



# How To Prepare For A Machine Learning Interview



May 25, 2016

(https://twitter.com/https://google.tokachi/search.php?article?url=https://blogs.hongkongfp.com/2016/05/prepare-machine-learning-interview-questions/&text=How%20to%20Prepare%20for%20Machine%20Learning%20Interview)	
To Prepare For A Machine Learning Interview)	To Prepare For A Machine Learning Interview)



Getting ready for a job interview has been likened to everything from preparing for battle, to gearing up to ask someone out on a date, to lining up a putt on the 18th green at The Masters. Meaning, at best, it's nerve-racking, and at worse, it's terrifying! Preparing for a Machine Learning interview is no different. You know you've got something ahead with the potential to be either really great, or really terrible. But how do you ensure your result is the great one?

It's all about mindset, and preparation.

## Company and Role

Understanding the context of your pending interview—i.e. the reason WHY there's an open role in the first place—should be an integral part of your preparation. Knowing why you're being interviewed will help you contextualize *your* value to the company. For example, if a company is looking to hire a Machine Learning Engineer, it should be clear that they are

trying to solve a complex problem where traditional algorithmic solutions are hard to apply or simply do not work well enough. It should also be clear they are also extremely motivated to solve that problem.

## Identify a Core Problem

The first thing you need to do when *applying* for such a role is to imagine yourself *in* that roll. To do this, you need to find out as much as possible about the company and position. To organize your research, ask yourself: *What is one core problem I can solve for this company?* Pursuing an answer to this question should excite you, and drive you to find out more about the problem—existing approaches, recent developments in that domain—and lead you to a bunch of more specific challenges. If you know what team you are being interviewed for, picking an appropriate problem might be easy; otherwise, choose something that is essential for the company. Put another way, think about the challenges facing the company, and then try to determine the questions they're likely asking.

## Explore Potential Data Sources

The next step in your preparations should be to think about what data you need to answer those questions. Some of this may be readily available, while you may have to build in additional hooks to gather certain pieces of information. Dig into the company's infrastructure and operations—what stack do they operate on, what APIs do they have, what data are they already collecting, etc. Most companies today have a blog where they often discuss their challenges, approaches, successes and failures. This should give you further insight into how they operate, and what products and services they might have in the pipeline.

## Prepare to Discuss Machine Learning Solutions

Now you need to make a fairly big conceptual jump: *How does machine learning fit into all this?* Given what you're trying to achieve, and the data you think might be available, can you cast it into a learning problem? What is an appropriate model to use? How would you go about training and evaluating it? To give you an example, the primary challenge that a lot of

recommendation systems like Netflix and Amazon face is clustering, not prediction—i.e. once you are able to figure out groups of users who seem to have similar preferences and behaviors, it becomes a whole lot easier to recommend products that they may find useful.

This thought process will help you be prepared to talk about issues that matter to the company the most. Nobody expects you to walk into an interview and lay out a complete solution for something they've been working hard on for months or years! But everybody likes a candidate who shows genuine interest, motivation, and curiosity for a problem that is close to their hearts.

Depending on your interviewer and the stage of your interview, you may be asked more technical questions, but you should try to use any opportunity you get to demonstrate that you have thought about the company and role. When asked more open-ended questions such as *"Describe a technical challenge you faced when working on a project and how you solved it?"* try to pick something that aligns well with the company's interests.

## Skills and Sample Questions

I recently wrote a piece for the Udacity blog entitled [5 Skills You Need to Become a Machine Learning Engineer](http://blog.udacity.com/2016/04/5-skills-you-need-to-become-a-machine-learning-engineer.html) (<http://blog.udacity.com/2016/04/5-skills-you-need-to-become-a-machine-learning-engineer.html>). In that article I identified five groupings for the essential skills that a Machine Learning Engineer needs:

1. Computer Science Fundamentals and Programming
2. Probability and Statistics
3. Data Modeling and Evaluation
4. Applying Machine Learning Algorithms and Libraries
5. Software Engineering and System Design

I encourage you to read that post for further detail about these groups. What I wish to focus on here are the kinds of questions you're likely to face in a Machine Learning interview, so I'll use these groupings simply as an organizing principle.

## Computer Science Fundamentals and Programming

### Sample Questions

- How would you check if a linked list has cycles?
- Given two elements in a binary search tree, find their lowest common ancestor.
- Write a function to sort a given stack.
- What is the time complexity of any comparison-based sorting algorithm? Can you prove it?
- How will you find the shortest path from one node to another in a weighted graph? What if some weights are negative?
- Find all palindromic substrings in a given string.

For all such questions, you should be able to reason about the time and space complexity of your approach (usually in big-O notation), and try to aim for the lowest complexity possible.

Extensive practice is the only way to familiarize yourself with the different classes of problems, so that you can quickly converge on an efficient solution. Coding/interview prep platforms like [InterviewBit](https://www.interviewbit.com/) (<https://www.interviewbit.com/>), [LeetCode](https://leetcode.com/) (<https://leetcode.com/>), [Interview Cake](https://www.interviewcake.com/) (<https://www.interviewcake.com/>), [Pramp](https://www.pramp.com/) (<https://www.pramp.com/>) and [interviewing.io](http://interviewing.io/) (<http://interviewing.io/>) are highly beneficial for this purpose.

## Probability and Statistics

### Sample Questions

- The mean heights of men and women in a population were calculated to be  $\mu_M$  and  $\mu_W$ . What is the mean height of the total population?

- A recent poll revealed that a third of the cars in Italy are Ferraris, and that half of those cars are red. If you spot a red car approaching from a distance, what is the likelihood that it is a Ferrari?
- You're trying to find the best place to put in an advertisement banner on your website. You can make the size (thickness) small, medium or large, and choose vertical position top, middle or bottom. At least how many total page visits ( $n$ ) and ad clicks ( $m$ ) do you need to say with 95% confidence that one of the designs performs better than all the other possibilities?

Remember that many machine learning algorithms have a basis in probability and statistics. Conceptual clarity of these fundamentals is extremely important, but at the same time, you must be able to relate abstract formulae with real-world quantities.

## Data Modeling and Evaluation

### Sample Questions

- A dairy farmer is trying to understand the factors that affect milk production of her cattle. She has been keeping logs of the daily temperature (usually 30-40°C), humidity (60-90%), feed consumption (2000-2500 kgs), and milk produced (500-1000 liters).
  - *How would you begin processing the data in order to model it, with the goal of predicting liters of milk produced in a day?*
  - *What kind of machine learning problem is this?*
- Your company is building a facial expression coding system, which needs to take input images from a standard HD 1920×1080 pixel webcam, and continuously tell whether the user is in one of the following states: neutral, happy, sad, angry or afraid. When the user's face is not visible in the camera frame, it should indicate a special state: none.
  - *What class of machine learning problems does this belong to?*
  - *If each pixel is made up of 3 values (for red, green, blue channels), what is the raw input data complexity (no. of dimensions) for processing each image? Is there a way to reduce the no. of dimensions?*
  - *How would you encode the output of the system? Explain why.*
- Climate data collected over the past century reveals a cyclic pattern of rising and falling temperatures. How would you model this data (a sequence of average annual

temperature values) to predict the average temperature over the next 5 years?

- Your job at an online news service is to collect text reports from around the world, and present each story as a single article with content aggregated from different sources. How would you go about designing such a system? What ML techniques would you apply?

## Applying Machine Learning Algorithms and Libraries

### Sample Questions

- I'm trying to fit a single hidden layer neural network to a given dataset, and I find that the weights are oscillating a lot over training iterations (varying wildly, often swinging between positive and negative values). What parameter do I need to tune to address this issue?
- When training a support vector machine, what value are you optimizing for?
- Lasso regression uses the L1-norm of coefficients as a penalty term, while ridge regression uses the L2-norm. Which of these regularization methods is more likely to result in *sparse* solutions, where one or more coefficients are exactly zero?
- When training a 10-layer neural net using backpropagation, I find that the weights for the top 3 layers are not changing at all! The next few layers (4-6) are changing, but very slowly. What's going on and how do I fix this?
- I've found some data about wheat-growing regions in Europe that includes annual rainfall (R, in inches), mean altitude (A, in meters) and wheat output (O, in kgs/km<sup>2</sup>). A rough analysis and some plots make me believe that output is related to the square of rainfall, and log of altitude:  $O = \beta_0 + \beta_1 \times R^2 + \beta_2 \times \log_e(A)$   
Can I fit the coefficients ( $\beta$ ) in my model to the data using linear regression?

Data science and Machine Learning challenges such as those on [Kaggle](https://www.kaggle.com/)

(<http://www.kaggle.com/>) are a great way to get exposed to different kinds of problems and their nuances. Try to participate in as many as you can, and apply different machine learning models.

## Software Engineering and System Design

### Sample Questions

- You run an ecommerce website. When a user clicks on an item to open its details page, you would like to suggest 5 more items that the user may be interested in, based on item features as well as the user's purchase history, and display them at the bottom of the page. What services and database tables would you need to support this behavior? Assuming they're available, write a query or procedure to fetch the 5 items to suggest.
- What data would you like to collect from an online video player (like YouTube) to measure user engagement and video popularity?
- A very simple spam detection system works as follows: It processes one email at a time and counts the number of occurrences of each unique word in it (term frequency), and then it compares those counts with those of previously seen emails which have been marked as spam or not. In order to scale up this system to handle a large volume of email traffic, can you design a map-reduce scheme that can run on a cluster of computers?
- You want to generate a live visualization of what portion of a webpage users are currently viewing and clicking, sort of like a heat map. What components/services/APIs do you need in place, on the client and server end, to enable this?

## Conclusion

What I've tried to present here are two sides of the Machine Learning interview experience—call them the Contextual side and the Technical side. If there's one message I'd like to stress, it's that you should resist the temptation to focus on the latter at the expense of the former. It is far too common for aspiring Machine Learners to immerse themselves in technical preparations, while giving very little thought to the *why* of their interview—*why* is there an open role, *why* is the company pursuing Machine Learning talent (and Machine Learning solutions!), *why* are they interested in you? Understanding these questions will give meaning and context to the technical challenges you'll need to address, and answering them will set you apart as the candidate best suited to move the company forward.

~



For more career resources (including the original resource guide this post is based on), please visit the [Udacity Career Resource Center \(http://career.udacity.com/resource-center/index.html\)](http://career.udacity.com/resource-center/index.html).

~

## Machine Learning Nanodegree Program

**ENROLL NOW**

[https://www.udacity.com/degrees/machine-learning-engineer-nanodegree--nd009?utm\\_source=blog&utm\\_medium=referral&utm\\_campaign=sb\\_widget\\_enroll](https://www.udacity.com/degrees/machine-learning-engineer-nanodegree--nd009?utm_source=blog&utm_medium=referral&utm_campaign=sb_widget_enroll)

## Get Job Ready with Udacity

Master in-demand skills. Build and design amazing projects. Earn a valued credential. Launch your career in Data Science, Machine Learning, Android, iOS, and more.

**START LEARNING TODAY**

<https://www.udacity.com/pathfinder>

## Popular Nanodegree Programs

Intro to Programming (<https://www.udacity.com/course/intro-to-programming-nanodegree--nd000>)

Android Basics Nanodegree by Google (<https://www.udacity.com/course/android-basics-nanodegree-by-google--nd803>)

Front-End Web Developer Nanodegree (<https://www.udacity.com/course/front-end-web-developer-nanodegree--nd001>)

iOS Developer (<https://www.udacity.com/course/ios-developer-nanodegree--nd003>)

Machine Learning Engineer (<https://www.udacity.com/course/machine-learning-engineer-nanodegree--nd009>)

Android Developer (<https://www.udacity.com/course/android-developer-nanodegree--nd801>)

## Latest Posts

How to Barter Your Way to Career Success (<https://blog.udacity.com/2018/08/how-to-barter-your-way-to-career-success.html>)

The Google Developer Scholarship and the Power of Community (<https://blog.udacity.com/2018/08/the-google-developer-scholarship-and-the-power-of-community.html>)

Your Employers Are Talking about Workforce Transformation, but What Does That Mean for You? (<https://blog.udacity.com/2018/08/workforce-transformation.html>)

Introducing Nanodegree Program Learning Plans (<https://blog.udacity.com/2018/08/introducing-nanodegree-program-learning-plans.html>)

This Udacity Graduate Has A New Artificial Intelligence Career! (<https://blog.udacity.com/2018/08/udacity-graduate-new-artificial-intelligence-career.html>)

## SUBSCRIBE TO THE BLOG

### FREE DATA CAREER GUIDE

Interested in pursuing a career in data? Enter your email address for free access to *The Complete Guide to Landing a Career in Data*.

Enter your email for access:

Email address...

### GET THE FREE GUIDE

By providing your email, you consent to receiving marketing communications from us.



## Arpan Chakraborty

Arpan likes to find computing solutions to everyday problems. He is interested in human-computer interaction, robotics and cognitive science. He obtained his PhD from North Carolina State University, focusing on biologically-inspired computer vision. At Udacity, he develops content for artificial intelligence and machine learning courses.

### POSTS BY ARPAN

Designing Our Artificial Intelligence Curriculum (<https://blog.udacity.com/2016/11/artificial-intelligence-curriculum.html>)

How Companies Are Using Kaggle To Find The Best Machine Learning Talent (<https://blog.udacity.com/2016/07/companies-kaggle-machine-learning-talent.html>)

## Share This Post

TWITTER

([HTTPS://TWITTER.COM/INTENT/TWEET?URL=HTTPS://BLOG.UDACITY.COM/2016/05/PREPARE-MACHINE-LEARNING-INTERVIEW.HTML&VIA=UDACITY&TEXT=HOW TO PREPARE FOR A MACHINE](https://twitter.com/intent/tweet?url=https://blog.udacity.com/2016/05/prepare-machine-learning-interview.html&via=udacity&text=How%20to%20prepare%20for%20a%20machine%20learning%20interview)

**LEARNING INTERVIEW)**

FACEBOOK

([HTTP://WWW.FACEBOOK.COM/SHARER/SHARER.PHP?](http://www.facebook.com/sharer/sharer.php?u=https://blog.udacity.com/2016/05/prepare-machine-learning-interview.html)  
[U=HTTPS://BLOG.UDACITY.COM/2016/05/PREPARE-MACHINE-LEARNING-INTERVIEW.HTML](https://blog.udacity.com/2016/05/prepare-machine-learning-interview.html))

GOOGLE+

([HTTPS://PLUS.GOOGLE.COM/SHARE?URL=HTTPS://BLOG.UDACITY.COM/2016/05/PREPARE-MACHINE-LEARNING-INTERVIEW.HTML](https://plus.google.com/share?url=https://blog.udacity.com/2016/05/prepare-machine-learning-interview.html))

LINKEDIN

([HTTPS://WWW.LINKEDIN.COM/SHAREARTICLE?](https://www.linkedin.com/sharearticle?mini=true&url=https://blog.udacity.com/2016/05/prepare-machine-learning-interview.html&title=How%20to%20prepare%20for%20a%20machine%20learning%20interview)  
[MINI=TRUE&URL=HTTPS://BLOG.UDACITY.COM/2016/05/PREPARE-MACHINE-LEARNING-INTERVIEW.HTML&TITLE=HOW TO PREPARE FOR A MACHINE LEARNING INTERVIEW](https://blog.udacity.com/2016/05/prepare-machine-learning-interview.html))

PREVIOUS POST

← **Announcing: Objective-C for Swift Developers**  
(<https://blog.udacity.com/2016/05/announcing-objective-c-swift-developers.html>)

NEXT POST

**7 Courses Built With Google – All For Free!**  
(<https://blog.udacity.com/2016/05/7-courses-built-with-google-all-for-free.html>) →

[POPULAR NANODEGREE PROGRAMS](#)[STUDENT RESOURCES](#)[UDACITY](#)[INQUIRIES](#)

Nanodegree is a trademark of Udacity.

© 2011-2018 Udacity, Inc.



(<https://twitter.com/udacity>) (<https://www.facebook.com/udacity>) (<https://plus.google.com/u/0/+Udacity/posts>) (<https://www.linkedin.com/company/udacity>)

## Get The Free Google Ads Glossary!



### Create Successful Google Ads Campaigns

**Download the Guide**

By providing your email, you consent to receiving marketing communications from us.