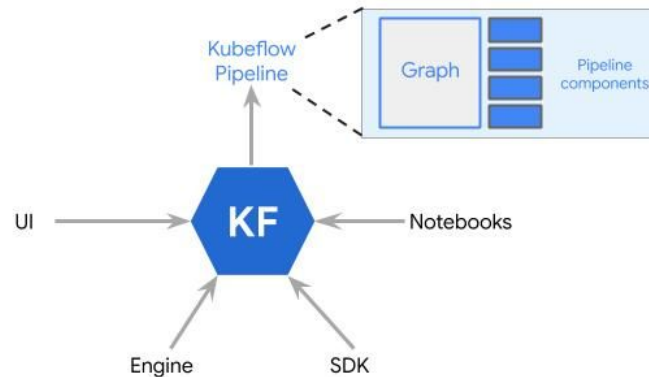


Kubeflow Pipelines

Proprietary + Confidential

Kubeflow Pipelines package multi-step ML workflows



Google Cloud

Training and Certification

Kubeflow is an open project that packages machine learning code for Kubernetes.

Kubeflow Pipelines is a platform for composing, deploying, and managing end-to-end machine learning workflows.

The main components of Kubeflow Pipelines includes:

A user interface (UI) for managing and tracking experiments, jobs, and runs.

An engine for scheduling multi-step ML workflows.

An SDK for defining and manipulating pipelines and components.

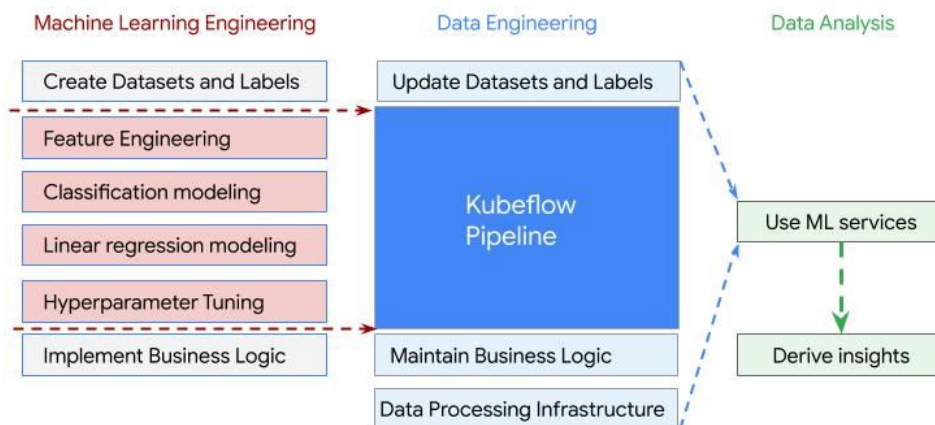
Notebooks for interacting with the system using the SDK.

These tools are used to define, experiment with, run, and share pipelines.

The pipeline consists of pipeline components, which are ML steps, that are assembled into a graph that describes the execution pattern.

Key benefits are reusability and portability.

Kubeflow Pipelines provides reusability and portability



Google Cloud

Training and Certification

Kubeflow Pipelines reusability helps separate the work which enables people to specialize.

In this example a Machine Learning Engineer can focus on Feature Engineering and Linear Regression Modeling and Hyperparameter Tuning.

Their solutions are bundled up into the Kubeflow Pipeline which can be used by a Data Engineer as part of a data engineering solution.

The solution appears as a service that is used by a Data Analyst to derive business insights.