CHAPTER 46

Kubeflow and Kubeflow Pipelines

Machine learning is often and rightly viewed as the use of mathematical algorithms to teach the computer to learn tasks that are computationally infeasible to program as a set of specified instructions. However, it turns out that these algorithms constitute only a small fraction of the overall learning pipeline from an engineering perspective. Building high-performant and dynamic learning models includes a number of other critical components. These components actually dominate the space of concerns for delivering an end-to-end machine learning product.

A typical machine learning production pipeline looks like the illustration in Figure 46-1.

E. Bisong, *Building Machine Learning and Deep Learning Models on Google Cloud Platform*,
https://doi.org/10.1007/978-1-4842-4470-8_46

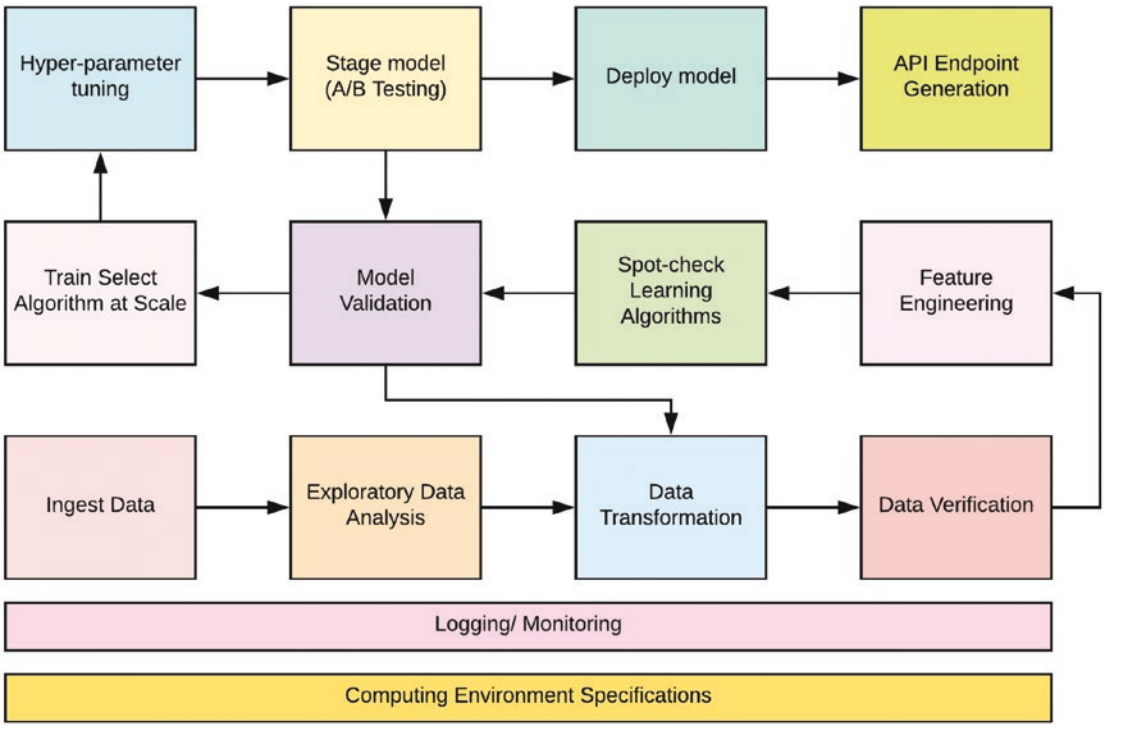

Figure 46-1. *Machine learning production pipeline*

From the preceding diagram, observe that the process flow in the pipeline is iterative. This repetitive pattern is central to machine learning experimentation, design, and deployment.

The Efficiency Challenge

It is easy to recognize that the pipeline requires a significant amount of development operations for the seamless transition from one component to another when building a learning model. This interoperability of parts has given rise to Machine Learning Ops, also known as MLOps. The term is coined as an amalgam of Machine Learning and DevOps.

The conventional way of doing machine learning is to perform all of the Experiment and development work in Jupyter notebooks, and the model is exported and sent off to the software development team for deployment and endpoint generation for integration