

Overview of ML Pipelines

Congratulations! Your model is ready for deployment to a production ML pipeline. This section of the course introduces testing guidelines for ML pipelines. However, this section *does not* demonstrate these guidelines because such a demo is not possible in a sandboxed environment.

You will learn about:

- Writing appropriate tests for launch and production.
- Detecting failure modes in your ML pipeline using tests.
- Evaluating your model quality in production.

What is an ML Pipeline?

An ML pipeline consists of several components, as the diagram shows. We'll become familiar with these components later. For now, notice that the “Model” (the black box) is a small part of the pipeline infrastructure necessary for production ML.

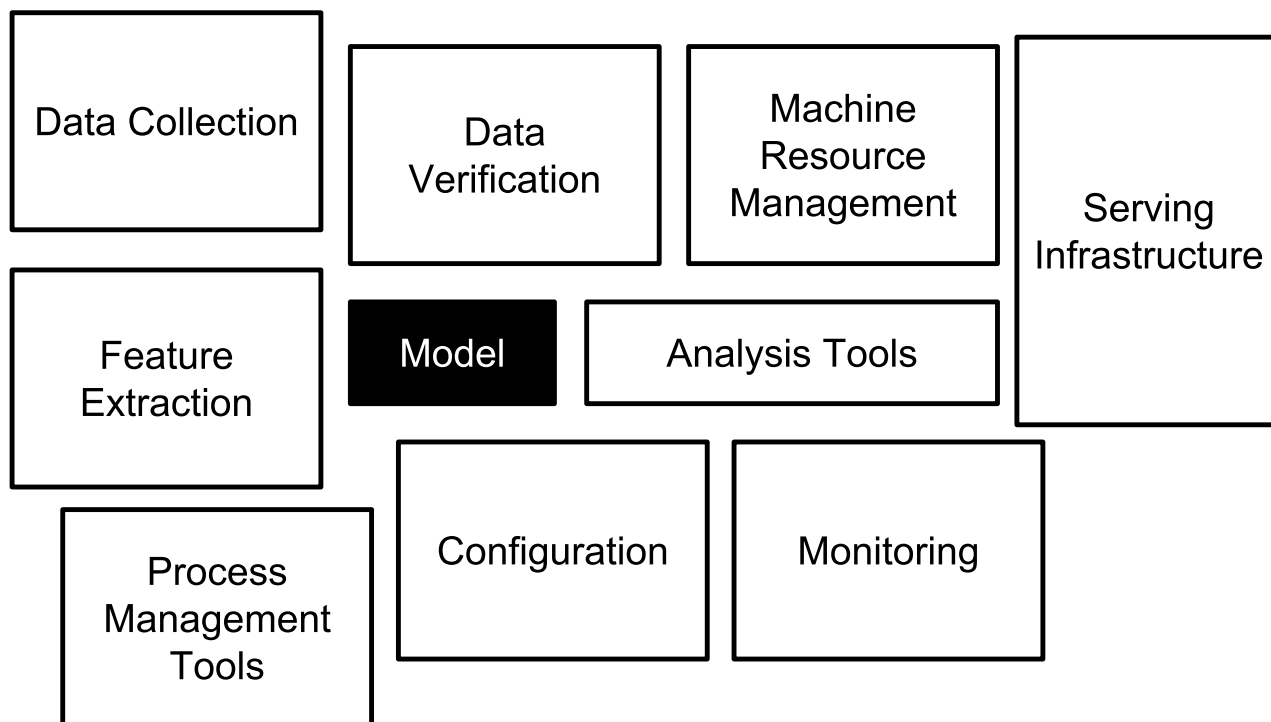


Figure 1: A schematic of a typical machine learning pipeline.

Role of Testing in ML Pipelines

In software development, the ideal workflow follows test-driven development (TDD). However, in ML, starting with tests is not straightforward. Your tests depend on your data, model, and problem. For example, before training your model, you cannot write a test to validate the loss. Instead, you discover the achievable loss during model development and *then* test new model versions against the achievable loss.

You need tests for:

- Validating input data.
- Validating feature engineering.
- Validating quality of new model versions.
- Validating serving infrastructure.
- Testing integration between pipeline components.

[Previous](#)

← [Programming Exercise](#)

(/machine-learning/testing-debugging/common/programming-exercise)

[Next](#)

[Testing for Deployment](#) (/machine-learning/testing-debugging/pipeline/deploying) →

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2019-03-06 UTC.