

Big Book of MLOps

Jamie Ralph

2024-09-24

Table of contents

Preface	3
1 Introduction	4
2 Summary	5

Preface

My journey into MLOps started in late 2023 when I needed to deploy an R Shiny app at the company I was working for. Up to that point there had been no easy way to do this, but luckily for me the MLOps team had put together a Helm chart for deploying containerised applications. I wasn't part of the data science team, but MLOps very kindly helped me get my app deployed. At the time I had no idea what Kubernetes was or how Helm charts worked - I just containerised my app with Docker, filled in some YAML files, and there it was. I was given access to this mystical platform called 'ArgoCD' where I could look at little tiles that represented things I didn't understand. The only thing I knew for certain was that the little green heart was good, and the broken red heart was bad.

Around 6 months later I was given the chance to transition over to the MLOps team. I jumped at the chance and never regretted my decision, but I quickly found the learning curve to be steep (somewhere around 89 degrees?) and was overwhelmed with where to begin. A few months into the job I started a google doc - I called it "the big book of MLOPs" (O'Reilly got there first). This was my place to write down notes from the tutorials/courses/books etc. that I was using to gain the knowledge I needed for my job. I decided these notes should sit in a more public forum because a) they could help someone else and b) if my notes could be improved, someone can open a PR and improve them!

I hope you might find these notes useful in some, especially if you're just starting out in MLOps or looking to move into the field.

1 Introduction

This is a book created from markdown and executable code.

See @knuth84 for additional discussion of literate programming.

2 Summary

In summary, this book has no content whatsoever.