✖

Anomaly Detection

5 questions

1 point

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

- [ ] Given an image of a face, determine whether or not it is the face of a particular famous individual.
- [ ] Given data from credit card transactions, classify each transaction according to type of purchase (for example: food, transportation, clothing).
- [ ] Given a dataset of credit card transactions, identify unusual transactions to flag them as possibly fraudulent.
- [ ] From a large set of primary care patient records, identify individuals who might have unusual health conditions.

---

1 point

2.
Suppose you have trained an anomaly detection system for fraud detection, and your system that flags anomalies when $p(x)$ is less than ε, and you find on the cross-validation set that it is missing many fradulent transactions (i.e., failing to flag them as anomalies). What should you do?

- ○ Decrease ε
- ○ Increase ε

1
point

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

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

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:

- $x_3 = \frac{1}{x_1}$
- $x_3 = x_1 + x_2$
- $x_3 = \frac{1}{x_2}$
- $x_3 = \frac{x_1}{x_2}$

1
point

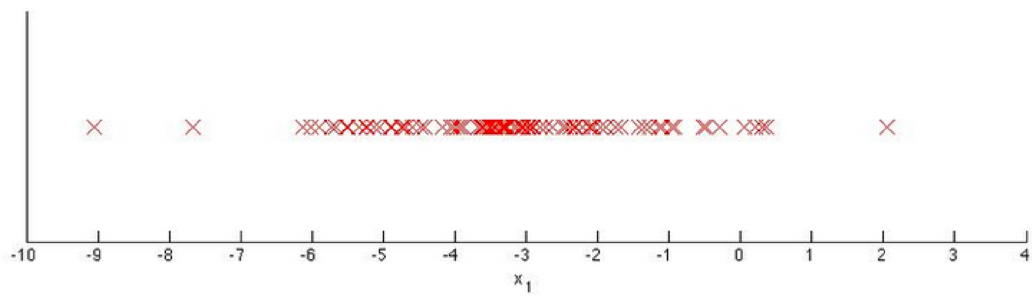
4.
Which of the following are true? Check all that apply.

- [ ] In a typical anomaly detection setting, we have a large number of anomalous examples, and a relatively small number of normal/non-anomalous examples.
- [ ] When developing an anomaly detection system, it is often useful to select an appropriate numerical performance metric to evaluate the effectiveness of the learning algorithm.
- [ ] In anomaly detection, we fit a model $p(x)$ to a set of negative ($y = 0$) examples, without using any positive examples we may have collected of previously observed anomalies.
- [ ]

When evaluating an anomaly detection algorithm on the cross validation set (containing some positive and some negative examples), classification accuracy is usually a good evaluation metric to use.

1
point

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



Suppose you fit the gaussian distribution parameters μ_1 and σ_1^2 to this dataset. Which of the following values for μ_1 and σ_1^2 might you get?

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

I understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.
Learn more about Coursera's Honor Code

Anderson Banihirwe

Submit Quiz