

Springboard - Blog

<https://www.springboard.com/blog/machine-learning-interview-questions/>

1. Bias vs Variance

- **Bias** is due to erroneous or overly simplistic assumptions in learning algorithm
- Usually underfitting your data
- **Variance** typically due to too much complexity in learning algorithm
- Makes the model sensitive to high degrees of variation in training data.
- Too much noise from training data
- If you make data more complex and add more variables, you'll lose bias but gain variance.

2. Supervised vs Unsupervised learning

- Supervised requires labeled data. Unsupervised does not.

3. How is KNN different from k-means clustering?

- KNN is a supervised classification algorithm.
- K-means clustering is unsupervised.
- Works very similarly
- KNN required labelled data
- K means clustering requires only a set of unlabeled point and a threshold
- The algorithm will gradually *learn* how to cluster them by computing mean of the distance between different points.

4. How does a ROC curve work

- graphical representation of contrast between true and false positive rate at various thresholds.
- Used as a proxy for trade-off between sensitivity of model (true positive) vs the fall-out or probability it will trigger a false alarm (false positives)
- Think about recall and precision in this case.
 - *ex.* You'd have perfect recall (there are actually 10 apples, and you predicted there would be 10) but 66.7% precision because out of the 15 events you predicted, only 10 (the apples) are correct.

5. Baye's Theorem?

-

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- Leads to a branch of ML called Naive Bayes classifier

6. Why is "Naive" Bayes naive?

- Used a lot in text mining
- It's naive because it makes an assumption that is virtually impossible in real-life data.
 - conditional probability is calculated as the pure product of the individual probabilities of components.
 - This implies absolute independence of features - condition probably never met in real life.
- Another way put, if a Naive Bayes classifier figured that you liked pickles and ice-cream would probably naively recommend you a pickle ice-cream.

7. Difference between L1 and L2 regularization

- Regularization helps solve over-fitting problems in ML
 - Simple model will be very poor generalization of data.
 - Complex model may not perform well in test due to over-fitting.
 - Regularization refers to adding a penalty term to objective function and control model complexity using that penalty term.
 - Ridge regression used L_2 norm for regularization.
-
-