, the city narrows down its criteria to: 1 / 1점

- "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."

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

(●)

Test Accuracy	Runtime	Memory size
98%	9 sec	9MB

( )

Test Accuracy	Runtime	Memory size
99%	13 sec	9MB

( )

Test Accuracy	Runtime	Memory size
97%	1 sec	3MB

Test Accuracy	Runtime	Memory size
97%	3 sec	2MB

더 보기

맞습니다
Correct! This model has the highest test accuracy, the prominent criteria you are looking for, compared with other models, and also has a runtime <10 seconds and memory size < 10MB.

3. Which of the following best answers why it is important to identify optimizing and satisficing metrics? 1 / 1점

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

더 보기

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

4. With 10,000,000 data points, what is the best option for train/dev/test splits? 1 / 1점

- train - 95%, dev - 2.5%, test - 2.5%
- train - 33.3%, dev - 33.3%, test - 33.3%
- train - 60%, dev - 30%, test - 10%
- train - 60%, dev - 10%, test - 30%

더 보기

맞습니다

Yes. The size of the data set allows for bias and variance evaluation with smaller data sets.

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. You should add the citizens' data to the training set. True/False? **0 / 1점**

- ○ True
- ◉ False

더 보기

틀립니다

No. Adding this data to the training set will change the training set distribution. However, it is not a problem to have different training and dev distributions. In contrast, it would be very problematic to have different dev and test set 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 to the dev set. You object because: (Choose all that apply) **1 / 1점**

- ☐ A bigger test set will slow down the speed of iterating because of the computational expense of evaluating models on the test set.
- ☑ The dev set no longer reflects the distribution of data (security cameras) you most care about.

Correct

Yes. The performance of the model should be evaluated on the same distribution of images it will see in production.

- ☑ 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

Yes. Adding a different distribution to the dev set will skew bias.

- ☐ The 1,000,000 citizens' data images do not have a consistent x-->y mapping as the rest of the data.

더 보기

맞습니다
Great, you got all the right answers.

7. You train a system, and the train/dev set errors are 3.5% and 4.0% respectively. You decide to try regularization to close the train/dev accuracy gap. Do you agree? **0 / 1점**

- (●) No, because this shows your variance is higher than your bias.
- ( ) No, because you do not know what the human performance level is.
- ( ) Yes, because this shows your bias is higher than your variance.
- ( ) Yes, because having a 4.0% training error shows you have a high bias.

더 보기

틀립니다
No. Test accuracy is not given so we can't speak about variance.

8. If your goal is to have "human-level performance" be a proxy (or estimate) for Bayes error, how would you define "human-level performance"? **1 / 1점**

- ( ) The performance of the head of the City Council.
- ( ) The performance of the average citizen of Peacetopia.
- (●) The best performance of a specialist (ornithologist) or possibly a group of specialists.
- ( ) The performance of their volunteer amateur ornithologists.

더 보기

맞습니다
Yes. This is the peak of human performance in this task.

9. Which of the following statements do you agree with? **1 / 1점**

- A learning algorithm's performance can never be better than human-level performance nor better than Bayes error.
- (●) 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.

더 보기

맞습니다

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 / 1점

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.)

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

Correct

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

Correct

- [ ] Try increasing regularization.

더 보기

맞습니다
Great, you got all the right answers.

11. You also evaluate your model on the test set, and find the following: 1 / 1점

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.
- [x] You have overfit to the dev set.

✓ **Correct**

- [ ] You have underfitted to the dev set.
- [x] You should try to get a bigger dev set.

✓ **Correct**

더 보기

맞습니다
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 / 1점

- [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.

- [x] You are close to Bayes error and possible overfitting.

✓ **Correct**
Yes. By definition, Bayes error cannot be exceeded except for overfitting.

- [ ] 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

This is a statistical anomaly (or must be the result of statistical noise) since it should not be possible to surpass human-level performance.

↙↗ 더 보기

맞습니다
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 / 1점

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

↙↗ 더 보기

맞습니다
Yes. 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 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 / 1점

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

더 보기

맞습니다
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 / 1점

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

  Correct
  Yes. The 10x size increase adds a small amount of accuracy but takes too much time.

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

  This should not be selected
  No. Although you may have gained many insights that may reduce the number of iterations needed, a DeepLearning model requires multiple iterations when working in a new dataset.

- [x] Given a significant budget for cloud GPUs, you could mitigate the training time.

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

더 보기

틀립니다
You didn't select all the correct answers