Data Science Projects That Will Get You The Job

How I got an internship four months into self-studying data science

Natassha Selvaraj

Jul 28, 2020·6 min read

Photo by [Free To Use Sounds](#) on [Unsplash](#)

Just a few weeks ago, I had an interview for a data science internship with one of the largest telecommunications company in Asia.

I start next week.

How did I get the interview?

Hint: It wasn't due to my education background (I haven't even completed my CS degree yet). Neither was it through the data science certificates I once so diligently tried to collect.

I got called in for an interview because of my data science portfolio, and the projects I showcased there.

Applying for data science jobs can be pretty overwhelming, simply because it is nearly impossible to be a perfect fit for the company you are applying to.

My advice — Rather than trying to learn every tool there is out there, use what you know to build something useful. That is where you learn the most. Then, showcase your project to the rest of the world. Tell a story around it.

Creating projects you are passionate about does more than display your skill. It shows your love for what you do.

You will have to spend weeks (sometimes even months), trying to bring an idea to life. It is going to be hours of staring at your laptop, trying to fix your code that throws an error at you every time you try to run it.

Sometimes, you will realize that it **isn't possible to finish your project** after spending a month on it.

There is no way to improve your model's accuracy on actual user data because your data set is flawed.

That is a month's worth of hard work down the drain, and you will have to start again.

Only a person who is truly passionate about what they're doing can stick it through.

This will give an employer the confidence that you can get the job done. Even if you don't completely fit the job description, you are someone who is capable of solving the task at hand.

In this article, I will take you through some data science projects you should do. These projects will not only improve your skills as a data scientist, but will also look really good on your resume.

I am going to assume you have taken a course or two in data science, and have a grasp on some programming and data analysis skills. *(If you don't, you can click [here](#) to find a list of free courses that will help you gain the required expertise).*

Kaggle

[Kaggle](#) projects are great to start off with because clean and structured data is handed to you.

It is pretty easy to analyze data that is well structured, and run machine learning models on them.

If you are a beginner, Kaggle is the best place for you to get your hands dirty!

Photo by [jesse orrico](#) on [Unsplash](#)

The best thing about doing Kaggle projects is the **amount of guidance available online.**

Everyone's doing them, and competing to get the highest accuracy. People make blog posts, GitHub repositories, and YouTube videos showcasing their approach towards solving the problem.

Kaggle also provides tutorials on some of their beginner-friendly projects, and these will be very helpful when starting out.

[Here](#) is a list of some Kaggle projects that you should do if you are new to the field.

Kaggle projects are great to start out with when you are a beginner. Less time is wasted on data collection and cleaning, and more time can be spent developing your machine learning skills.

However, if your goal is to create an impressive data science portfolio, stay away from Kaggle.

Why?

Real world data is messy.

When working for a company, you won't be handed a clean and structured data set like in Kaggle.

In fact, around 80% of a data scientist's job is spent cleaning the data, and only around 20% is spent on the model.

By showcasing your Kaggle projects, you have displayed only a small portion of the skill required to do the job — creating highly accurate models (at times).

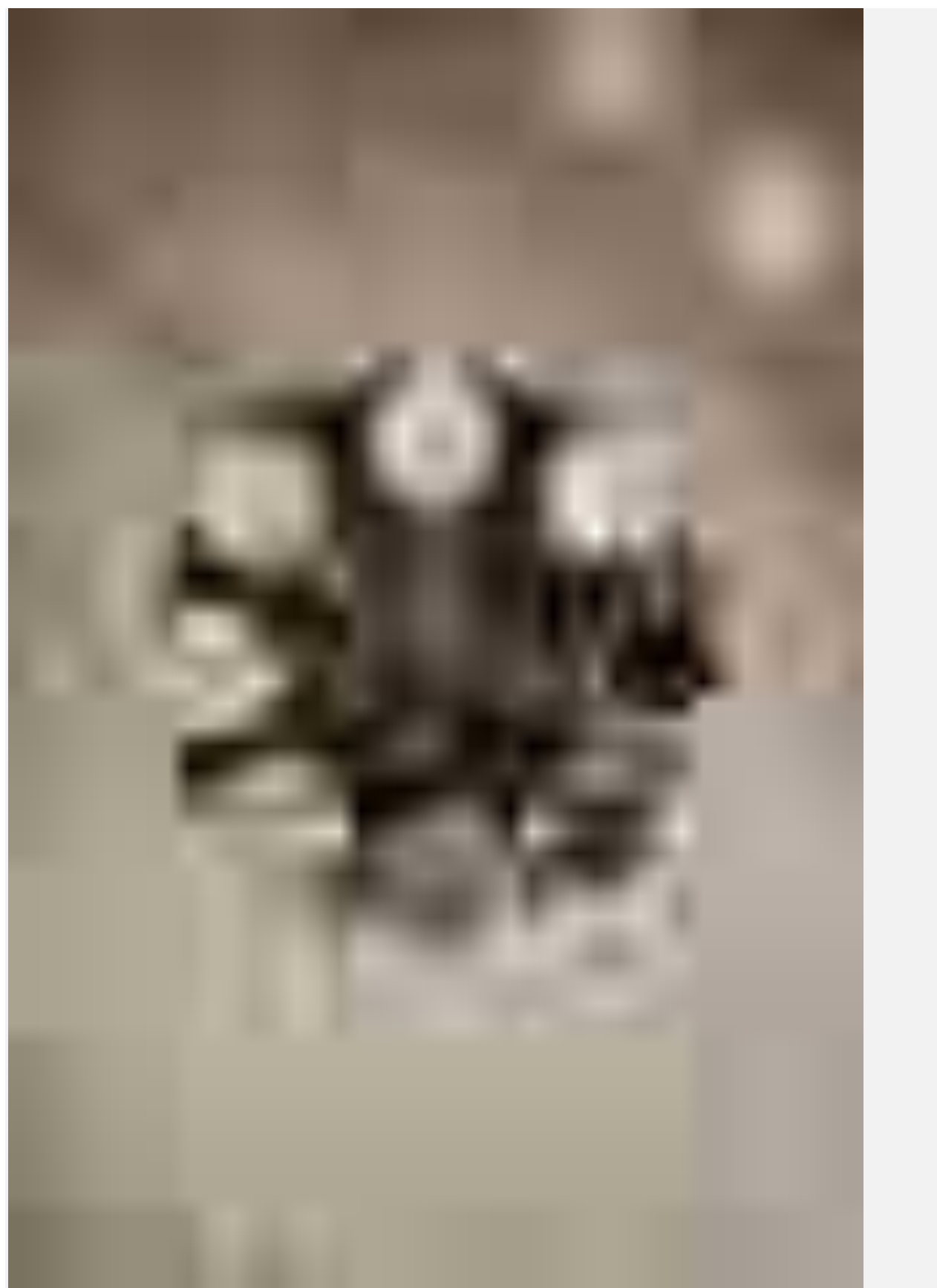So what should you do instead?

Showcase a variety of skills

Yes, now a potential employer knows that you can create great supervised learning models that make highly accurate predictions.

And... What else?

Make sure to create different types of projects that showcase a large variety of skills — data cleaning, analysis, visualization, machine learning, and communication.

If you have domain specific knowledge in another field, then you have a much higher chance of standing out if you integrate that knowledge into your data science projects.

Showcase your passion

Photo by [Ian Schneider](#) on [Unsplash](#)

You need to love what you do. And you may not necessarily love coding, or mathematics, or machine learning, but that's okay.

What you need to do is this — find a way to use data in a field you are passionate about.

Maybe you love music (*I mean, who doesn't?*)

You can try creating a music analysis project, by analyzing trends and predicting the next big hit in the industry!

In fact, there is a whole [page](#) dedicated to music analytics on Towards Data Science.

If you create data science projects related to something you are passionate about, you are far more likely to keep going even when it gets difficult. It will be a lot easier to complete than a project you aren't interested in.

Examples

I have always been interested in bringing to light social issues, such as climate change, gender inequality, and racial disparity.

What better way to gain insight into these issues than data?

This led to me creating data analysis projects, such as:

- [An analysis of gender disparity in Hollywood](#)

- [Driving While Black — Data Analysis](#)

- [An analysis of gender inequality in tech](#)

Also, I love reading. The Harry Potter series was my favourite when I was younger, so I created a [Harry Potter personality prediction model](#).

These are just some examples on how you can combine data and your passion to create a data science portfolio.

Communicate

Most importantly, make sure to communicate your projects.

Creating an amazing product isn't enough if the code is going to sit in your GitHub repository with little to no explanation.

After creating a data science project, tell a story around it.

Why did you decide to start this project? What were the steps taken to complete it? What were your data findings?

Communicate your project in simple terms, so it can be easily understood by anyone who tries to read it.

You can create a [GitHub pages](#) site to explain your project, write to [Medium](#) publications, or even start your own blog.

Data Science Portfolio

Finally, you will need a place to put all your work together. I suggest creating a data science portfolio website to showcase all of it.

When I applied for the internship, all I sent was a [link to my portfolio website](#). I included links to my GitHub and Medium accounts, and wrote a brief summary about all the projects I had done.

You can click [here](#) to find out how to create a simple data science portfolio website. It only took me a day to make mine.

Just Start!

If you have taken a couple of online courses and aren't sure what to do next, I suggest you should start to work on a data science project.

Do some research, pick a data set, and start. Give yourself a deadline, and try to complete it by then.

Good luck!