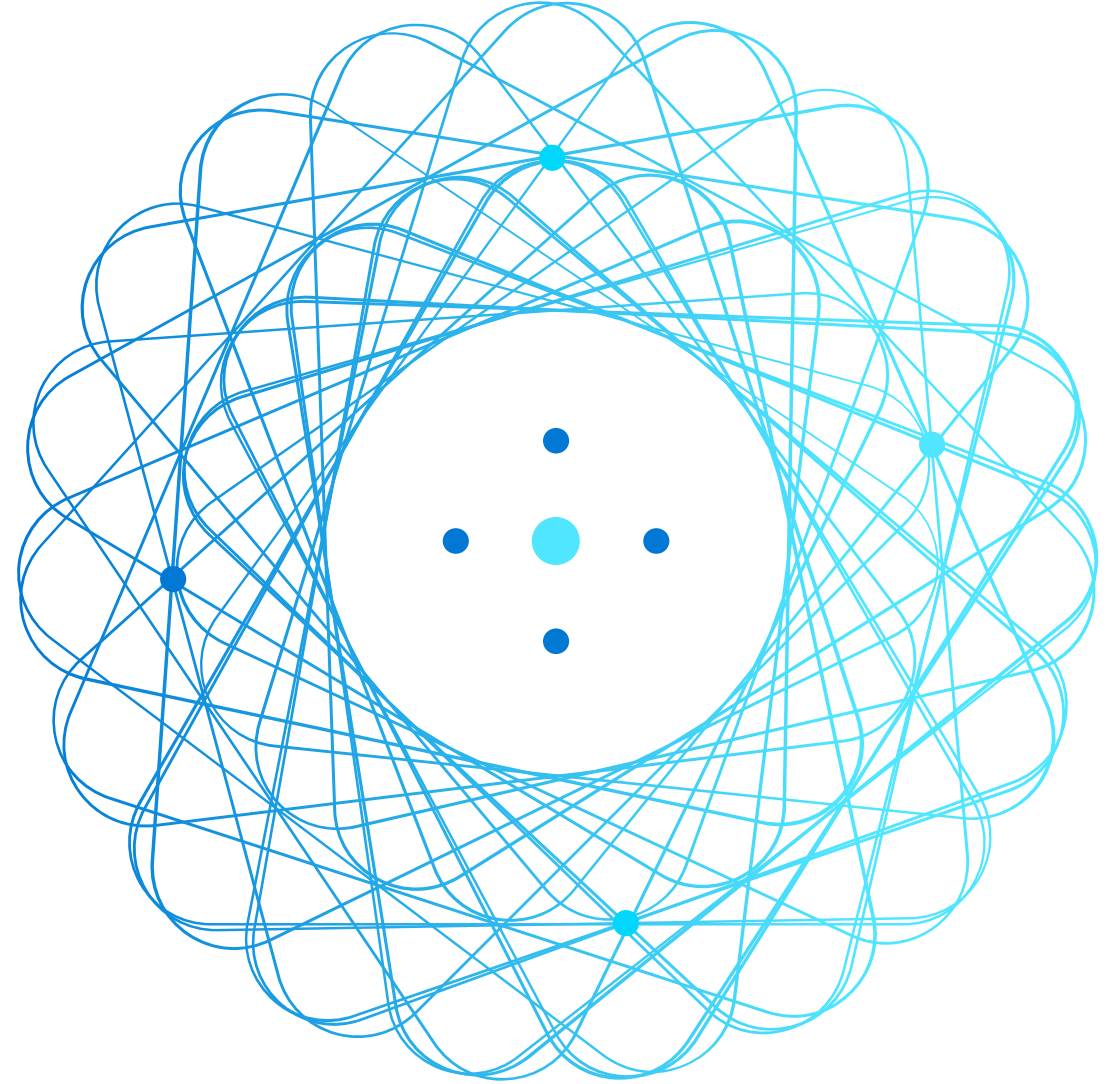


# Module 2: Visual Tools for Machine Learning

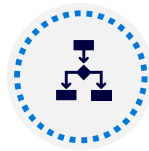
Mohammed Arif



# Agenda

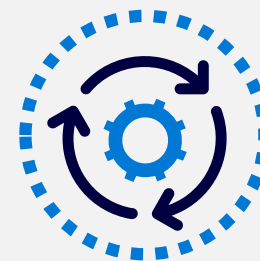


Automated Machine Learning



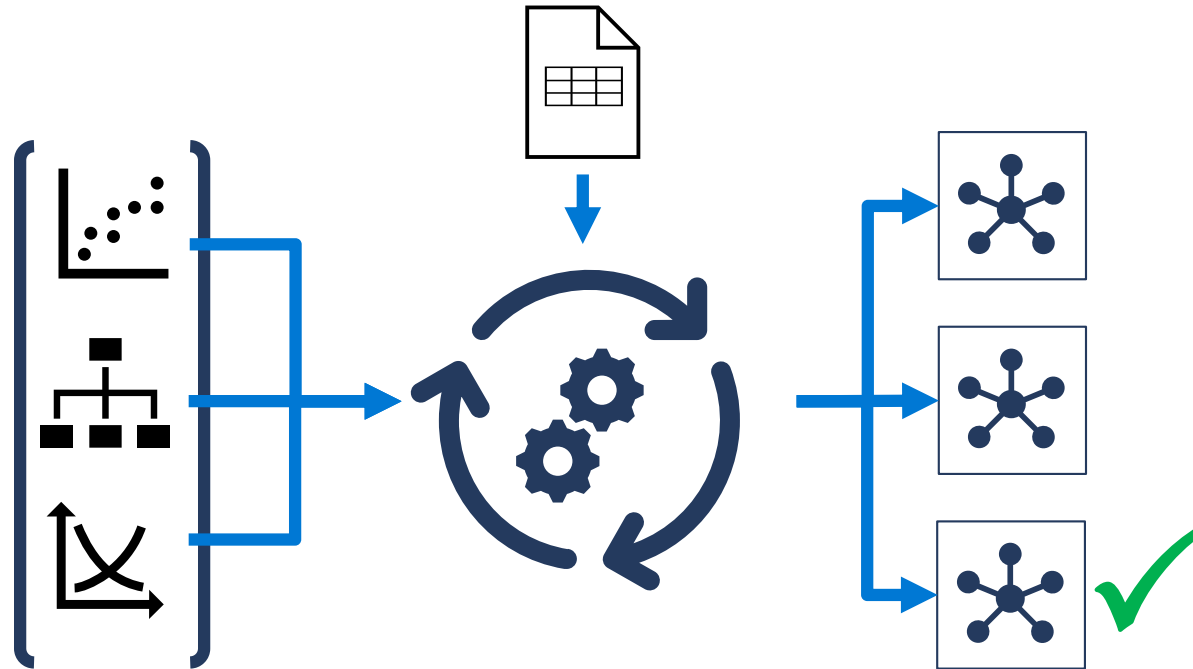
Azure Machine Learning Designer

# Automated Machine Learning



# What is Automated Machine Learning?

Train multiple models in parallel, varying algorithm and preprocessing  
Find the "best" model based on a specific performance metric



# Automated ML in Azure Machine Learning Studio

## 1. Select dataset

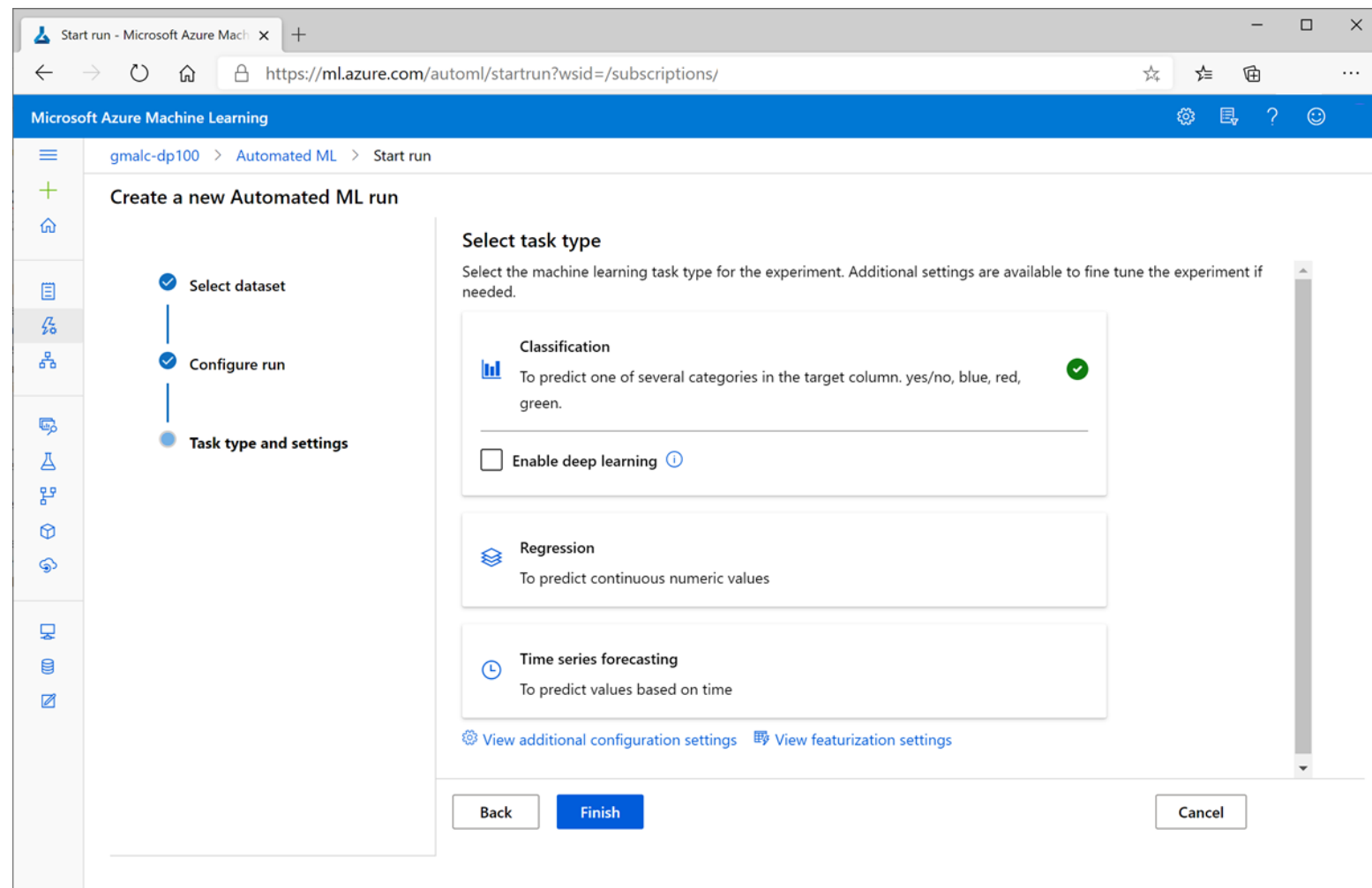
- Upload files
- Import from Web
- Register data source

## 2. Configure run

- Experiment name
- Target label
- Compute

## 3. Task type and settings

- Classification
- Regression
- Time Series



# Configuration and Featurization

## Configuration Options

- **Primary metric** (used to evaluate the best model)
- **Explain best model** (generates feature importance)
- **Blocked algorithms** (restricts training algorithms)
- **Exit criterion** (enables early-stopping)
- **Validation** (sets cross-validation technique)
- **Concurrency** (sets number of parallel iterations)

## Featurization

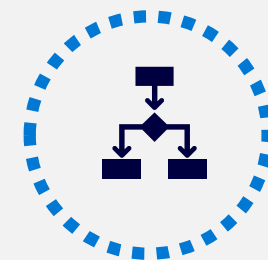
- **Normalization / scaling** is automatic
- **Optional featurization includes:**
  - Dropping high-cardinality features
  - Imputing missing values
  - Categorical encoding
  - Derived feature generation
- **Data guardrails** mitigate unbalanced data

# Lab: Use Automated Machine Learning



1. View the lab instructions at <https://aka.ms/mslearn-dp100>
2. Complete the **Use Automated Machine Learning** exercise

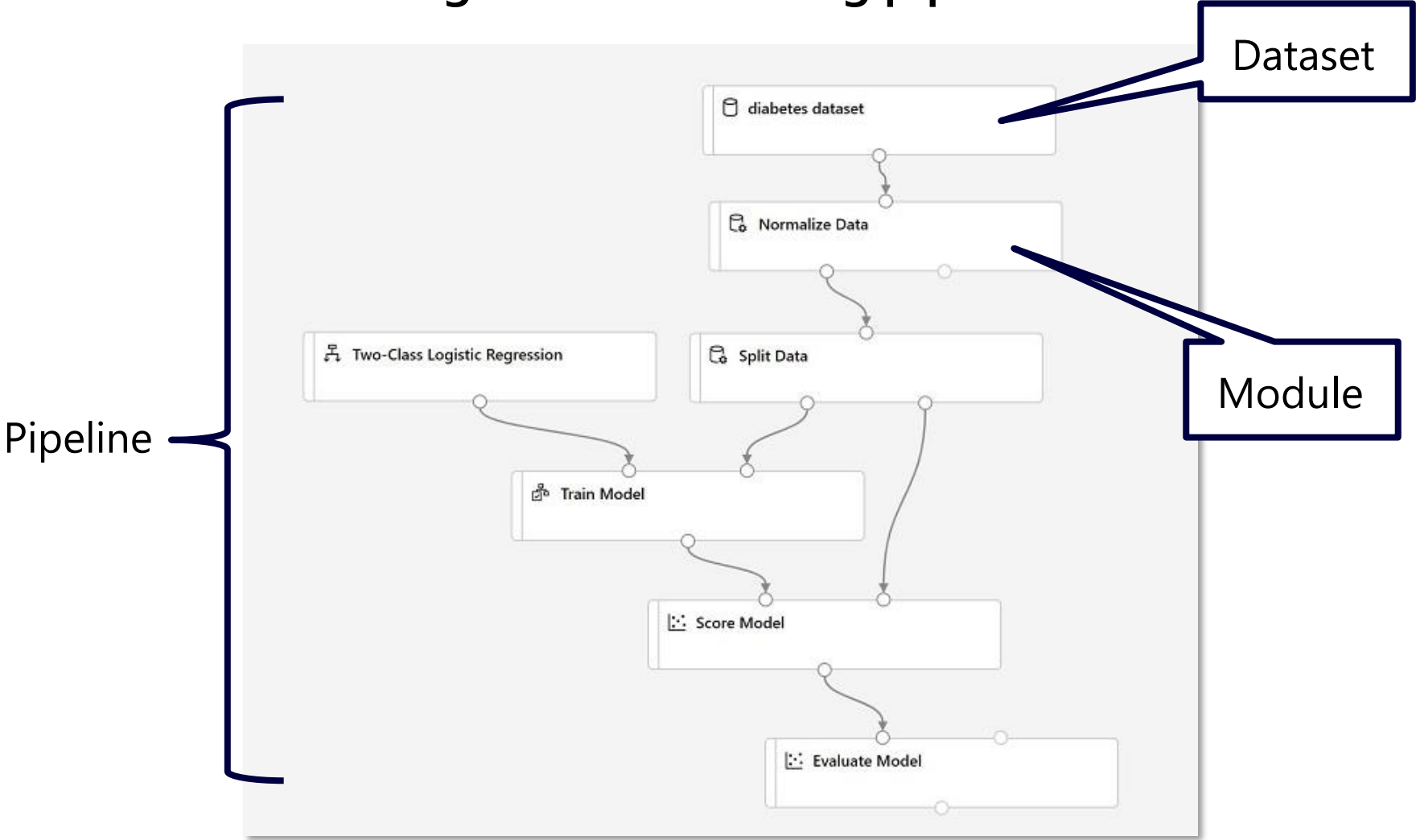
# Azure Machine Learning Designer





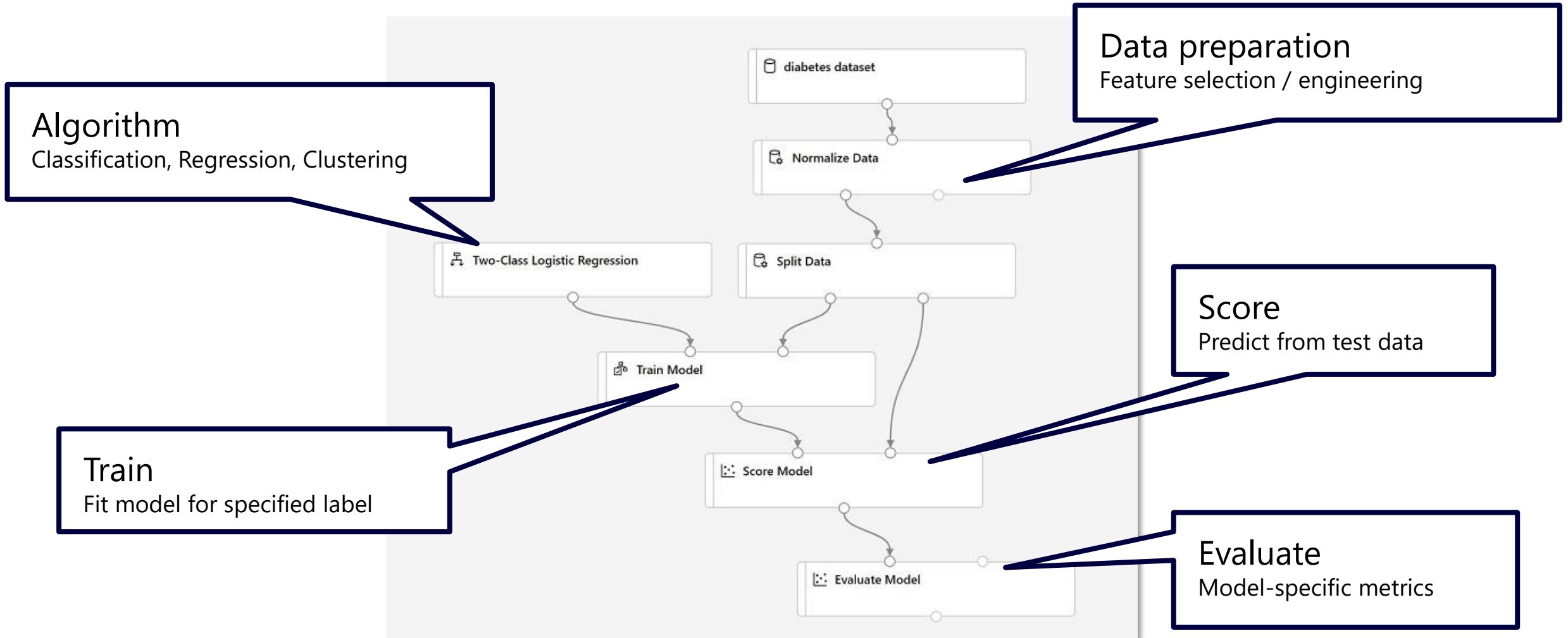
# What is Azure Machine Learning Designer?

A visual interface for creating machine learning pipelines



# Training Pipelines

Data preparation, model training, scoring, and evaluation



# Inference Pipelines

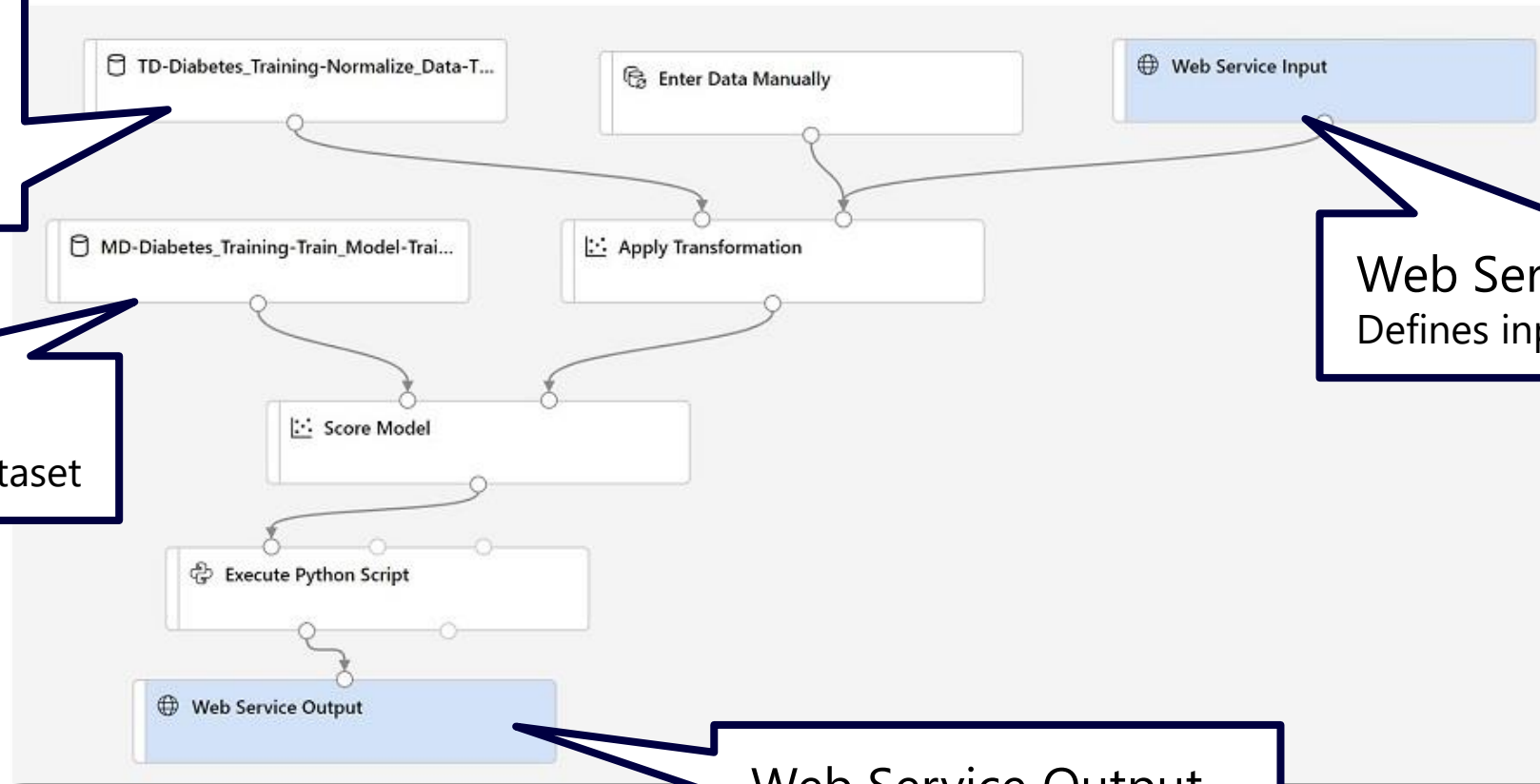
Use the trained model to get predictions from new data

Transformations  
Encapsulated in  
datasets based on  
training data statistics

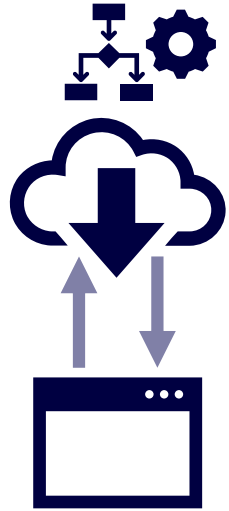
Trained Model  
Encapsulated in a Dataset

Web Service Input  
Defines input schema

Web Service Output  
Defines output schema



# Publishing a Service Endpoint

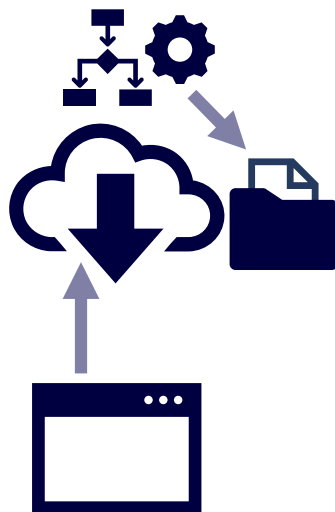


## Deploy a Real-Time Pipeline:

Specify deployment target:

- Azure Container Instance
- Azure Kubernetes Services Inference Compute

Submit new data to an HTTP endpoint for immediate results



## Publish a Batch Pipeline

Runs on Azure Machine Learning Training Compute

Initiate a pipeline experiment run through an HTTP endpoint

Results are saved in the run output

# Lab: Use Azure Machine Learning Designer



1. View the lab instructions at <https://aka.ms/mslearn-dp100>
2. Complete the **Use Azure Machine Learning Designer** exercise

# Knowledge check



You want to use automated machine learning with car sales data to train a machine learning model that predicts the price of a car based on its make, model, engine size, and mileage.

What task type should you select?

- Classification
  - Regression
  - Time-series
- 



You are creating a training pipeline using a dataset that has multiple numeric columns. You want to transform the numeric columns so that the values are all on a similar scale.

Which module should you add to the pipeline?

- Select Columns in a Dataset
- Clean Missing Data
- Normalize Data

# References

Microsoft Learn: Use visual tools to create machine learning models with Azure Machine Learning

<https://docs.microsoft.com/learn/paths/create-no-code-predictive-models-azure-machine-learning>

Automated Machine Learning documentation

<https://docs.microsoft.com/azure/machine-learning/concept-automated-ml>

Designer documentation

<https://docs.microsoft.com/azure/machine-learning/concept-designer>

