

Mistakes That Newbie Machine Learning and Data Science Enthusiasts Make And How To Avoid Them

Mistakes made by beginners

1. Spending a lot of time on theory
2. Jumping directly into coding ML algorithms without learning prerequisites
3. Thinking to build the future without learning the basics.
4. Not Spending Enough Time on Exploring and Visualizing the Data
5. Giving importance to tools and libraries over the Business Problem
6. Focusing on Accuracy over Understanding how model works
7. Assuming The Algorithm Is More Important Than Domain Knowledge
8. Not Having a Structured Approach to Problem Solving
9. Learning Multiple Tools at Once
10. Not Learning/Working Consistently
11. Less Communication

Avoiding Mistakes

Spending a lot of time on theory

- It is very important to have a proper balance between theoretical and practical.
- Whatever you learned try practicing them (Just go to Google and search a dataset and practice on it)
- Try participating in hackathons/competitions.
- Learn to be comfortable with partial knowledge. You'll naturally fill in the gaps as you progress.

Jumping directly into coding ML algorithms without learning prerequisites

- Five important components you need to know Linear Algebra, Calculus, Statistics, Probability and Optimization.
- Whatever you build, it is a collection of all individual parts. Data scientist is also same.
- If you do code an algorithm from scratch, do so with the intention of learning instead of perfecting your implementation.

Avoiding Mistakes

Focusing on Accuracy over Understanding how model works

- The best way to prevent yourself from making this mistake is speaking to people working in the industry and understand how their project works.
- There is no better teacher than experience.
- Try to build simpler models and explain it to someone, then slowly add complexity to it and keep doing this until even you don't understand what's going on beneath your data model.

Assuming The Algorithm Is More Important Than Domain Knowledge

- Pushing data randomly without exploring it much will show biases in the results of your model.
- Hence, it is needed much to do some exploratory data analysis which will help you to make some hypothesis of the model and you can be informed about what you are doing.

Avoiding Mistakes

Thinking to build the future without learning the basics

- Learn how classical machine learning algorithms works and master them.
- Basics are the building blocks for advanced topics.
- Learn a systematic approach to structuring machine learning projects.

Not Spending Enough Time on Exploring and Visualizing the Data

- Practice!
- You will be stunned at the amount of insight it will generate for you.
- Understanding your dataset is the foremost task that an aspiring data scientist should do as it will later reflect in your model.
- Ask questions!
- search for solutions on the internet and if you don't find any, ask on social media.

Avoiding Mistakes

Learning Multiple Tools at Once

- Stick to one and master it.
- You will learn more with this approach.
- Each tool has a great community where you can ask your questions whenever you get stuck.

Giving importance to tools and libraries over the Business Problem

- Learn how industries in that domain which you are interested are using data science.
- If possible, search for datasets in a specific industry and try to work on them. This will be a MASSIVE standout point in your resume

Avoiding Mistakes

Not Having a Structured Approach to Problem Solving

- You can acquire a structured thinking mindset through simple training and a disciplined approach.
- Start small, aim big
- Tackle the techniques from the top down
- Avoid getting sloppy with your logic.

Not Learning/Working Consistently

- Set goals and make a time table and stick to it.
- Plan how and what you want to study and set deadlines for yourself.
- If you want to become something you should be ready to put in the hours.
- If you continually keep finding excuses not to study, this might not be the field for you.
- If you practice continuously it will give you a clarity in decision making, gives you control over future and a sense of personal satisfaction.

Avoiding Mistakes

Less Communication

- Practice explaining technical concepts to non-technical audiences.
- To be a keen data scientist, you have to be a good communicator. It really helps a lot.
- One should always keep in mind that data scientists are meant to solve other people's issues and without communicating whether it be inside the organisation or some outside business clients,
- Start participating in discussions and competitions, It is also a way to improve your communication skills.