Chapter 13

Coda

> "Begin at the beginning," the King said, gravely, "and go on till you come to the end: then stop."
>
> – Lewis Carroll

Hopefully you the reader have been at least partially enlightened by this book, and remain excited by the power of data. The most common path to employing these skills is to take a job in industry. This is a noble calling, but be aware there are also other possibilities.

13.1 Get a Job!

There are very rosy predictions of the job prospects for future data scientists. The McKinsey Global Institute projects that demand for "deep analytical talent in the United States could be 50% to 60% greater than its projected supply by 2018." The job placement site `www.glassdoor.com` informs me that as of today, the average data scientist salary is precisely $113,436. *Harvard Business Review* declared that being a data scientist is "the sexiest job of the 21st century" [DP12]. That sounds like the place where I want to be!

But all this testimony would be much more convincing if there was some widely shared understanding of what exactly a data scientist is. It is less obvious to me that there are ever destined to be vast numbers of jobs with the official title of *data scientist* the way there are for, say *software engineer* or *computer programmer*. But don't panic.

It is fair to say that there are several different types of jobs that relate to data science, distinguished by the relative importance of applications knowledge and technical strength. I see the following basic career tracks related to data science:

- *Software engineering for data science*: A substantial fraction of high-end software development positions are at big data companies like Google,

S.S. Skiena, *The Data Science Design Manual,*
Texts in Computer Science, https://doi.org/10.1007/978-3-319-55444-0_13

Facebook, and Amazon, or data-centric companies in the financial sector, like banks and hedge funds. These jobs revolve around building large-scale software infrastructures for managing data, and generally require a degree in computer science to acquire the necessary technical skills and experience.

- *Statistician/data scientists*: There has always been a diverse job market for trained statisticians, especially in health care, manufacturing, business, education, and the government/non-profit sectors. This world will continue to grow and thrive, although I suspect it will demand stronger computational skills than in the past. These computational-oriented statistical analysts will have training or experience in data science, building on a strong foundation in statistics.

- *Quantitative business analysts*: A large cohort of business professionals work in marketing, sales, advertising, and management, providing essential functions at any product-based or consulting company. These careers require a greater degree of business domain knowledge than the previous two categories, but increasingly expect quantitative skills. They may be hiring you to work in marketing, but demand a background or experience in data science/analytics. Or they hire you to work in human resources, but expect you to be able to develop metrics for job performance and satisfaction.

The material covered in this book is essential for all three of these career tracks, but obviously you have more to learn. The careers which are easiest to train for also prove to be the quickest to saturate, so keep developing your skills through coursework, projects, and practice.

13.2 Go to Graduate School!

If you find the ideas and methods presented in this book interesting, perhaps you are the kind of person who should think about going to graduate school.[1] Technical skills age quickly without advanced training, and it can be difficult to find the time for professional training after joining the working world.

Graduate programs in data science are rapidly emerging from host departments of computer science, statistics, business, applied mathematics, and the like. Which type of program is most appropriate for you depends upon your undergraduate training and life experiences. Depending upon their focus, data science programs will differ wildly in the computational and statistical background they expect. Generally speaking, the technically-hardest programs in terms of programming, machine learning, and statistics provide the best preparation for the future. Be aware of grandiose claims from programs which minimize these demands.

[1] Good for you if you are already there!

Most of these programs are at the masters level, but outstanding students who are able to make the life commitment should consider the possibility of undertaking a Ph.D degree. Graduate study in computer science, machine learning, or statistics involves courses in advanced topics that build upon what you learned as an undergraduate, but more importantly you will be doing new and original research in the area of your choice. All reasonable American doctoral programs will pay tuition and fees for all accepted Ph.D students, plus enough of a stipend to live comfortably if not lavishly.

If you have a strong computer science background and the right stuff, I would encourage you to continue your studies, ideally by coming to work with us at Stony Brook! My group does research in a variety of interesting topics in data science, as you can tell from the war stories. Please check us out at `http://www.data-manual.com/gradstudy`.

13.3 Professional Consulting Services

Algorist Technologies is a consulting firm that provides its clients with short-term, expert help in data science and algorithm design. Typically, an Algorist consultant is called in for one to three days worth of intensive on site discussion and analysis with the client's own development staff. Algorist has built an impressive record of performance improvements with several companies and applications, as well as expert witness services and longer-term consulting.

Visit `www.algorist.com/consulting` for more information on the services provided by Algorist Technologies.