Congratulations! You passed!

Grade received 100% | Latest Submission Grade 100% | To pass 80% or higher

Retake the assignment in **24h 48m**

Go to next item

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 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 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

Expand

Correct

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? 1 / 1 point

- ( ) Suggest that they purchase more infrastructure to ensure the model runs quickly and accurately.

- Suggest to them that they define which criterion is most important. Then, set thresholds for the other two.
- Suggest to them that they focus on whichever criterion is important and then eliminate the other two.

Expand

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

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

- Identifying the optimizing metric informs the team which models they should try first.
- 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.
- Knowing the metrics provides input for efficient project planning.

Expand

Correct
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 / 1 point**

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

Expand

Correct
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. Which of the following is the best use of that additional data? **1 / 1 point**

- Do not use the data. It will change the distribution of any set it is added to.
- Split it among train/dev/test equally.
- Add it to the dev set to evaluate how well the model generalizes across a broader set.
- Add it to the training set.

Expand

Correct
Yes. It is not a problem to have different training and dev distributions. Different dev and test distributions would be an issue.

6. One member of the City Council knows a little about machine learning and thinks you should add the 1,000,000 **1 / 1 point**

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: 1 / 1 point

- The additional data would significantly slow down training time.
- The training set will not be as accurate because of the different distributions.
- If we add the images to the test set then it won't reflect the distribution of data expected in production.
- The 1,000,000 citizens' data images do not have a consistent x-->y mapping as the rest of the data.

Expand

Correct
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. Human performance for identifying birds is < 1%, training set error is 5.2% and dev set error is 7.3%. Which of the options below is the best next step? 1 / 1 point

- Get more data or apply regularization to reduce variance.
- Try an ensemble model to reduce bias and 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 drive down the >4.0% training error.

Expand

Correct
Yes. Avoidable bias is >4.2% which is larger than the 2.1% 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 / 1 point

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

Expand

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

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

Expand

Correct

10. Which of the following best expresses how to evaluate the next steps in your project when your results for human-level performance, train, and dev set error are 0.1%, 2.0%, and 2.1% respectively? **1 / 1 point**

- (•) Based on differences between the three levels of performance, prioritize actions to decrease bias and iterate.
- ( ) Evaluate the test set to determine the magnitude of the variance.
- ( ) Port the code to the target devices to evaluate if your model meets or exceeds the satisficing metrics.
- ( ) Keep tuning until the train set accuracy is equal to human-level performance because it is the optimizing metric.

Expand

Correct
Yes. Always choose the area with the biggest opportunity for improvement.

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) **1 / 1 point**

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

  Correct
  Yes. The dev set performance versus the test set indicates it is overfitting.

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

  Correct
  Yes. The dev set performance versus the test set indicates it is overfitting.

- [ ] Try decreasing regularization for better generalization with the dev set.

Expand

Correct
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 likely? (Check all that apply.) **1 / 1 point**

- [ ] There is still avoidable bias.
- [x] The model has recognized emergent features that humans cannot. (Chess and Go for example)

  Correct
  Yes. When Google beat the world Go champion, it was recognized that it was making deeper moves than humans.

- [x] Pushing to even higher accuracy will be slow because you will not be able to easily identify sources of bias.

  Correct
  Yes. Exceeding human performance means you are close to Bayes error.

- [ ] This result is not possible since it should not be possible to surpass human-level performance.

Expand

Correct
Great, you got all the right answers.

13. Your system is now very accurate but has a higher false negative rate than the City Council of Peacetopia would like. What is your best next step? **1 / 1 point**

- (•) Reset your "target" (metric) for the team and tune to it.
- ( ) 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.
- ( ) Expand your model size to account for more corner cases.

Expand

Correct
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 / 1 point**

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

Expand

Correct
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.) **1 / 1 point**

- [x] This significantly impacts iteration speed.

  Correct
  Yes. This training time is an absolute constraint on iteration.

- [ ] Reducing the model complexity will allow the use of the larger data set but preserve accuracy.
- [x] Lowering the number of images will reduce training time and likely allow for an acceptable tradeoff between iteration speed and accuracy.

  Correct
  Yes. There is a sweet spot that allows development at a reasonable rate without significant accuracy loss.

Expand

Correct
Great, you got all the right answers.