

1. For which of the following problems would anomaly detection be a suitable algorithm?

1 / 1 point

- ☐ Given data from credit card transactions, classify each transaction according to type of purchase (for example: food, transportation, clothing).
- ☒ From a large set of primary care patient records, identify individuals who might have unusual health conditions.

✓ Correct

Since you are just looking for unusual conditions instead of a particular disease, this is a good application of anomaly detection.

- ☐ Given an image of a face, determine whether or not it is the face of a particular famous individual.
- ☒ Given a dataset of credit card transactions, identify unusual transactions to flag them as possibly fraudulent.

✓ Correct

By modeling "normal" credit card transactions, you can then use anomaly detection to flag the unusual ones which might be fraudulent.

2. Suppose you have trained an anomaly detection system that flags anomalies when  $p(x)$  is less than  $\epsilon$ , and you find on the cross-validation set that it has too many false negatives (failing to flag a lot of anomalies). What should you do?

1 / 1 point

- ☒ Increase  $\epsilon$
- ☐ Decrease  $\epsilon$

✓ Correct

By increasing  $\epsilon$ , you will flag more anomalies, as desired.

3. Suppose you are developing an anomaly detection system to catch manufacturing defects in airplane engines. Your model uses

1 / 1 point

$$p(x) = \prod_{j=1}^n p(x_j; \mu_j, \sigma_j^2).$$

You have two features  $x_1$  = vibration intensity, and  $x_2$  = heat generated. Both  $x_1$  and  $x_2$  take on values between 0 and 1 (and are strictly greater than 0), and for most "normal" engines you expect that  $x_1 \approx x_2$ . One of the suspected anomalies is that a flawed engine may vibrate very intensely even without generating much heat (large  $x_1$ , small  $x_2$ ), even though the particular values of  $x_1$  and  $x_2$  may not fall outside their typical ranges of values. What additional feature  $x_3$  should you create to capture these types of anomalies:

- ☐  $x_3 = x_1 + x_2$
- ☒  $x_3 = \frac{x_1}{x_2}$
- ☐  $x_3 = x_1^2 \times x_2$
- ☐  $x_3 = x_1 \times x_2$

✓ Correct

This is correct, as it will take on large values for anomalous examples and smaller values for normal examples.

4. Which of the following are true? Check all that apply.

1 / 1 point

- ☒ When choosing features for an anomaly detection system, it is a good idea to look for features that take on unusually large or small values for (mainly the) anomalous examples.

✓ Correct

These are good features, as they will lie outside the learned model, so you will have small values for  $p(x)$  with these examples.

- ☒ If you do not have any labeled data (or if all your data has label  $y = 0$ ), then it is still possible to learn  $p(x)$ , but it may be harder to evaluate the system or choose a good value of  $\epsilon$ .

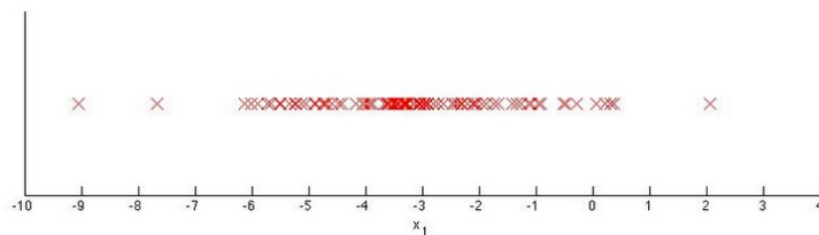
✓ Correct

Only negative examples are used in training, but it is good to have some labeled data of both types for cross-validation.

- ☐ If you have a large labeled training set with many positive examples and many negative examples, the anomaly detection algorithm will likely perform just as well as a supervised learning algorithm such as an SVM.
- ☐ If you are developing an anomaly detection system, there is no way to make use of labeled data to improve your system.

5. You have a 1-D dataset  $\{x^{(1)}, \dots, x^{(m)}\}$  and you want to detect outliers in the dataset. You first plot the dataset and it looks like this:

1 / 1 point



Suppose you fit the gaussian distribution parameters  $\mu_1$  and  $\sigma_1^2$  to this dataset. Which of the following values for  $\mu_1$  and  $\sigma_1^2$  might you get?

- ☒  $\mu_1 = -3, \sigma_1^2 = 4$
- ☐  $\mu_1 = -6, \sigma_1^2 = 4$
- ☐  $\mu_1 = -3, \sigma_1^2 = 2$
- ☐  $\mu_1 = -6, \sigma_1^2 = 2$

✓ Correct

This is correct, as the data are centered around -3 and tail most of the points lie in  $[-5, -1]$ .