**1.**

Pertanyaan #1

**Problem Statement**

This example is adapted from a real production application, but with details disguised to protect confidentiality.

A picture containing outdoor, cloud, sky, water

Description automatically generated

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

A screenshot of a computer

Description automatically generated**Salah**

No. The goal is to have one metric that focuses the development effort and increases iteration velocity.

**2.**

Pertanyaan #2

The city asks for your help in further defining the criteria for accuracy, runtime, and memory. How would you suggest they identify the criteria?

**A screenshot of a computer

Description automatically generated with medium confidenceBenar**

Yes. The thresholds provide a way to evaluate models head to head.

**3.**

Pertanyaan #3

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

**A screenshot of a computer

Description automatically generated with low confidenceBenar**

**4.**

Pertanyaan #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?

**A screenshot of a computer

Description automatically generated with medium confidenceBenar**

Yes.

**5.**

Pertanyaan #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?

**A picture containing text, screenshot, font

Description automatically generatedBenar**

Yes. This will cause the training and dev/test set distributions to become different, however as long as dev/test distributions are the same you are aiming at the same target.

**6.**

Pertanyaan #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 proportionately to the train/dev/test sets. You object because:

**A screenshot of a computer

Description automatically generated with low confidenceBenar**

Yes. Using the data in the training set could be beneficial, but you wouldn't want to include such images in your test set as they are not from the expected distribution of data you'll see in production.

**7.**

Pertanyaan #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?

**A screenshot of a computer

Description automatically generated with medium confidenceBenar**

**8.**

Pertanyaan #8

You want to define what human-level performance is to the city council. Which of the following is the best answer?

**A screenshot of a computer

Description automatically generated with medium confidenceBenar**

Yes. The best human performance is closest to Bayes' error.

**9.**

Pertanyaan #9

Which of the following statements do you agree with?

**A screenshot of a computer

Description automatically generated with low confidenceBenar**

**10.**

Pertanyaan #10

After working on your algorithm you have to decide the next steps. Currently, human-level performance is 0.1%, training is at 2.0% and the dev set is at 2.1%. Which statement below best describes your thought process?

**A screenshot of a computer screen

Description automatically generated with low confidenceBenar**

Great, you got all the right answers.

**11.**

Pertanyaan #11

After running your model with the test 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)

**A screenshot of a test

Description automatically generated with medium confidenceBenar**

Great, you got all the right answers.

**12.**

Pertanyaan #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.)

**A screenshot of a computer

Description automatically generated with medium confidenceBenar**

Great, you got all the right answers.

**13.**

Pertanyaan #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?

**A screenshot of a computer

Description automatically generated with medium confidenceBenar**

Yes. The target has shifted so an updated metric is required.

**14.**

Pertanyaan #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.

A picture containing bird, pigeon, outdoor, tree

Description automatically generated

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?

**A picture containing text, bird, pigeon, outdoor

Description automatically generatedBenar**

**15.**

Pertanyaan #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.)

**A screenshot of a chat

Description automatically generated with medium confidenceBenar**

Great, you got all the right answers.