

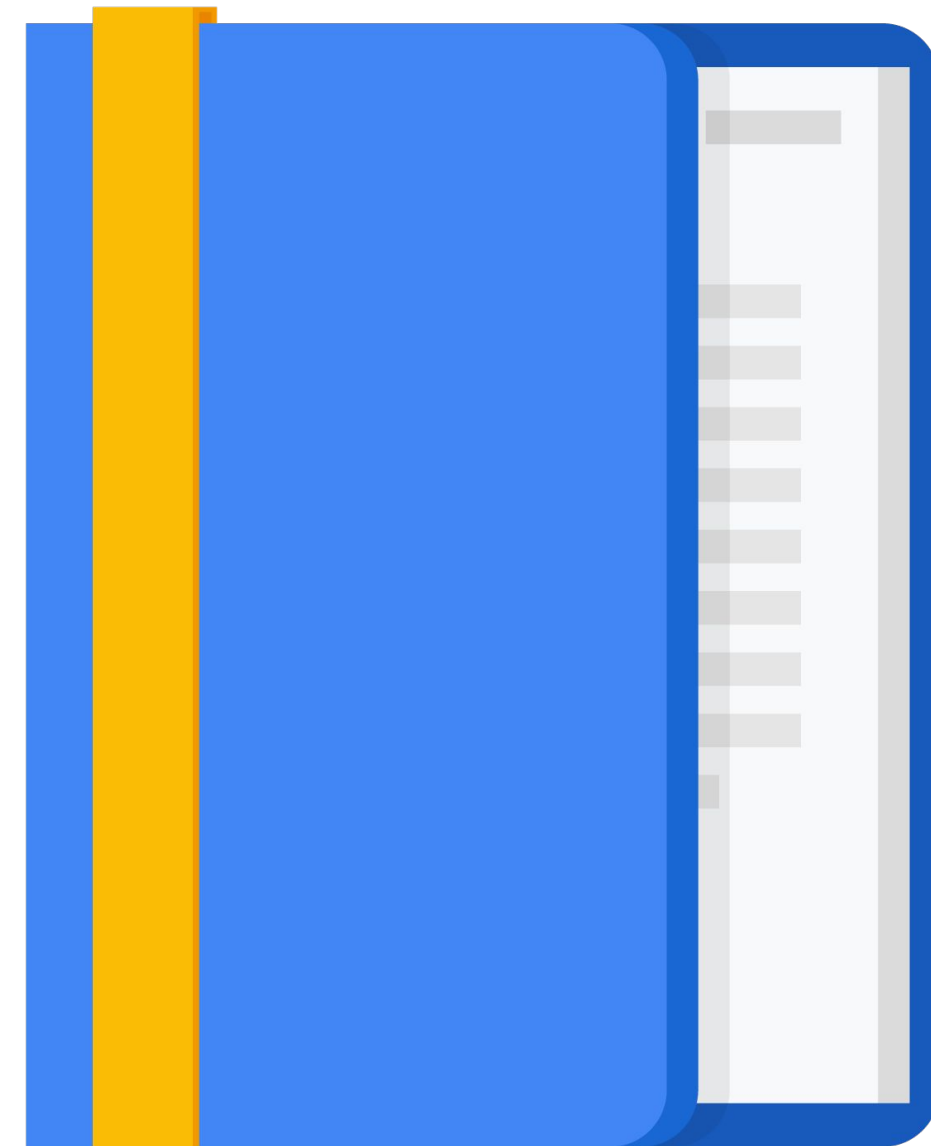


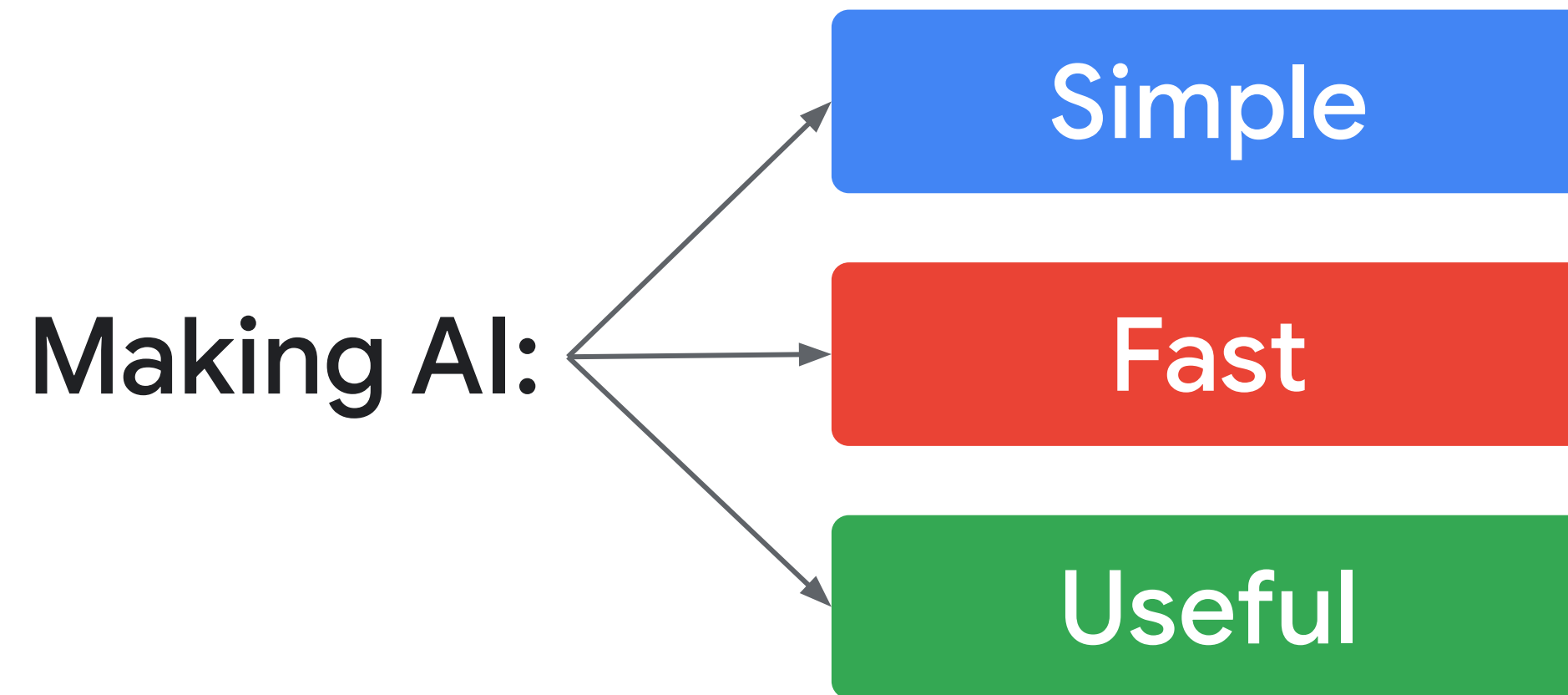
Introduction to ML Pipelines



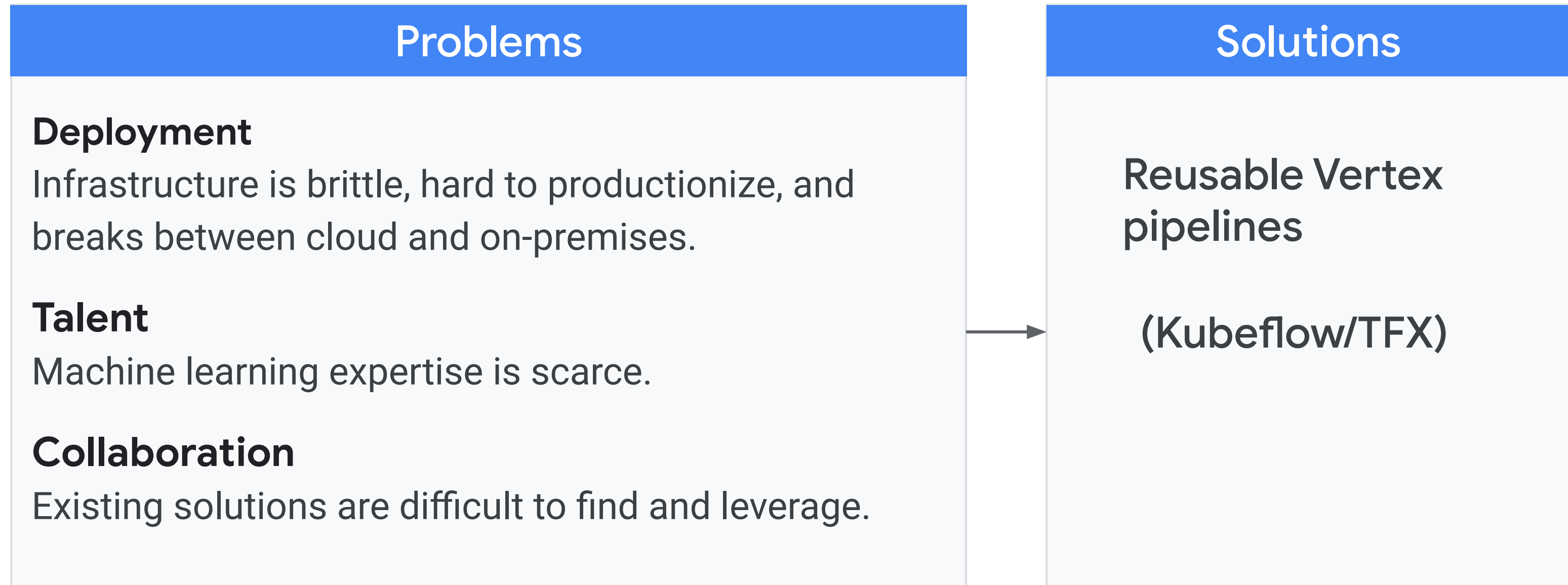
Agenda

- Overview
- Introduction to Vertex Pipelines
- Vertex Pipelines: Concepts
- Vertex Pipelines: When to Use?

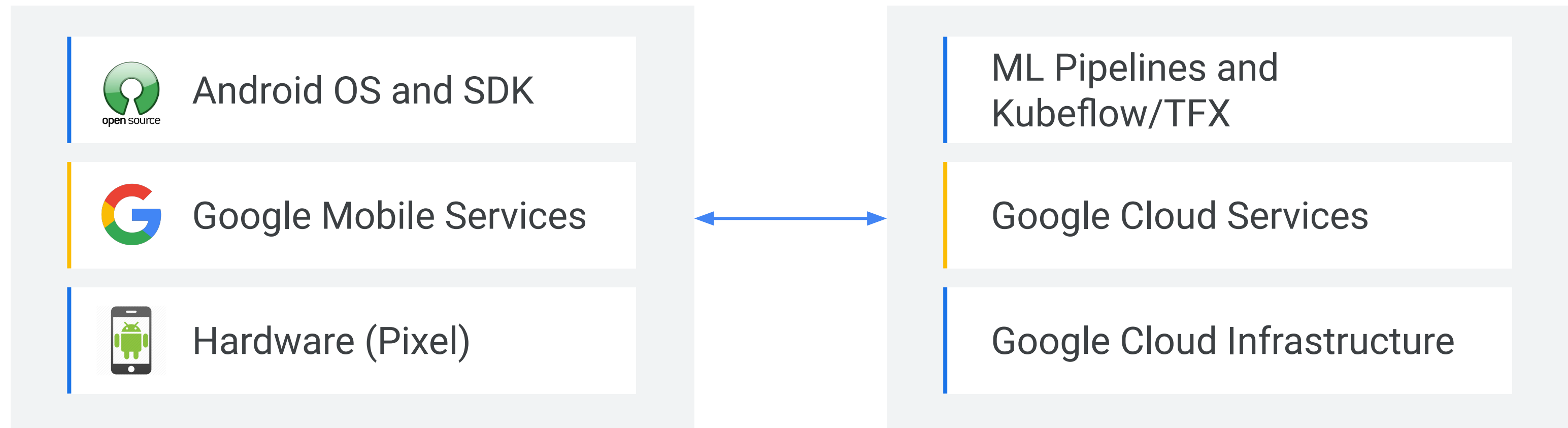




Few common AI problems

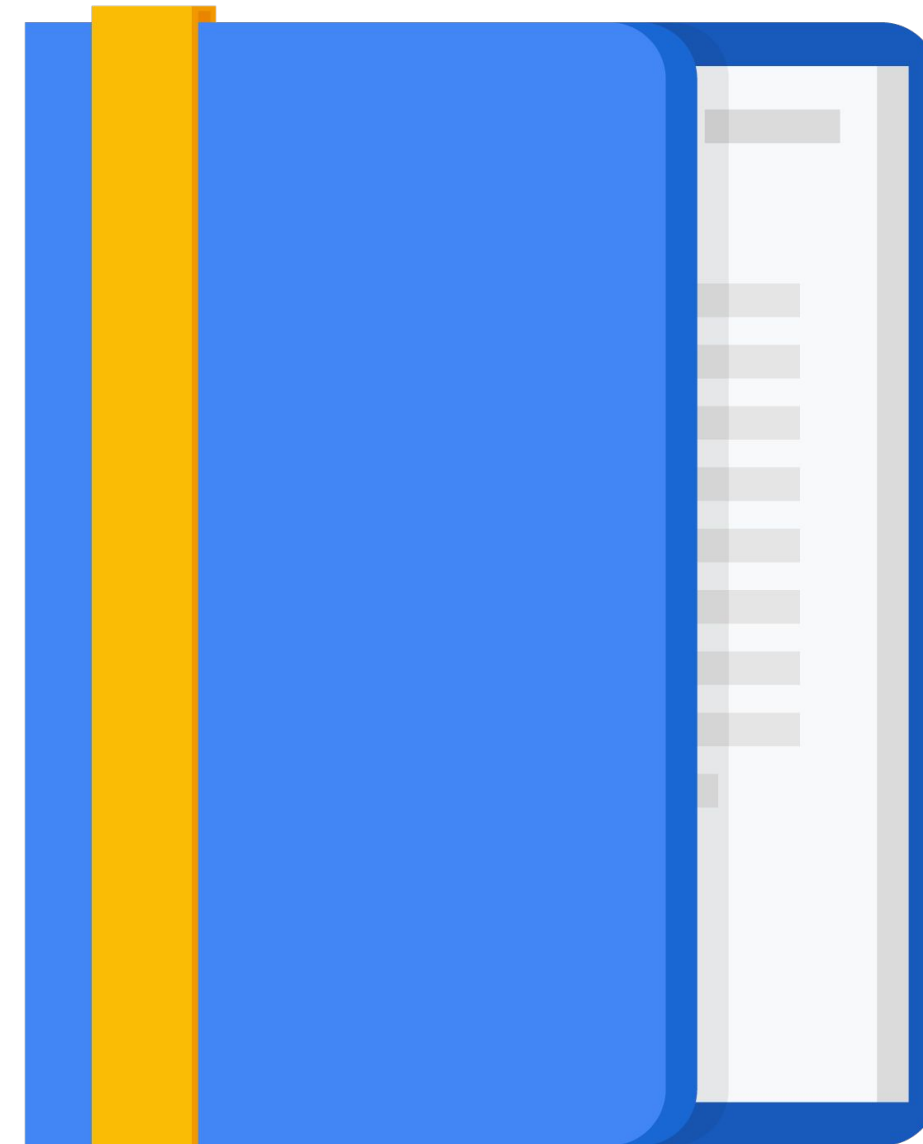


Android ecosystem analogy



Agenda

- Overview
- Introduction to Vertex Pipelines
- Vertex Pipelines: Concepts
- Vertex Pipelines: When to Use?



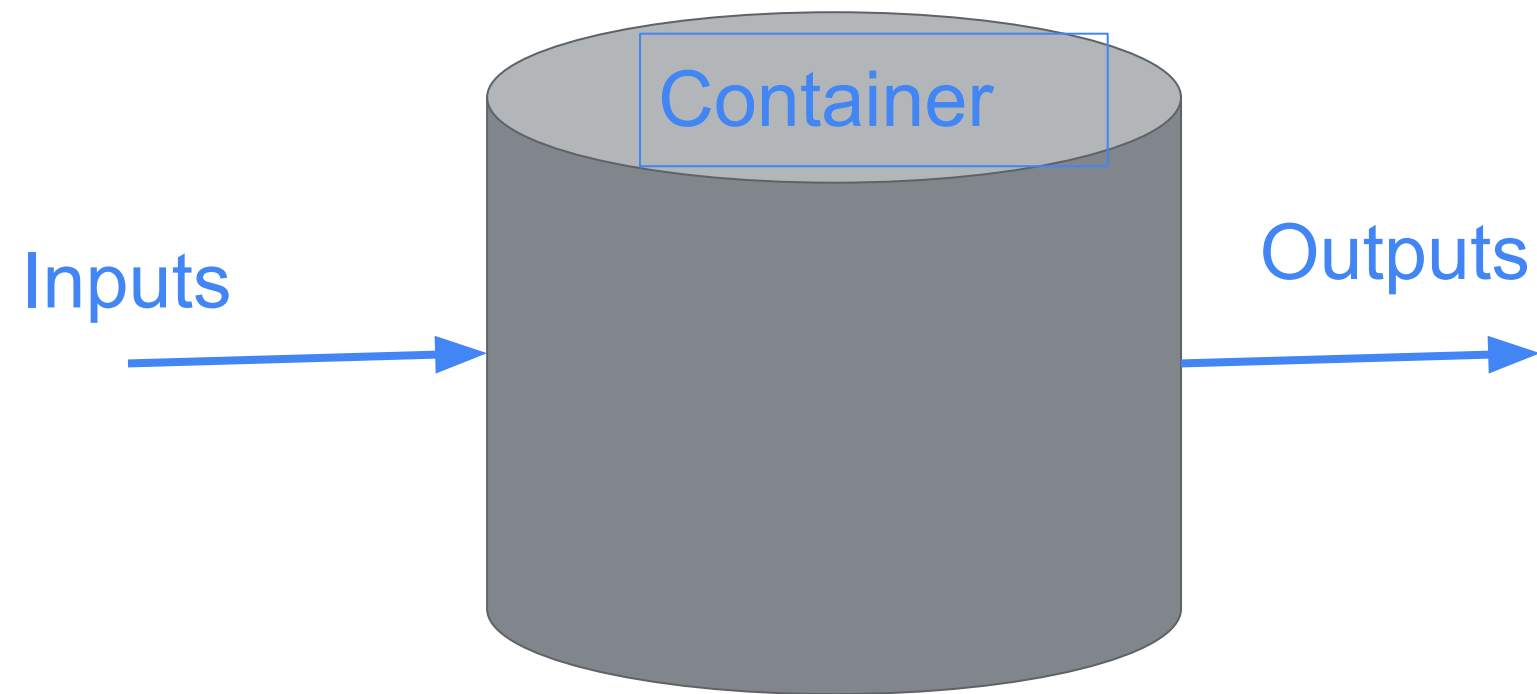
What is an ML pipeline?

ML task implemented as Docker containers orchestrated in a Directed Acyclic Graph of dependencies.

Kubeflow orchestrates the tasks and run the container on Kubernetes.

Task = Component = Operation (Op)

Understanding pipeline components



Example pipeline component

Container image location

gcr.io/example-components/...

Container image registry

Component container image

Inputs

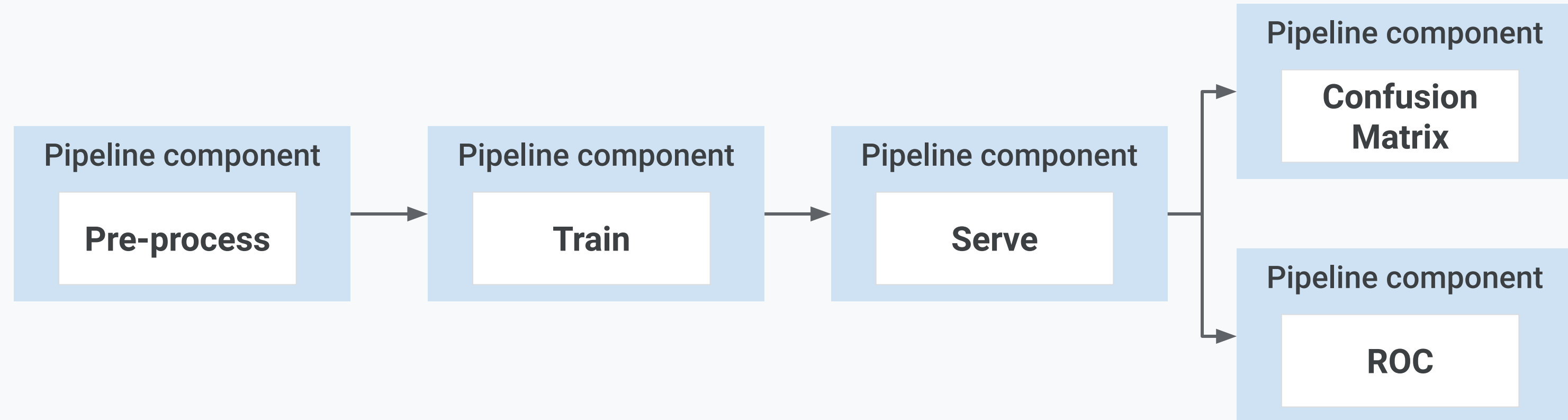
train_data=gs:examples/train.csv,...
eval_data=gs:examples/train.csv,...

Outputs

output=output.txt

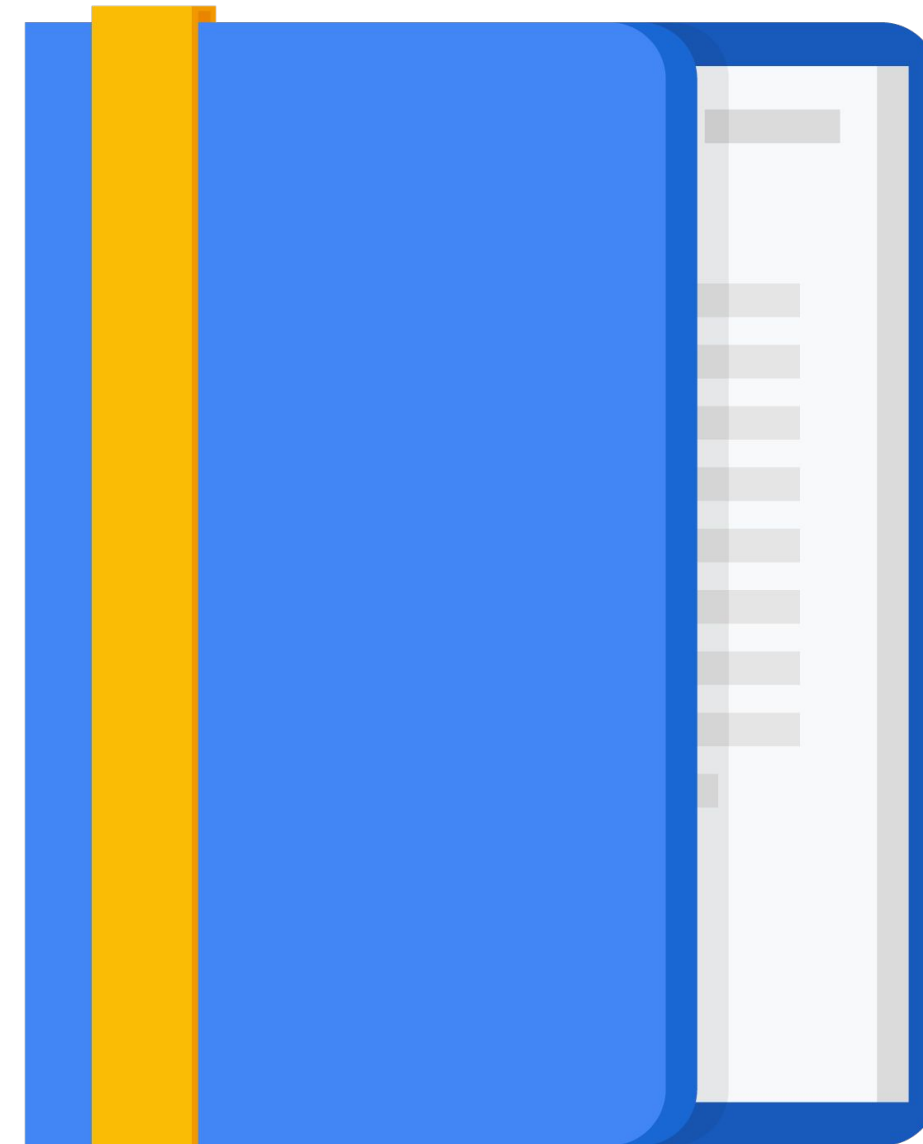
Understanding a pipeline workflow

Example pipeline



Agenda

- Overview
- Introduction to Vertex Pipelines
- Vertex Pipelines: Concepts
- Vertex Pipelines: When to Use?



AI Platform Pipelines implementation strategy

Kubeflow Pipelines

Through Kubeflow Pipelines SDK

- Lower-level ML framework–agnostic implementation
- Enables direct control of Kubernetes resource control
- Simple sharing of containerized components
- Use it for fully custom pipelines

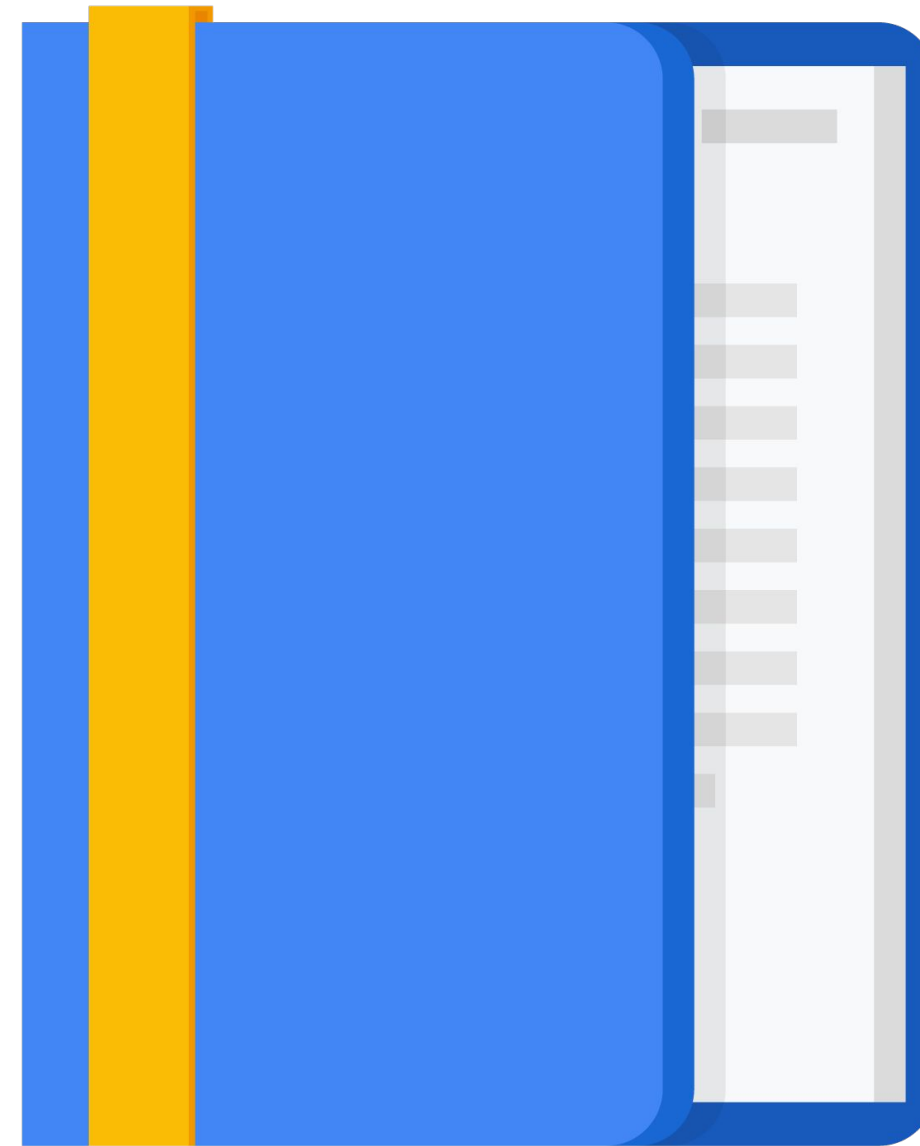
TensorFlow Extended (TFX)

Through TFX SDK

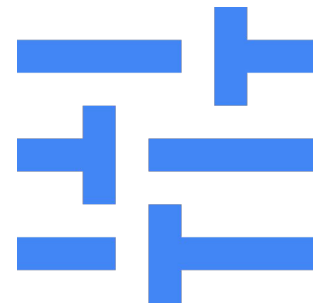
- Higher-level abstraction
- Prescriptive but customizable components with pre-defined ML types
- Brings Google best practices for robust/scalable ML workloads
- TF-based pipeline with customizable data pre-processing and training code

Agenda

- Overview
- Introduction to Vertex Pipelines
- Vertex Pipelines: Concepts
- Vertex Pipelines: When to Use?



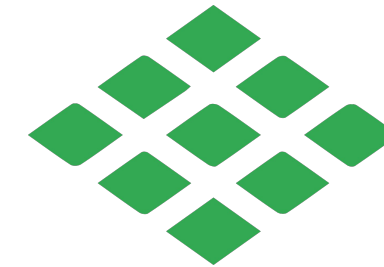
What does Vertex Pipelines enable?



**Workflow
orchestration**

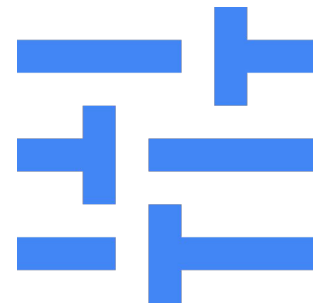


**Rapid, reliable,
repeatable
experimentation**



**Share, re-use,
and compose**

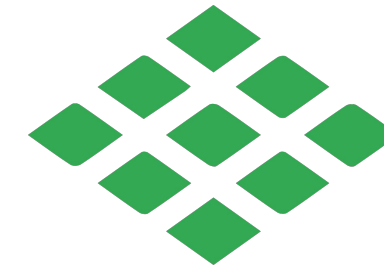
What does AI Platform Pipelines enable?



**Workflow
orchestration**

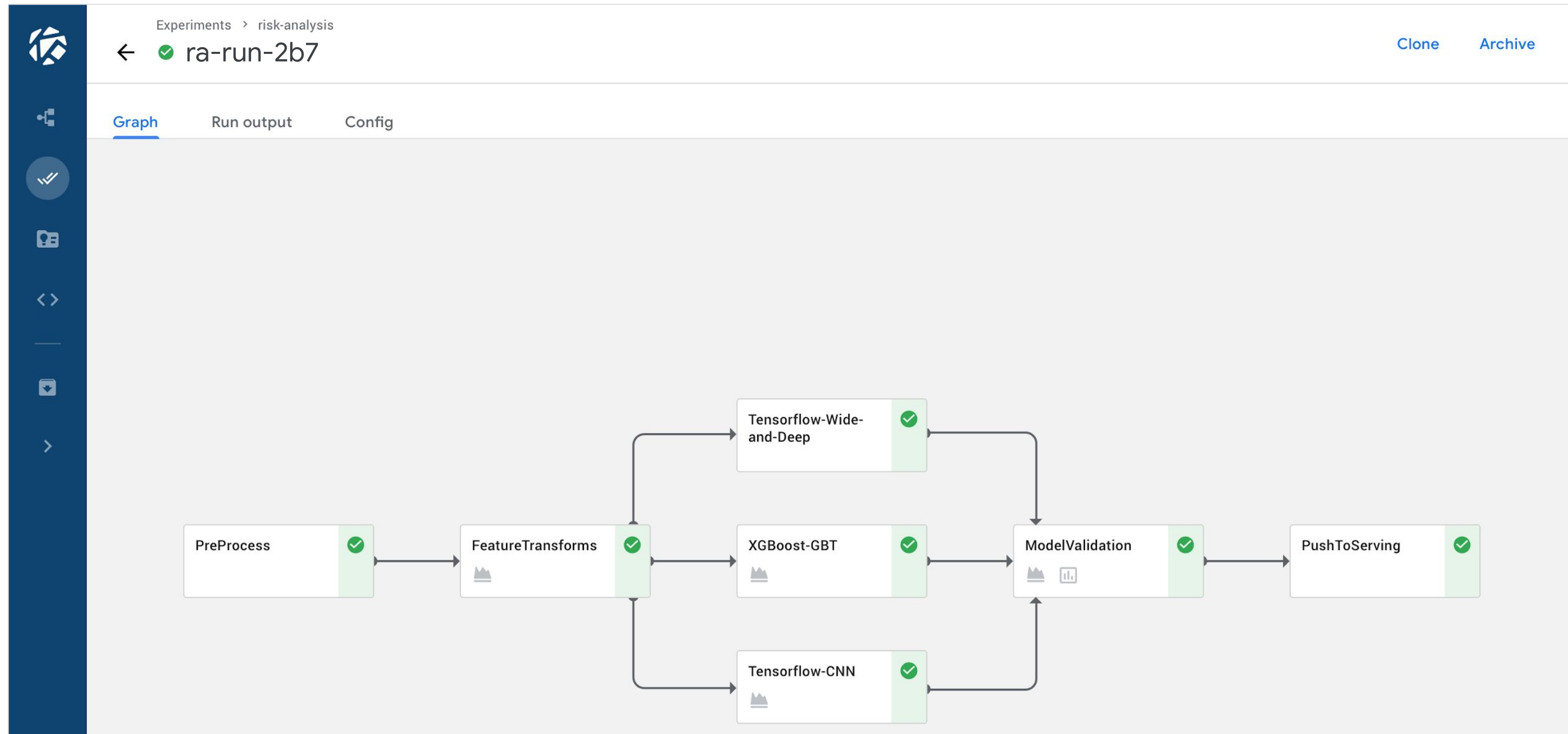


**Rapid, reliable,
repeatable
experimentation**

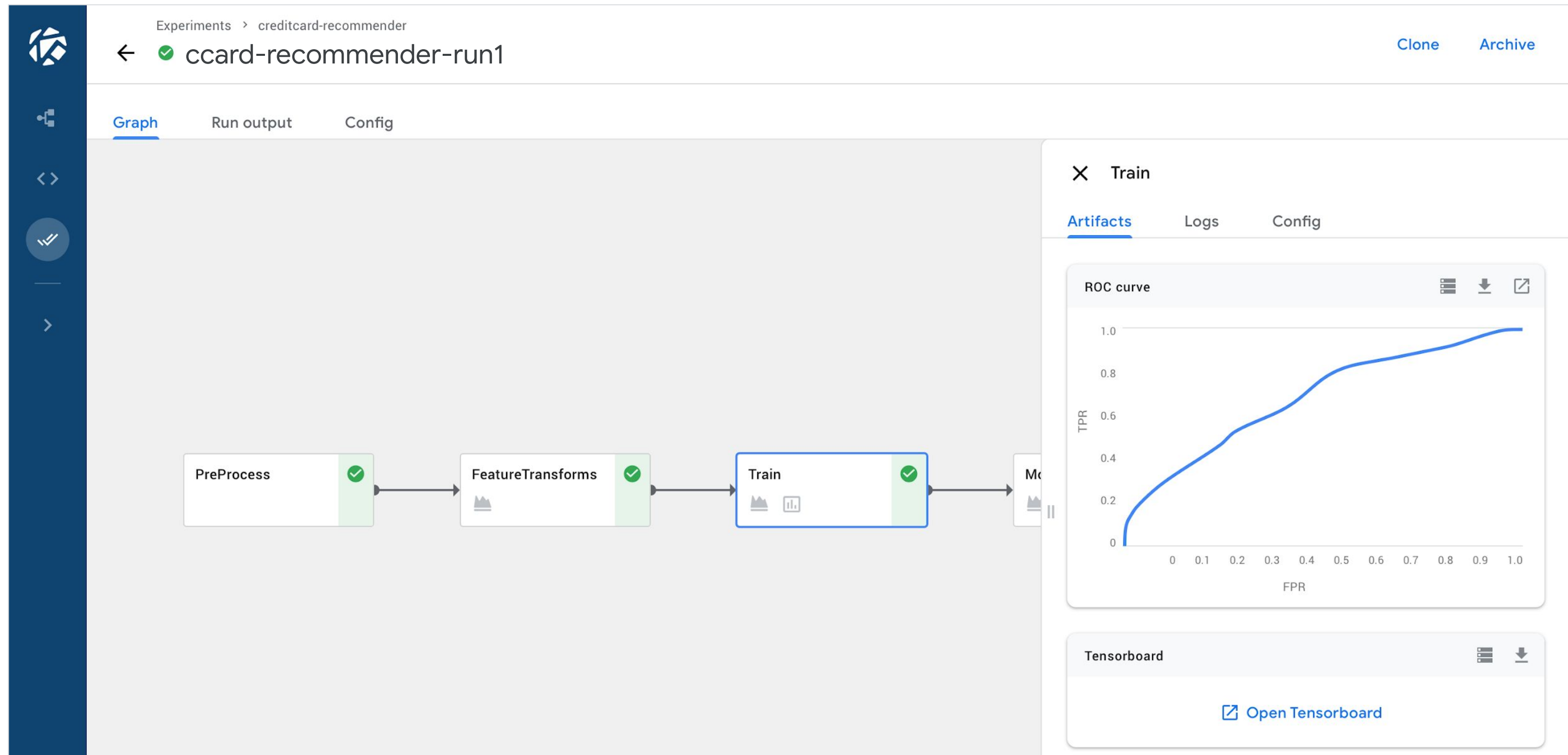


**Share, re-use,
and compose**






Visual depiction of pipeline topology



Rich visualizations of metrics



Access to all
config params,
inputs, and
outputs for each
run



Experiments > Product Image Classification

←

✔

Simple XGBoost Classifier

Graph [Config](#)

Run details

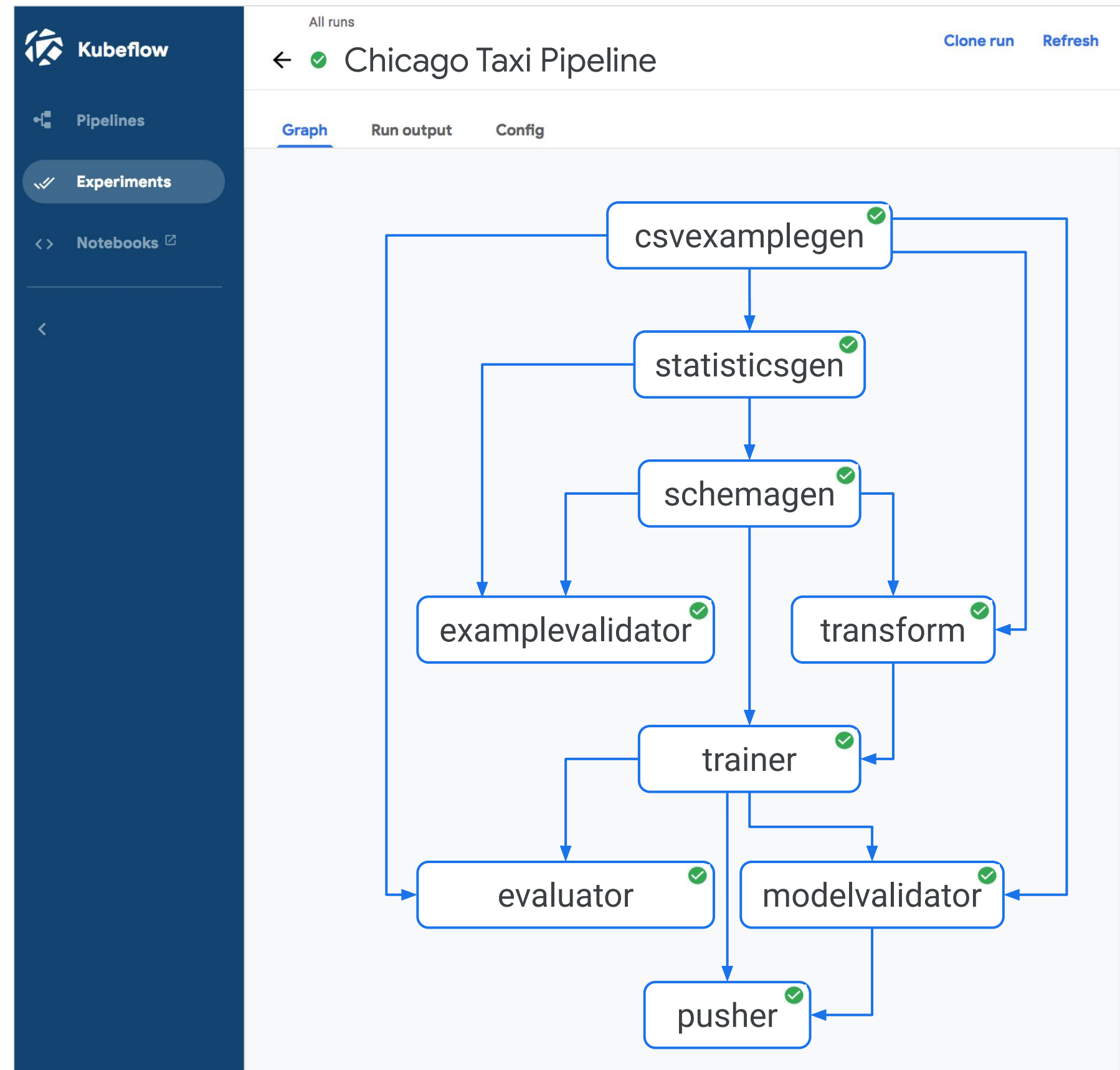
Status	Succeeded
Description	
Created at	11/25/2018, 12:56:44 PM
Started at	11/25/2018, 12:56:44 PM
Finished	11/25/2018, 12:16:37 PM
Duration	0:19:53

Run parameters

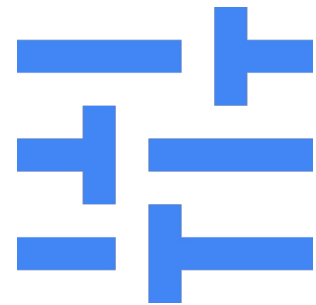
output	gs://mipipelines
project	foo2thebar
region	us-central1
train-data	gs://ml-pipeline-playground/sfpd/train.csv
eval-data	gs://ml-pipeline-playground/sfpd/eval.csv
schema	gs://ml-pipeline-playground/sfpd/schema.json
target	resolution
rounds	200

TFX + Kubeflow Pipelines

Robust extensible platform and **best-in-class** components for typical stages of good ML workflow



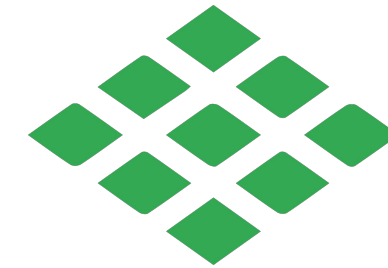
What does Vertex Pipelines enable?



Workflow
orchestration

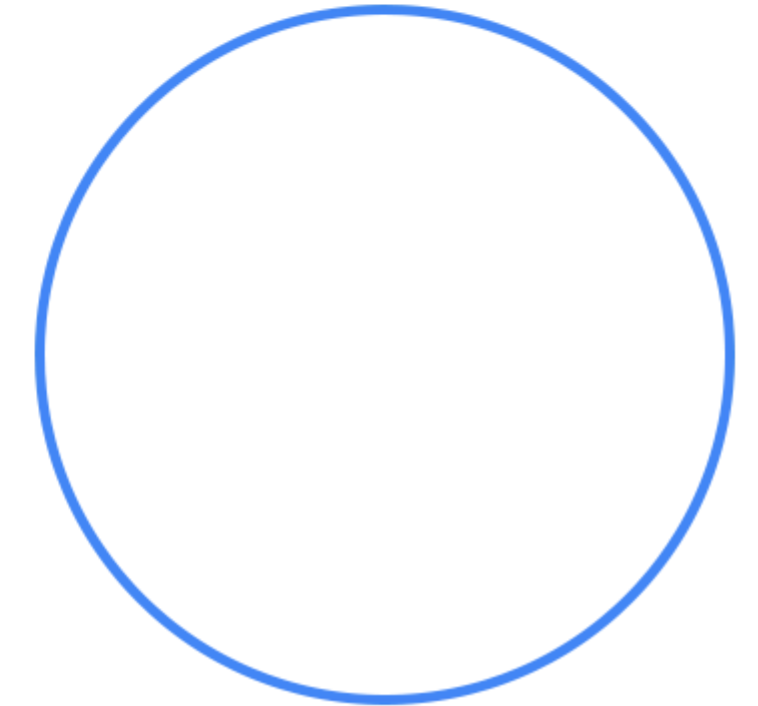
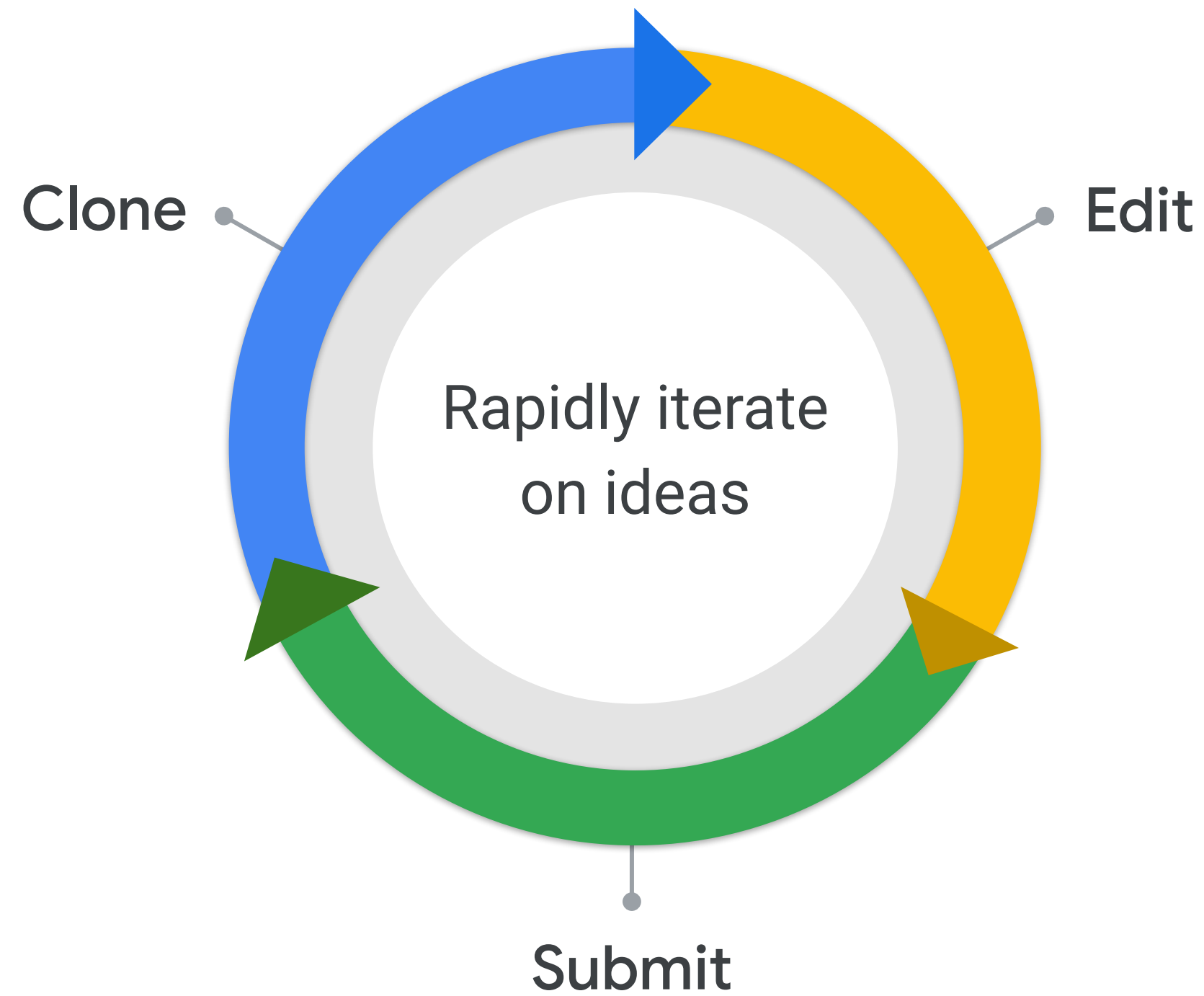


Rapid, reliable,
repeatable
experimentation




Share, re-use,
and compose

Rapid, reliable experimentation



Select any Run to see all params and metrics



Experiments > creditcard-recommender

← ✓ ccard-recommender-run1 Clone Archive

Graph Run output Config

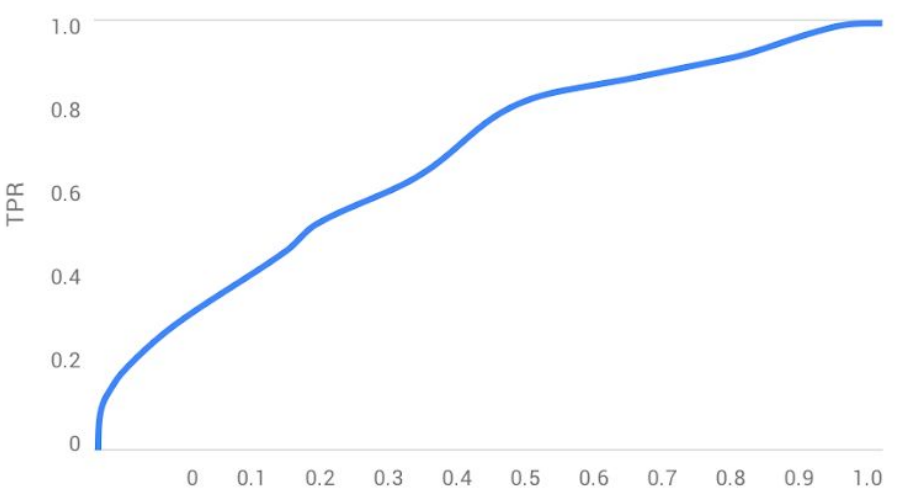
PreProcess ✓

FeatureTransforms ✓

Train ✓

Model ✓


ROC curve



Tensorboard

[Open Tensorboard](#)

Clone an existing pipeline



Experiments > creditcard-recommender

← ✓ ccard-recommender-run1

Clone

Archive

Graph

Run output

Config

PreProcess

FeatureTransforms

Train

Model

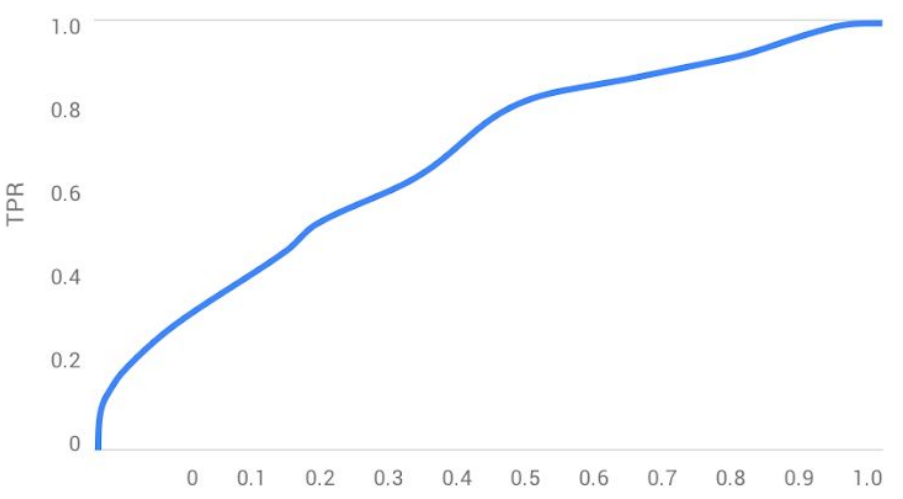
Train

Artifacts

Logs

Config

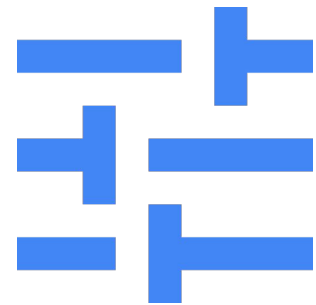
ROC curve



Tensorboard

Open Tensorboard

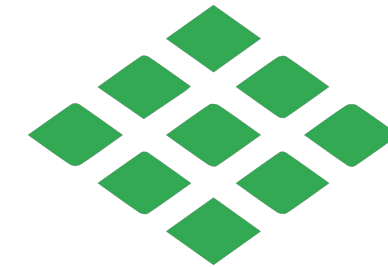
What does Vertex Pipelines enable?



**Workflow
orchestration**



**Rapid, reliable,
repeatable
experimentation**



**Share, re-use,
and compose**

Share, re-use, and compose

- Re-use instead of re-implement
- Easy to compose or swap
 - Standard artifact types
 - Standard interfaces
 - Conventions and best practices

A lot of reusable
components

+

Composable
pipelines

