

You'll often encounter the Machine Learning Algorithms interview during an onsite round for a Machine Learning Engineer (MLE), Data Scientist (DS) or Software Engineer-Machine Learning (SE-ML) position.

## What to expect in the interview

The interviewer will try to uncover how deeply you understand machine learning algorithms (usually classic ones).

Here's a list of questions you might be asked:

- Derive the binary cross-entropy loss function.
- How does Logistic Regression differ from Linear Regression?
- What is the difference between Gradient Descent and Stochastic Gradient Descent?
- Explain a classic machine learning algorithm, among the following list: Linear Regression, Logistic Regression, Decision Trees, Random Forest, XGBoost, Support Vector Machines, K-means, K-Nearest Neighbors, Neural Networks, Principal Component Analysis, Naive Bayes Classifier, L1/L2 regularization, etc.
- Why is the EM algorithm useful?
- Why is the Naive Bayes classifier called Naive?
- How does a discriminative model differ from a generative model?
- In K-Nearest Neighbors, how does the value of K impacts the bias and variance?
- In Support Vector Machines, what is the kernel trick?

## Resources

Useful content to prepare for this interview:

- CS229 lecture notes:
  - [Structuring your Machine Learning Project](#) (Coursera)
  - [Linear Regression and Logistic Regression](#)
  - [Generative Learning Algorithms, Naive Bayes](#)
  - [Support Vector Machines](#)
  - [Neural Networks and Deep Learning](#)
  - [Bias/variance tradeoff and Error analysis](#)
  - [Regularization and Model Selection](#)
  - [Unsupervised Learning, K-means clustering](#)
  - [Mixtures of Gaussians and the EM algorithm](#)
  - [Principal components analysis](#)
  - [Independent components analysis](#)

- Coursera: [Machine Learning](#)

## Recommended Plan

Here are useful rules of thumb to follow:

- Listen to the hints given by your interviewer. **Example:** You're explaining PCA and state that "we should find the eigenvalues and eigenvectors of the data matrix X". If your interviewer questions you with "are you sure?" or "can you interpret the eigenvalues of X?", there is a high chance your answer is imprecise or wrong. You should react by reconsidering and talking through your answer. In this case, the interviewer expects you to introduce the covariance matrix of X and find its eigenvalues/eigenvectors.
- Don't mention methods you're not able to explain. **Example:** You're explaining Logistic Regression and state that "we're using logistic regression for binary classification problems. For a multi-class problem we would use a softmax regression." In this scenario, you can expect the interviewer to ask: "could you explain softmax regression?"
- Write clearly, draw charts, and introduce a notation if necessary. The interviewer will judge your scientific rigor. **Example:** You're asked to write the binary cross entropy loss function. Instead of writing:  $L = - \sum [y \log y + (1 - y) \log (1 - y)]$  , write:  $L(y, y) = - \frac{1}{m} \sum_{i=1}^m y_i \log(y_i) + (1 - y_i) \log(1 - y_i)$
- Many teams will ask you questions about the methods the company is using. Before going onsite, read online about the product the company is building and try to infer the methods they might be using. **Example:** If you're interviewing with a fraud detection team, you might want to learn about the methods to deal with imbalanced datasets, precision, recall, and F1 score before going onsite.
- When you are not sure of your answer, be honest and say so. Good interviewers care about your ability rather than your knowledge.
- When you get stuck, think out loud rather than staying silent. Talking through your thought process will help the interviewer correct you and point you in the right direction.