ML Dream Job Checklist

Don't just start applying: You need to make sure you have the technical skills down cold. If not, you're wasting time on an application when you should have been studying.

Decide if you want to be a **data scientist**, or a **machine learning (ML) engineer**. If you're new to the field, I'd strongly recommend applying to data science. Here's why:

- The technical questions are easier to understand. It's easier to grasp the big picture of why you're solving a data science problem. The reason for solving an ML engineer problem may not be obvious at first glance.
- You can likely relate more of your past experiences to a data science role. If you've done any type of statistics at all, you are equipped to talk about the data science applications to that project.
- In order to become an ML engineer, companies likely want to see that you completed data science projects in the past anyways. In many companies, being a data scientist is a prerequisite for being an ML engineer.

When preparing for the technical interview, you should break up your study into two types of sessions.

- **80% of your time should be spent doing Individual Study Sessions (ISS)**: This is where you get a first look and first try at new technical concepts. Let's say you had a technical interview that required you to know how logistic regression works, but you didn't know how to explain it properly. You should use your individual study time to learn the topic inside and out. You do two types of study during individual study.
  - Problem sets (70%): example data science problems that cover a topic you wish to learn.
  - Information gathering (30%): looking at textbooks, YouTube videos, and blogs for an introduction to the information
- **20% of your time should be spent doing Group Study Sessions (GSS)**: This is when you practice the topics you learned from individual study amongst your peers, and with the guidance of a more experienced individual.
  - Oral recitation (60%): You talk about an algorithm/concept with a peer. This is important because technical interviews often have an in person component as well as a coding component.
  - Group problem set (40%): A data science problem is presented, and every single person talks to each other about a proposed solution

You generally know you're ready to start applying when you pass the two tests below. And if you fail them, continue with the ISS/GSS split.

- On a random sample of **20 model coding problems**, you can successfully code at least 17 of them while only getting online help for when you are looking up syntax from the package documentation.
- On a random sample of **50 non-coding ML questions**, you are at least getting 35 of them correct without any help at all.

It's all about the reps. Landing the job will be hard. But if you stick to the method, you got this...

Stick to the method! I applied to >140 ML jobs and I wish I figured this out earlier...