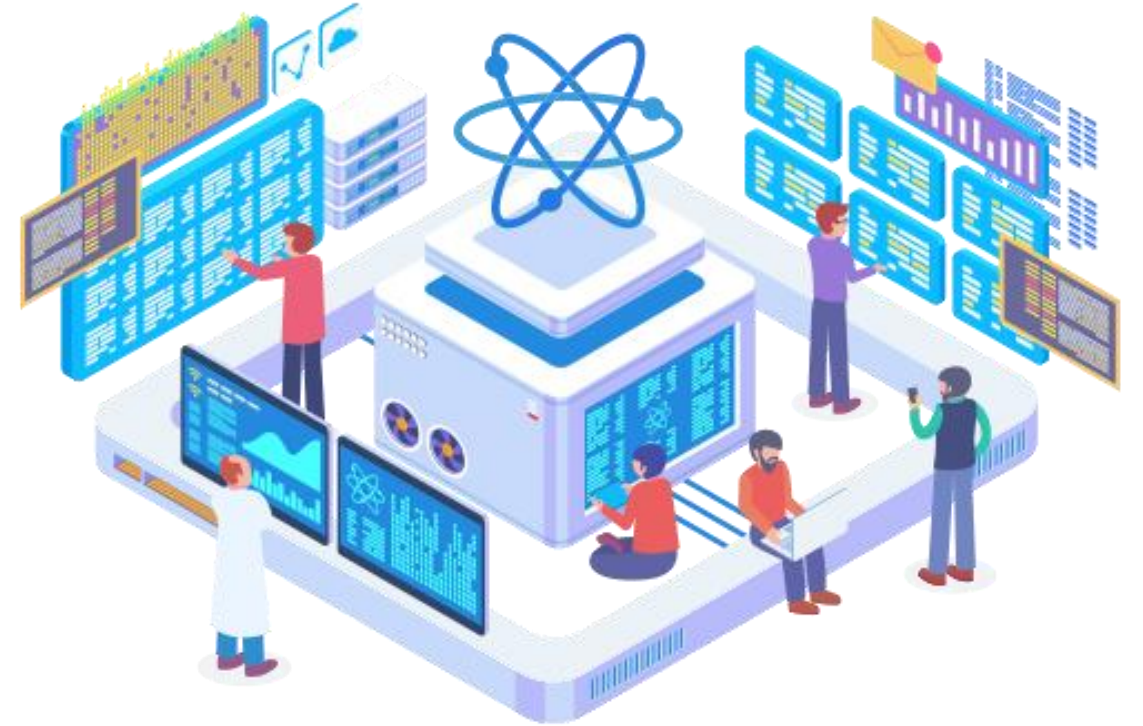


Advanced Certification in Data Science and Artificial Intelligence

Deploying Machine Learning Models on
Cloud (MLOps)



Agenda

01 What is MLOPs

02 Why do you need MLOPs

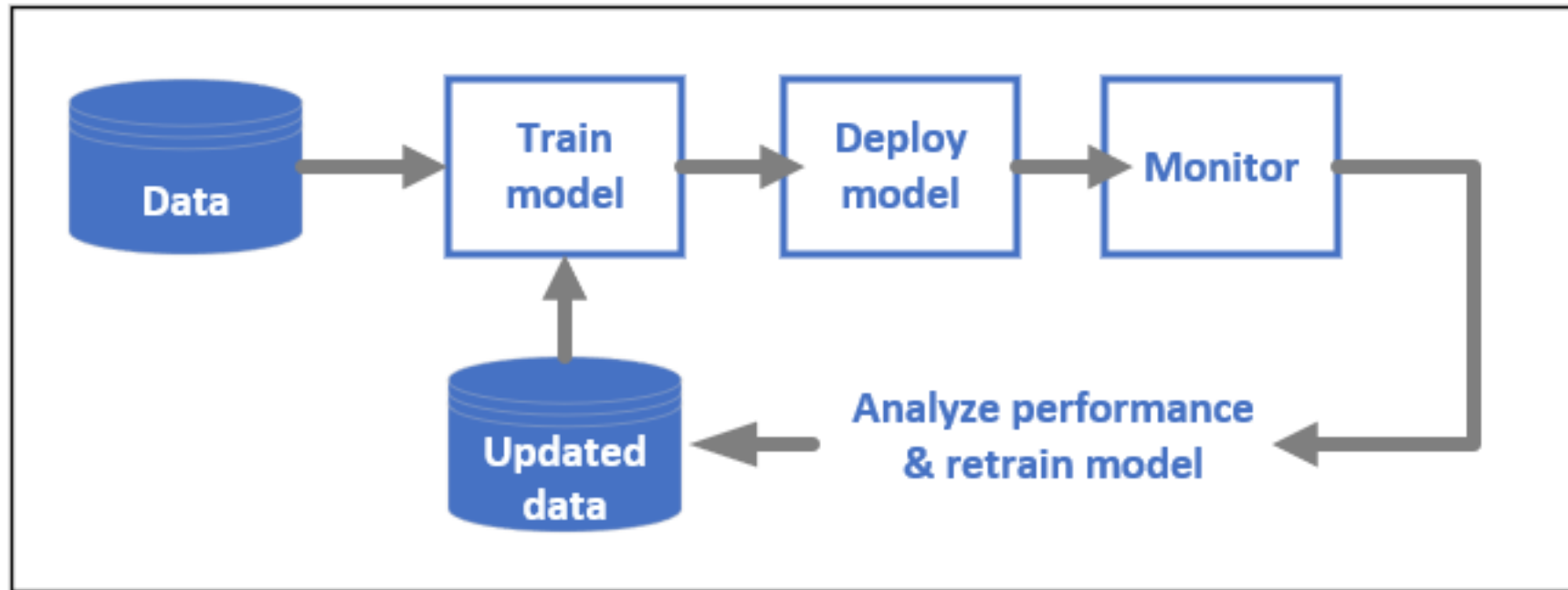
03 Process of MLOPs

04 Azure Machine Learning

What is MLOps?

What is MLOPs?

DevOps for machine learning models, often called MLOps, is a process for developing models for production. A model's lifecycle from training to deployment must be auditable if not reproducible.



What is MLOPs?

MLOps is a set of practices for collaboration and communication between data scientists and operations professionals. Applying these practices increases the quality, simplifies the management process, and automates the deployment of Machine Learning and Deep Learning models in large-scale production environments. It's easier to align models with business needs, as well as regulatory requirements.



Why do you need MLOPs?

Machine learning helps individuals and businesses deploy solutions that unlock previously untapped sources of revenue, save time, and reduce cost. Automating model development and deployment with MLOps means faster go-to-market times and lower operational costs. MLOps serves as the map to guide individuals, small teams, and even businesses to achieve their goals with minimal troubles. Besides that MLOps can help in resolving



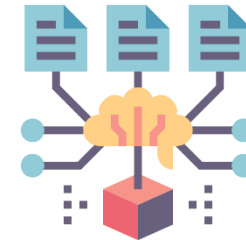
**Deployment
Issues**



**Monitoring
Issues**

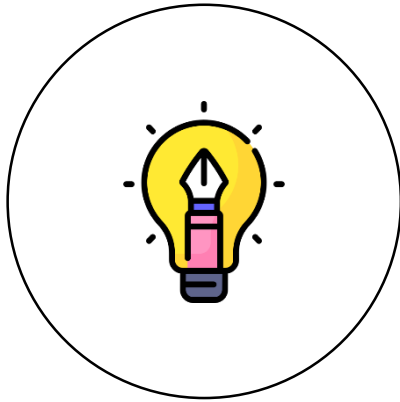


**Lifecycle
management issues**



**Model Governance
issues**

The complete MLOps process was divided into three broad phases



Design

- Requirements Engineering
- ML use-case prioritization
- Data availability



Model development

- Data Engineering
- ML Model Engineering
- Model Test and Validation

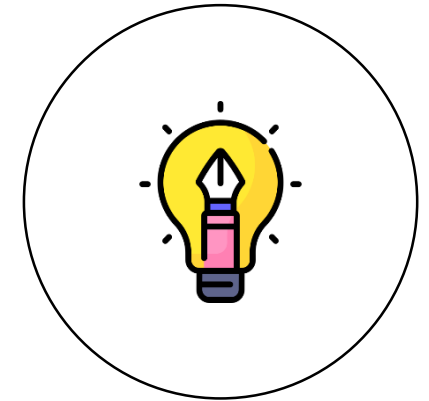


Operations

- ML Model deployment
- CI/CD pipelines
- Monitoring and Triggering

MLOps Process - Design

- The first phase is devoted to business understanding, data understanding and designing the ML-powered software.
- In this stage, we identify our potential user, design the machine learning solution to solve its problem, and assess the further development of the project. Mostly, we would act within two categories of problems - either increasing the productivity of the user or increasing the interactivity of our application.
- Initially, we define ML use-cases and prioritize them. The best practice for ML projects is to work on one ML use case at a time. Furthermore, the design phase aims to inspect the available data that will be needed to train our model and to specify the functional and non-functional requirements of our ML model.
- We should use these requirements to design the architecture of the ML-application, establish the serving strategy, and create a test suite for the future ML model.



MLOps Process – Model Development

- The follow-up phase “ML Experimentation and Development” is devoted to verifying the applicability of ML for our problem by implementing Proof-of-Concept for ML Model.
- Here, we run iteratively different steps, such as identifying or polishing the suitable ML algorithm for our problem, data engineering, and model engineering. The primary goal in this phase is to deliver a stable quality ML model that we will run in production.



MLOps Process – Operations

- The main focus of the “ML Operations” phase is to deliver the previously developed ML model in production by using established DevOps practices such as testing, versioning, continuous delivery, and monitoring.
- All three phases are interconnected and influence each other. For example, the design decision during the design stage will propagate into the experimentation phase and finally influence the deployment options during the final operations phase.

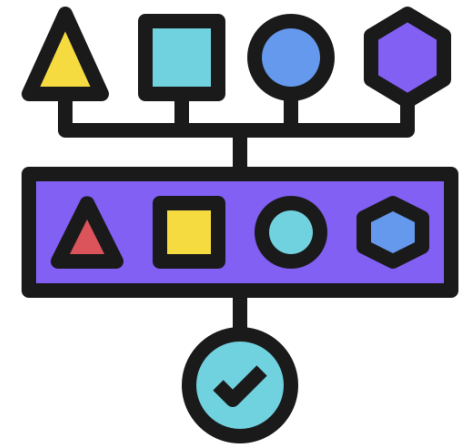


Understanding ML Pipelines

ML pipelines are portable, scalable ML workflows, based on containers. ML pipelines are composed of a set of input parameters and a list of tasks. Each task is an instance of a pipeline component.

ML Pipelines Can be used to

- Apply MLOps strategies to automate repeatable processes.
- Experiment by running an ML workflow with different sets of hyperparameters, number of training steps or iterations, etc.
- Reuse a pipeline's workflow to train a new model.

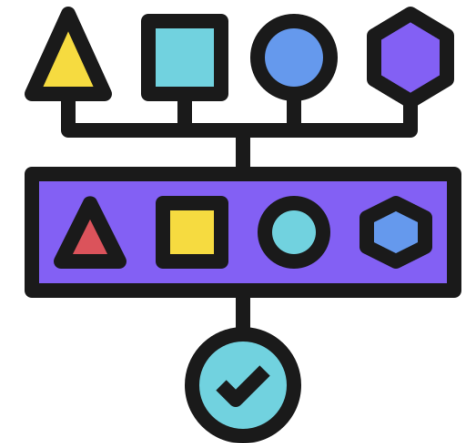


Understanding ML Pipelines

Pipeline Components

Pipeline components are self-contained sets of code that perform one step in a pipeline's workflow, such as data preprocessing, data transformation, model training, etc.

Components are composed of a set of input parameters, a set of outputs, and the location of a container image. A component's container image is a package that includes the component's executable code and a definition of the environment that the code runs in.



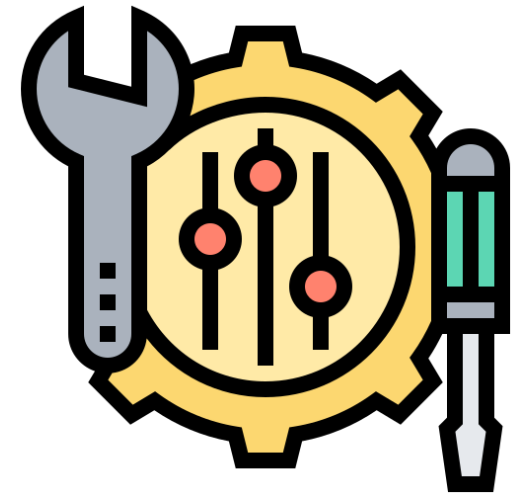
What is hyperparameter tuning?

Before we see the hyperparameter tuning, first let's see what is hyperparameter

What is hyperparameter?

Hyperparameters are the ones which contains the data that govern the training process itself. Usually when training the model, the training application handles three types of data

- **Input data**
 - Used during training to configure your model to accurately make predictions about new instances of similar data. Never directly became part of the model
- **The model's parameter**
 - Differentiates your particular model from other models of the same type working on similar data.
- **The hyperparameters**
 - Not directly related to the training data. They are configuration variables. It is important to note that parameters change during a training job, while hyperparameters are usually constant during a job.

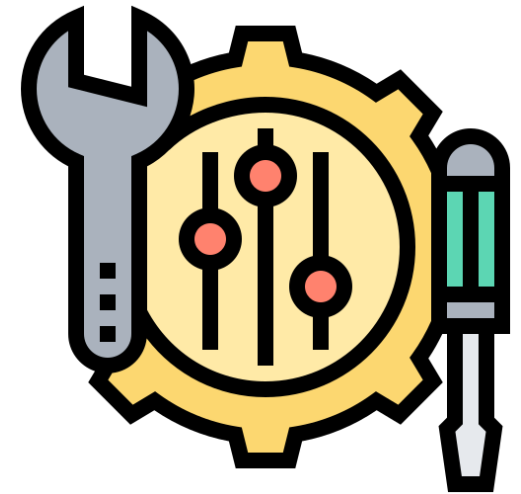


What is hyperparameter tuning?

How hyperparameter tuning works?

Hyperparameter tuning works by running multiple *trials* in a single training job. Each trial is a complete execution of your training application with values for your chosen hyperparameters, set within limits you specify. When the job is finished, you can get a summary of all the trials along with the most effective configuration of values according to the criteria you specify.

Hyperparameter optimization, or hyperparameter tuning, can be a tedious task. **Azure Machine Learning can automate this task for arbitrary parameterized commands with little modification to your job definition.** Results are visualized in the studio.



What is Azure Machine Learning?

What is Azure Machine Learning?



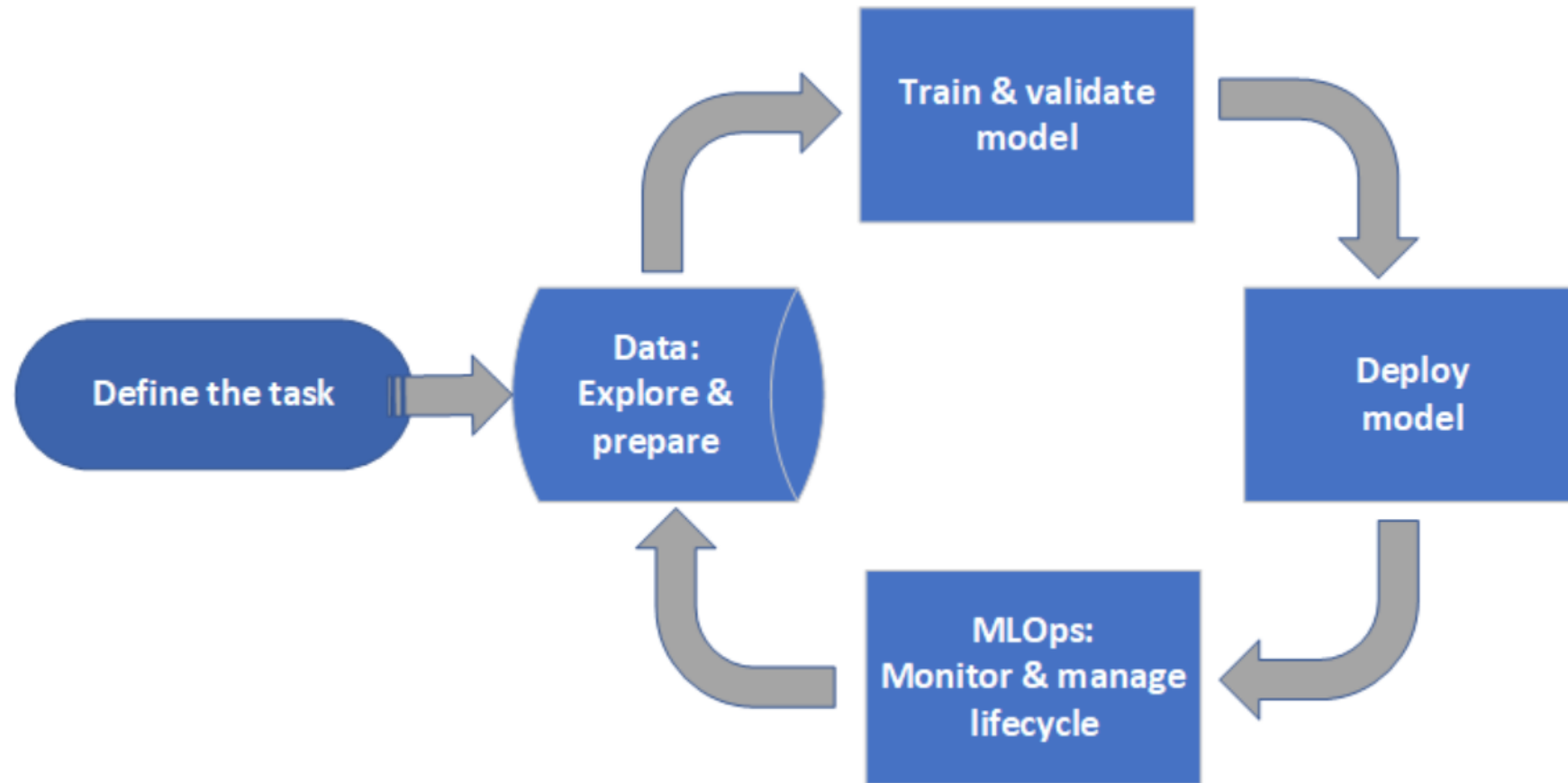
Azure Machine Learning is a cloud service for accelerating and managing the machine learning project lifecycle. Machine learning professionals, data scientists, and engineers can use it in their day-to-day workflows: Train and deploy models, and manage MLOps.



You can create a model in Azure Machine Learning or use a model built from an open-source platform, such as Pytorch, TensorFlow, or scikit-learn. MLOps tools help you monitor, retrain, and redeploy models.

Project Lifecycle

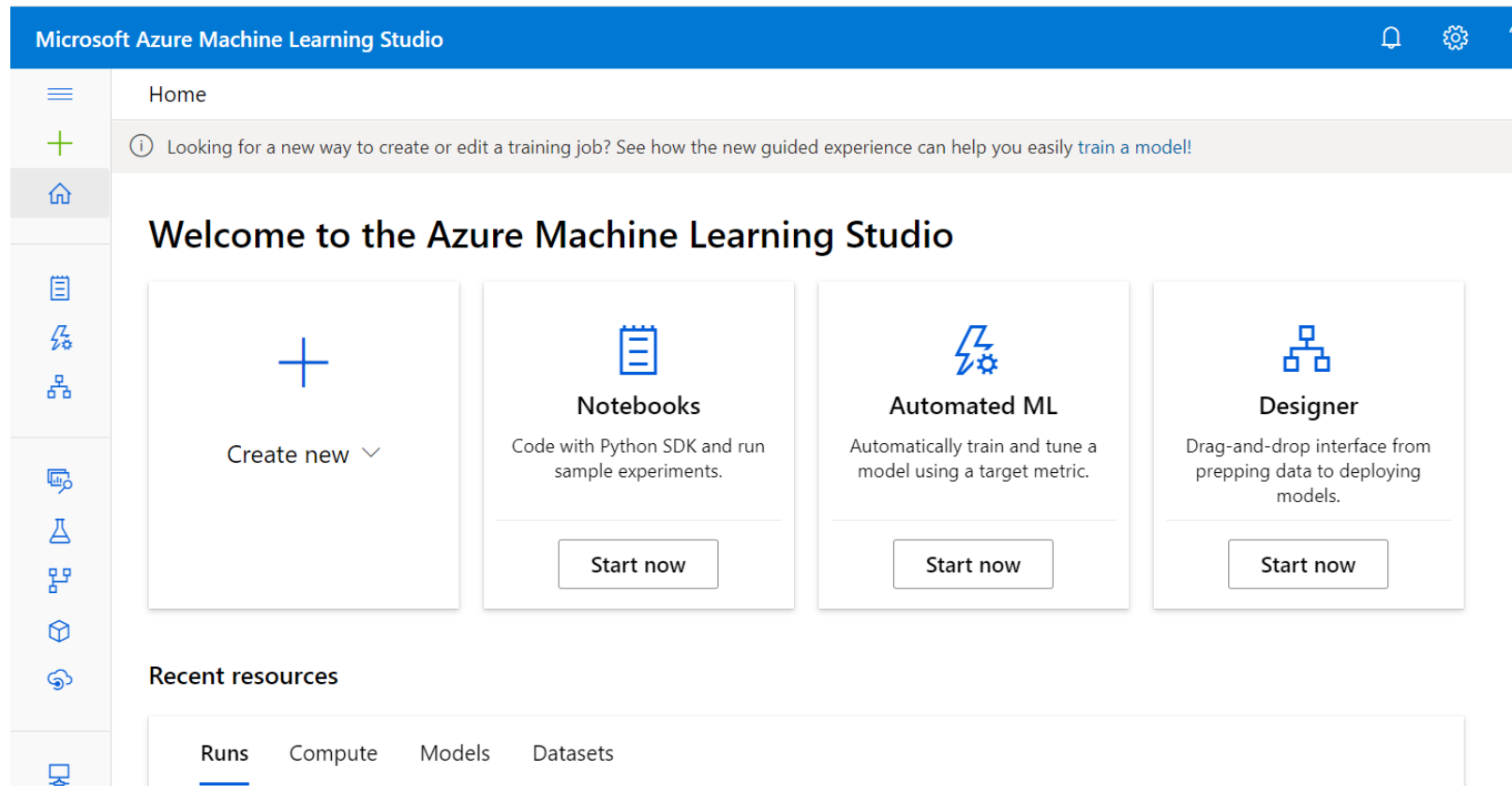
The project lifecycle can vary by project, but the common cycle looks like this:



What is Azure Machine Learning studio?

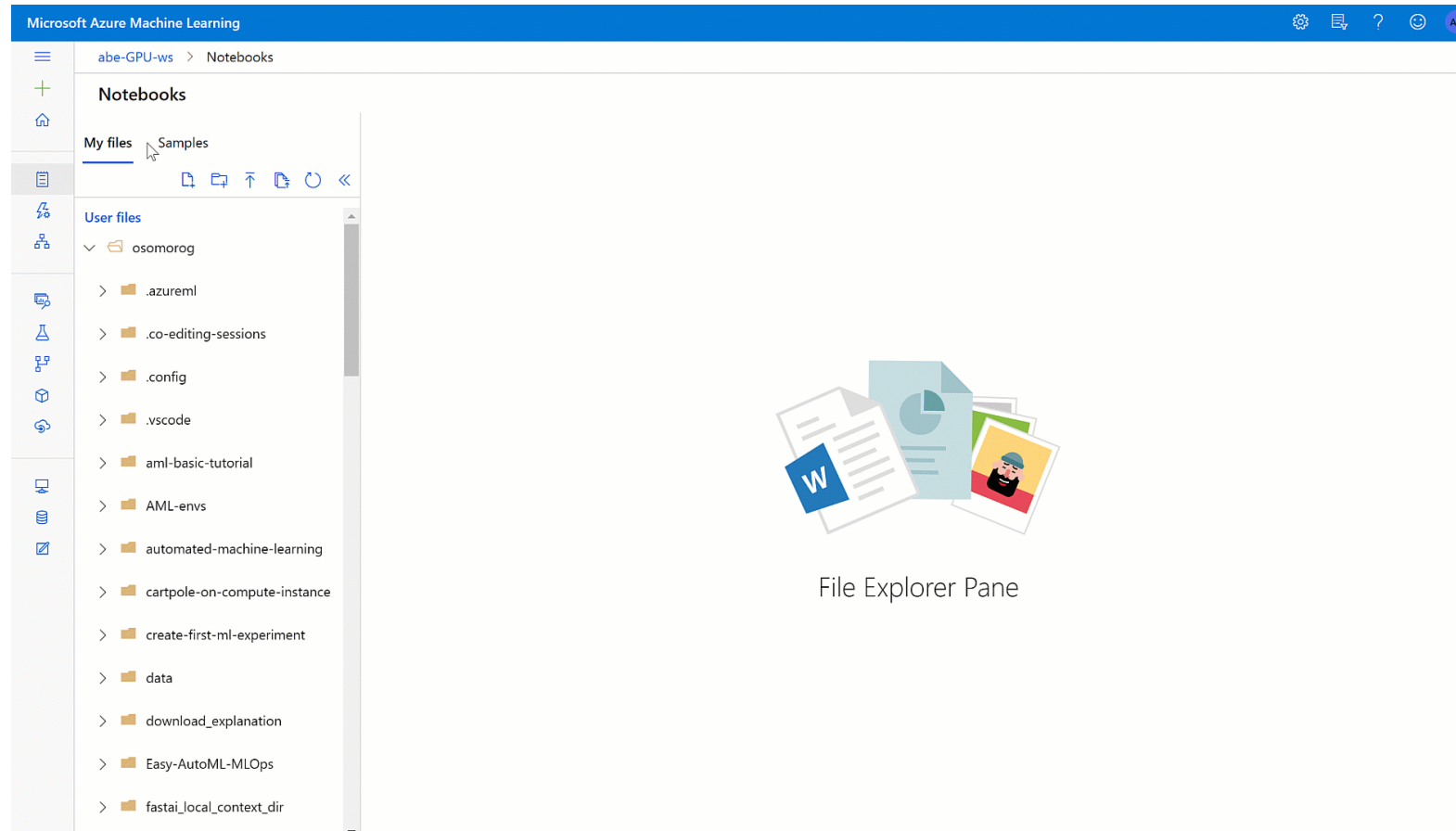


Azure Machine Learning studio is the web portal for data scientist developers in Azure Machine Learning. The studio combines no-code and code-first experiences for an inclusive data science platform.



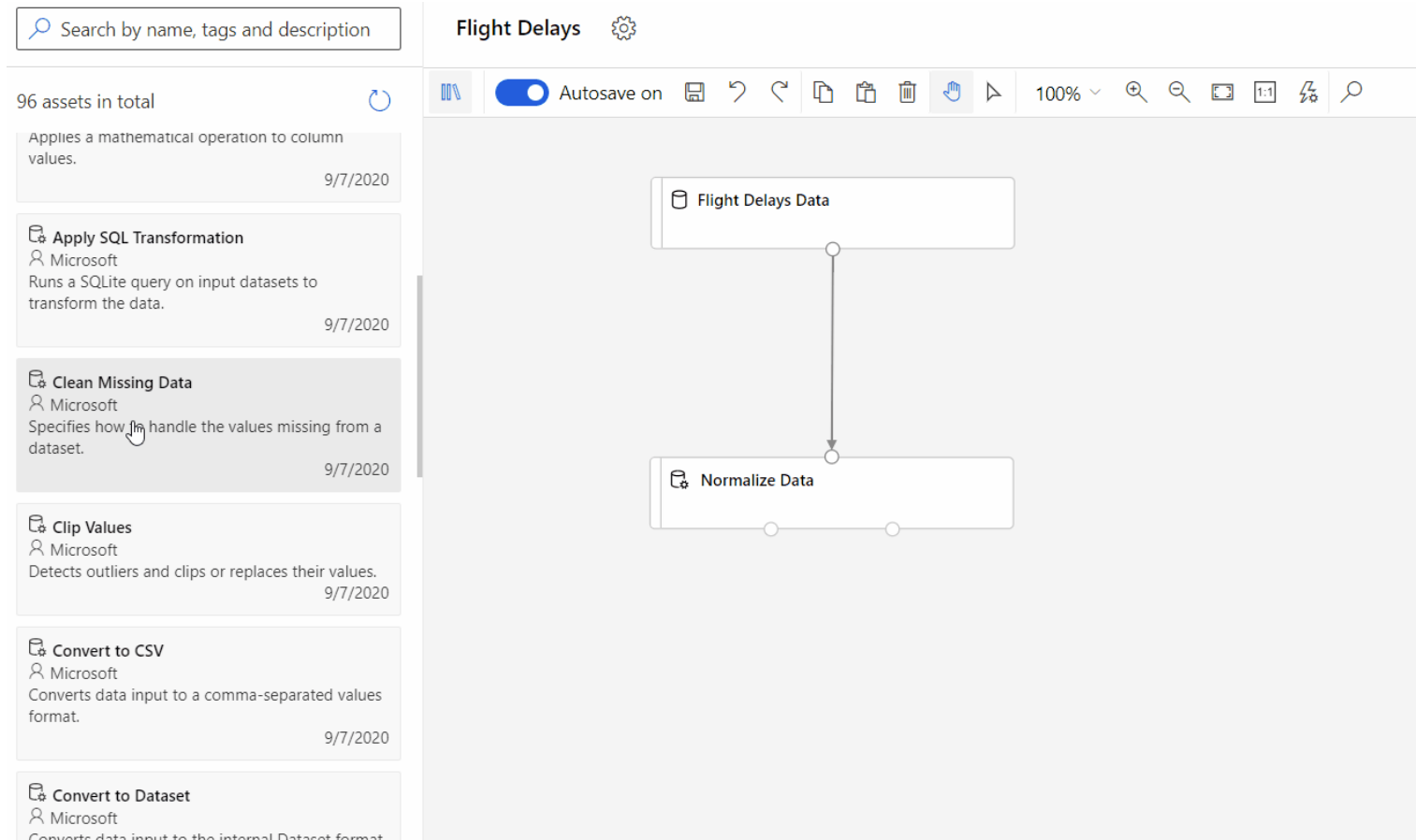
Azure Machine Learning Projects

Notebooks - Write and run your own code in managed Jupyter Notebook servers that are directly integrated in the studio.



Azure Machine Learning Projects

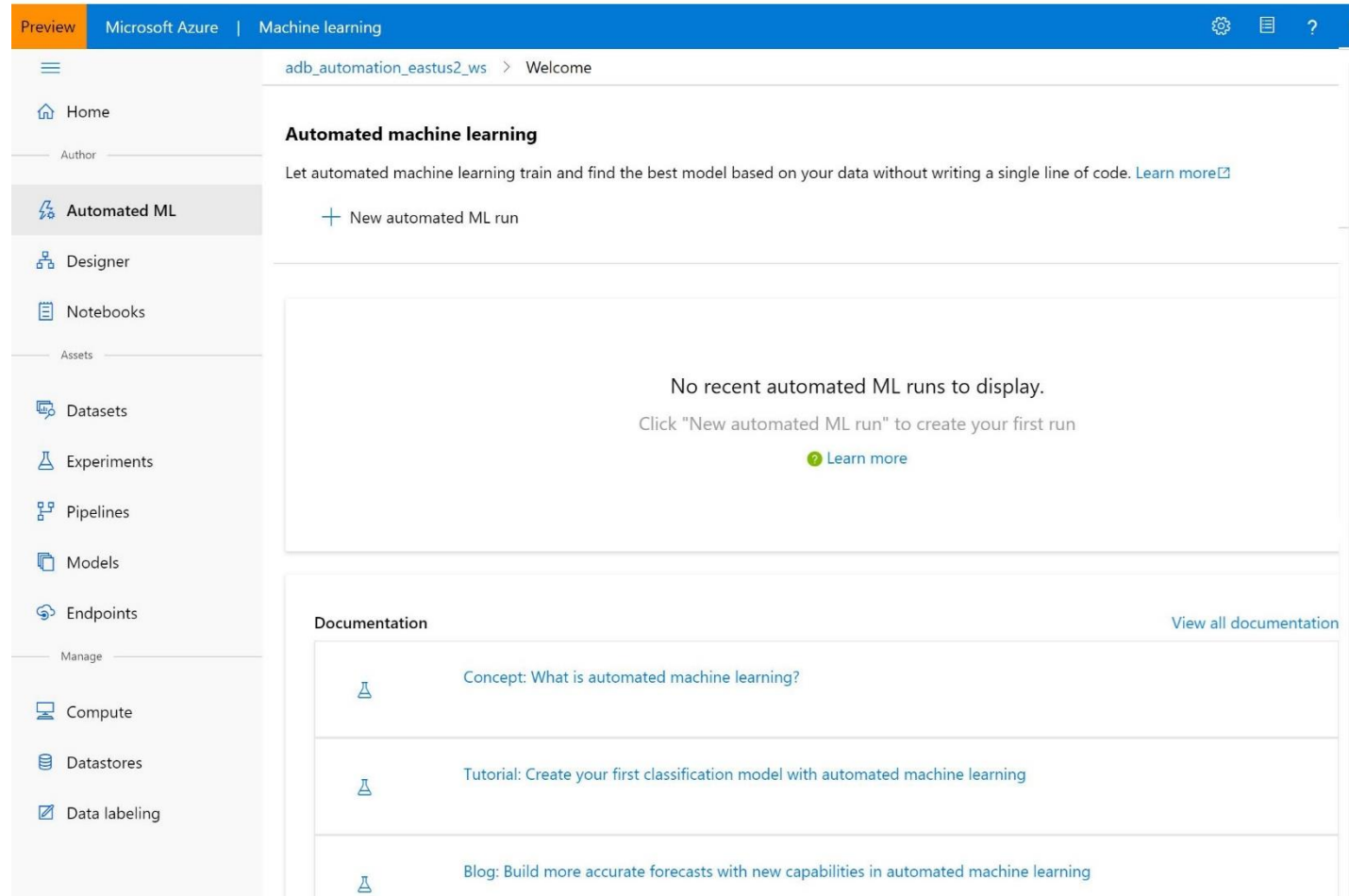
Azure ML Designer - Use the designer to train and deploy machine learning models without writing any code. Drag and drop datasets and modules to create ML pipelines.



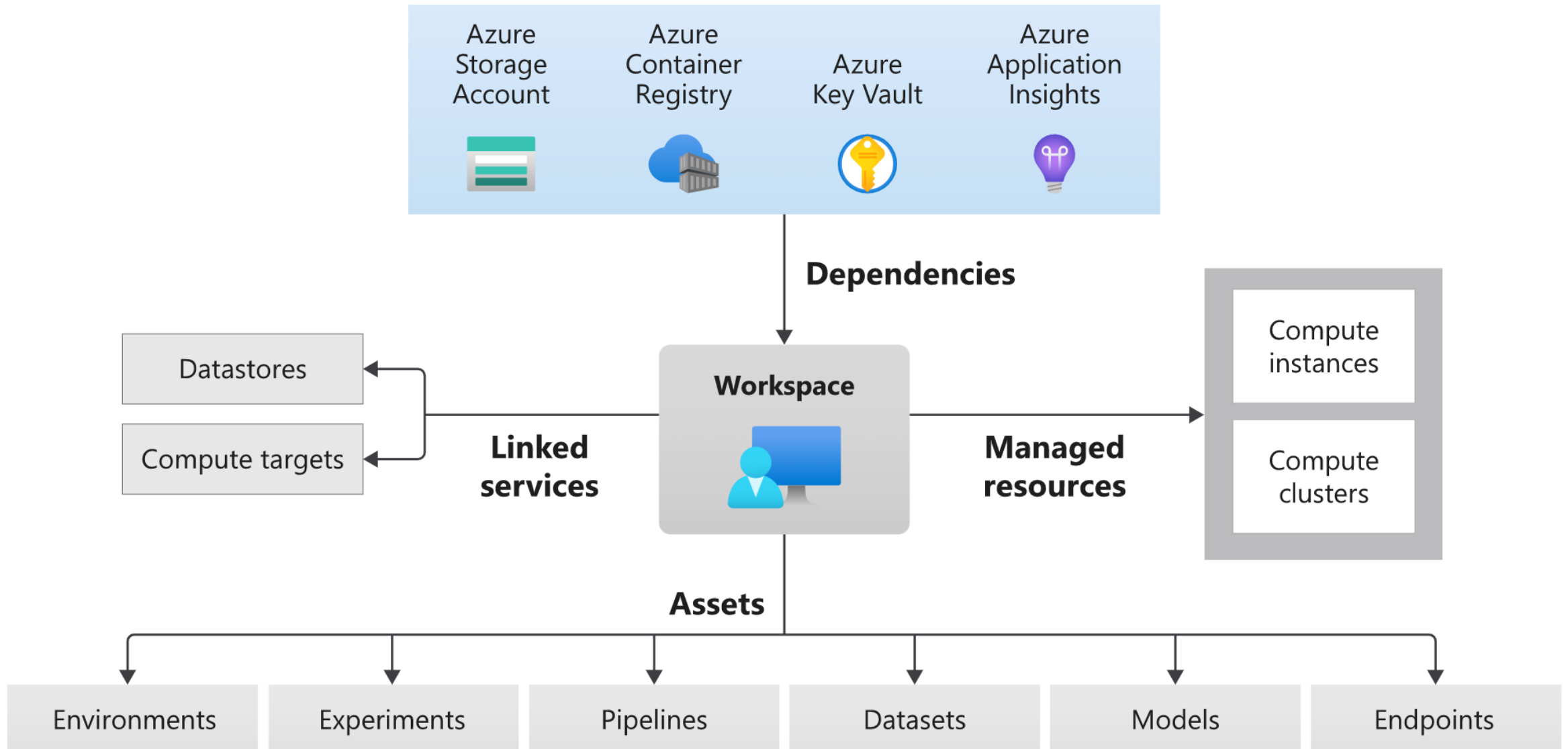
Azure Machine Learning Projects



Automated ML UI - Learn how to create automated ML experiments with an easy-to-use interface.

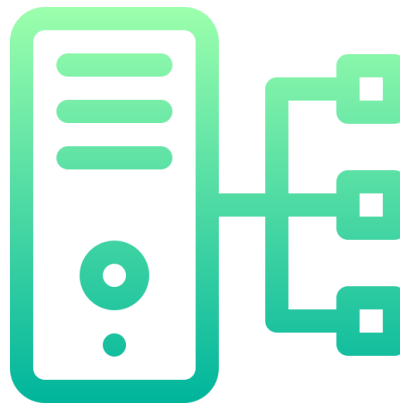


Azure Machine Learning Architecture



Computes - A compute target is any machine or set of machines you use to run your training script or host your service deployment. You can use your local machine or a remote compute resource as a compute target

- Compute instance: A compute instance is a VM that includes multiple tools and environments installed for machine learning. The primary use of a compute instance is for your development workstation.
- Compute clusters: Compute clusters are a cluster of VMs with multi-node scaling capabilities. Compute clusters are better suited for compute targets for large jobs and production.





Azure ML Designer Demo

Regression - Automobile Price Prediction

We will be creating a no-code regression model using Azure ML Designer. Below are the things we will learn from this demo

- Create a new pipeline.
- Import data.
- Prepare data.
- Train a machine learning model.
- Evaluate a machine learning model.

Azure MLOps Project



India: +91-7022374614

US: 1-800-216-8930 (TOLL FREE)



support@intellipaat.com



24/7 Chat with Our Course Advisor