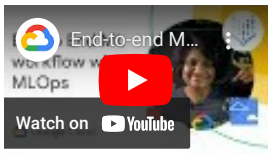
MLOps on Vertex AI

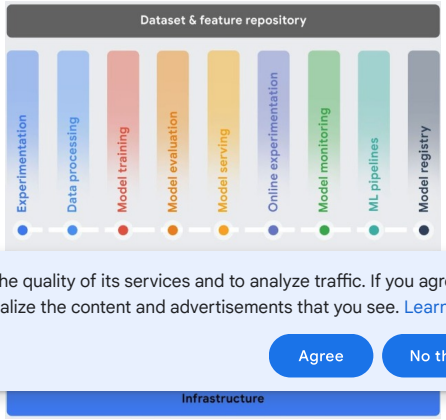
This section describes Vertex AI services that help you implement *Machine learning operations (MLOps)* with your machine learning (ML) workflow.



After your models are deployed, they must keep up with changing data from the environment to perform optimally and stay relevant. MLOps is a set of practices that improves the stability and reliability of your ML systems.

Vertex AI MLOps tools help you collaborate across AI teams and improve your models through predictive model monitoring, alerting, diagnosis, and actionable explanations. All the tools are modular, so you can integrate them into your existing systems as needed.

For more information about MLOps, see Continuous delivery and automation pipelines in machine learning and the Practitioners Guide to MLOps.



- **Orchestrate workflows**: Manually training and serving your models can be time-consuming and error-prone, especially if you need to repeat the processes many times.
  - Vertex AI Pipelines helps you automate, monitor, and govern your ML workflows.
- **Track the metadata used in your ML system**: In data science, it's important to track the parameters, artifacts, and metrics used in your ML workflow, especially when you repeat the workflow multiple times.
  - Vertex ML Metadata lets you record the metadata, parameters, and artifacts that are used in your ML system. You can then query that metadata to help analyze, debug, and audit the performance of your ML system or the artifacts that it produces.
- **Identify the best model for a use case**: When you try new training algorithms, you need to know which trained model performs the best.
  - Vertex AI Experiments lets you track and analyze different model architectures, hyper-parameters, and training environments to identify the best model for your use case.

    - Vertex AI TensorBoard helps you track, visualize, and compare ML experiments to measure how well your models perform.
- **Manage model versions**: Adding models to a central repository helps you keep track of model versions.
    - Vertex AI Model Registry provides an overview of your models so you can better organize, track, and train new versions. From Model Registry, you can evaluate models, deploy models to an endpoint, create batch predictions, and view details about specific models and model versions.
- **Manage features**: When you re-use ML features across multiple teams, you need a quick and efficient way to share and serve the features.
    - Vertex AI Feature Store provides a centralized repository for organizing, storing, and serving ML features. Using a central featurestore enables an organization to re-use ML features at scale and increase the velocity of developing and deploying new ML applications.
- **Monitor model quality**: A model deployed in production performs best on prediction input data that is similar to the training data. When the input data deviates from the data used to train the model, the model's performance can deteriorate, even if the model itself hasn't changed.
    - Vertex AI Model Monitoring monitors models for training-serving skew and prediction drift and sends you alerts when the incoming prediction data skews too far from the training baseline. You can use the alerts and feature distributions to evaluate whether you need to retrain your model.

What's next

- Vertex AI interfaces