#GlobalAzure

#CloudGenVerona

@cloudgen_verona

Thanks to all the sponsors

PREMIUM SPONSOR

4WARD

KIRATECH®
WE DEVOPS IT

BASIC SPONSOR

CODICEPLASTICO

dgroove
Tecnologia emozionale.

TOPIC

Azure Machine Learning Service

Automated Machine Learning

Marco Zamana

@marco_zamana

MarcoZama

marco-zamana

Microsoft Most Valuable Professional

Azure Machine Learning service

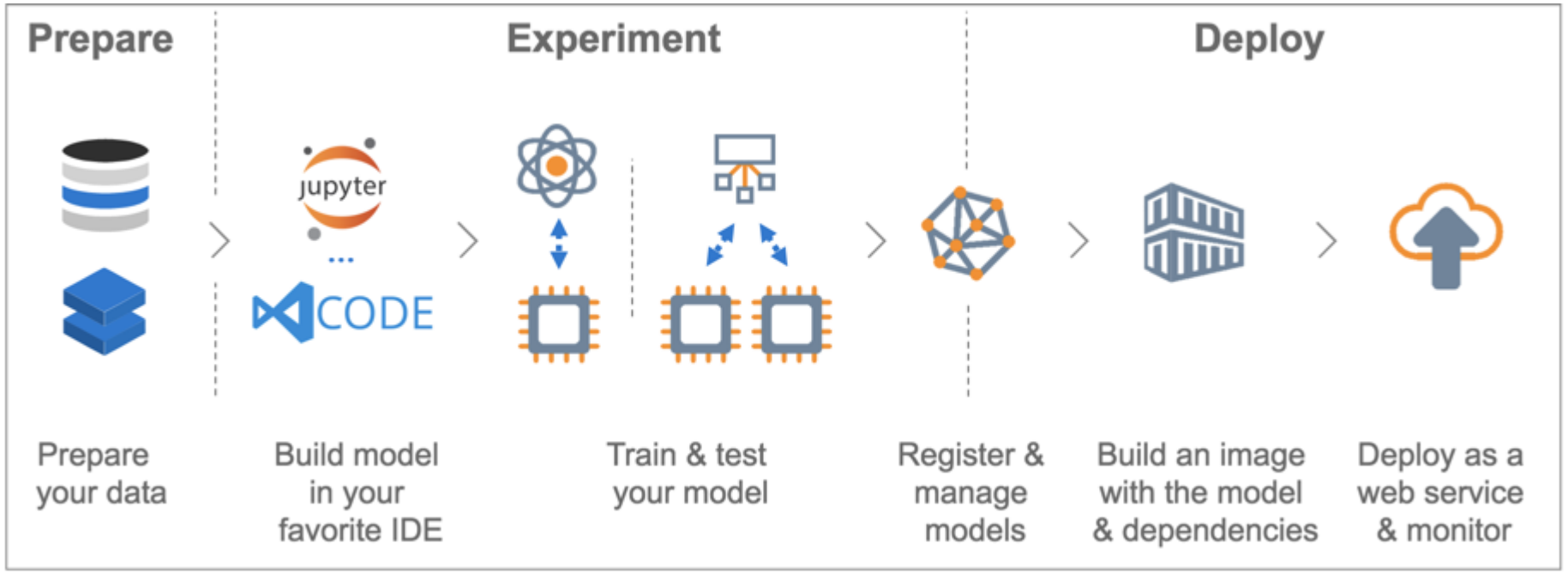
What is Machine Learning?

Machine learning is a data science technique that allows computers to use existing data to forecast future behaviors, outcomes, and trends. By using machine learning, computers learn without being explicitly programmed.

What is Azure Machine Learning service?

Azure Machine Learning service is a cloud service that you use to train, deploy, automate, and manage machine learning models, all at the broad scale that the cloud provides

What is Azure Machine Learning service?



What is Azure Machine Learning service?

Support Open Source Technologies

Azure Machine Learning service fully supports open-source technologies.

We can use tens of thousands of open-source Python packages with machine learning components.

What is Azure Machine Learning service?

From Local to Cloud

You can start training on your local machine and then scale out to the cloud.

With many available compute targets, like Azure Machine Learning Compute and Azure Databricks, and with advanced hyperparameter tuning services, you can build better models faster by using the power of the cloud.

What is Azure Machine Learning service?

Container deployment

When you have the right model, you can easily deploy it in a container such as Docker.

So it's simple to deploy to Azure Container Instances or Azure Kubernetes Service. Or you can use the container in your own deployments, either on-premises or in the cloud.

What is Azure Machine Learning service?

Manage the model

You can manage the deployed models and track multiple runs as you experiment to find the best solution.

After it's deployed, your model can return predictions in real time or asynchronously on large quantities of data.

What is Azure Machine Learning service?

Advanced pipeline

And with advanced machine learning pipelines, you can collaborate on all the steps of data preparation, model training and evaluation, and deployment.

What can I do with AMLS?

Build and train models

Using the main Python SDK and the Data Prep SDK for Azure Machine Learning as well as open-source Python packages, you can build and train highly accurate machine learning and deep-learning models yourself in an Azure Machine Learning service Workspace.

What can I do with AMLS?

Python package available

- Scikit-learn
- Tensorflow
- PyTorch
- MXNet

What can I do with AMLS?

But…

Azure Machine Learning service can also autotrain a model and autotune it for you.

With Automated ML (we'll see after! ;))

What can I do with AMLS?

Deploy the model

After you have a model, you use it to create a container, such as Docker, that can be deployed locally for testing. After testing is done, you can deploy the model as a production web service in either Azure Container Instances or Azure Kubernetes Service.

What can I do with AMLS?

Manage the models!

Then you can manage your deployed models by using the Azure Machine Learning SDK for Python or the Azure portal. You can evaluate model metrics, retrain, and redeploy new versions of the model, all while tracking the model's experiments.

AMLService vs AMLStudio

Azure Machine Learning Studio

Azure Machine Learning Studio is a collaborative, drag-and-drop visual workspace where you can build, test, and deploy machine learning solutions without needing to write code.

It uses prebuilt and preconfigured machine learning algorithms and data-handling modules.

AMLService vs AMLStudio

Use Azure Machine Learning studio: when you want to experiment with machine learning models quickly and easily, and the built-in machine learning algorithms are sufficient for your solutions.

AMLService vs AMLStudio

Use Machine Learning service if you work in a Python environment, you want more control over your machine learning algorithms, or you want to use open-source machine learning libraries.

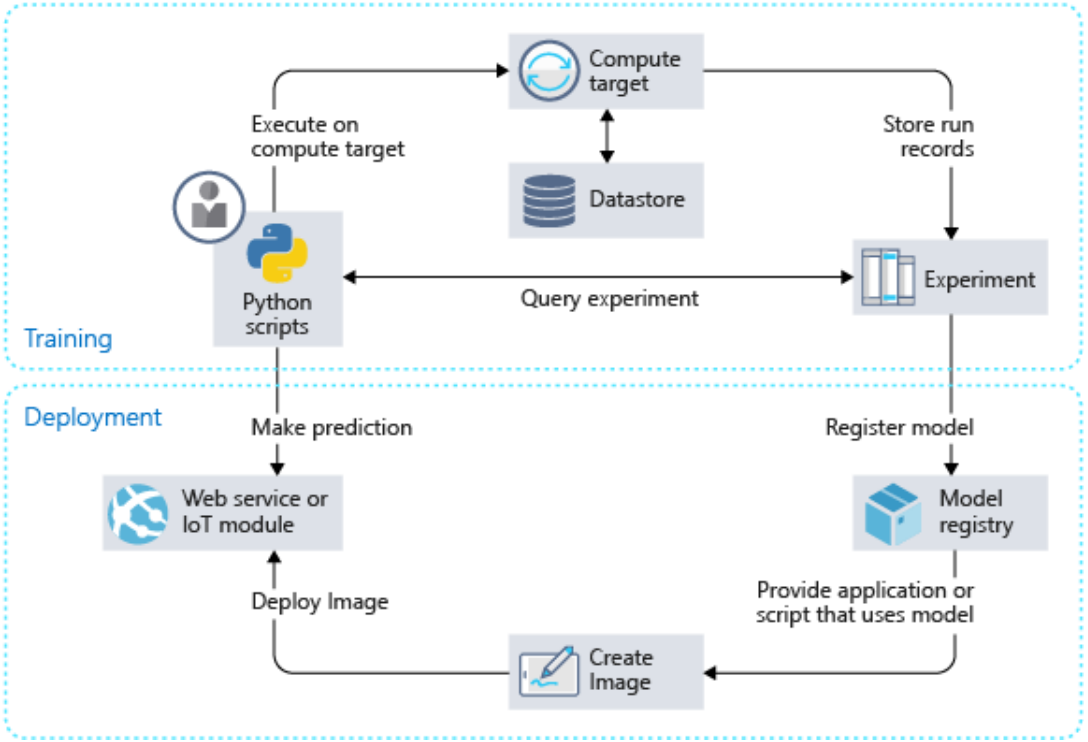
AMLService vs AMLStudio

Attention!!!

Models created in Azure Machine Learning Studio can't be deployed or managed by Azure Machine Learning service.

How AMLS works

Architecture and Workflow



1. Develop machine learning training scripts in **Python**.
2. Create and configure a **compute target**.
3. **Submit the scripts** to the configured compute target to run in that environment. During training, the scripts can read from or write to **datastore**. And the records of execution are saved as **runs** in the **workspace** and grouped under **experiments**.
4. **Query the experiment** for logged metrics from the current and past runs. If the metrics don't indicate a desired outcome, loop back to step 1 and iterate on your scripts.
5. After a satisfactory run is found, register the persisted model in the **model registry**.
6. Develop a scoring script.
7. **Create an image** and register it in the **image registry**.
8. **Deploy the image** as a **web service** in Azure.

How AMLS works

Fundamental concept and components

1. Workspace
2. Experiment
3. Model
4. Run configuration
5. Datastore
6. Compute target
7. Training script
8. Run
9. Snapshot
10. Activity
11. Image
12. Deployment
13. Pipeline
14. Logging

How AMLS works

Workspace

The workspace is the top-level resource for Azure Machine Learning service. It provides a centralized place to work with all the artifacts you create when you use Azure Machine Learning service.

The workspace keeps a list of compute targets that you can use to train your model. It also keeps a history of the training runs, including logs, metrics, output, and a snapshot of your scripts. You use this information to determine which training run produces the best model.

Experiment

An experiment is a grouping of many runs from a specified script. It always belongs to a workspace.

When you submit a run, you provide an experiment name. Information for the run is stored under that experiment.

If you submit a run and specify an experiment name that doesn't exist, a new experiment with that newly specified name is automatically created.

Model

A model is a piece of code that takes an input and produces output.

Creating a machine learning model involves selecting an algorithm, providing it with data, and tuning hyperparameters.

Training is an iterative process that produces a trained model, which encapsulates what the model learned during the training process.

Model

Azure Machine Learning service is framework agnostic

You can use any popular machine learning framework, such as Scikit-learn, XGBoost, PyTorch, TensorFlow, and Chainer.

Run configuration

A run configuration is a set of instructions that defines how a script should be run in a specified compute target.

The configuration includes a wide set of behavior definitions, such as whether to use an existing Python environment or to use a Conda environment that's built from a specification.

Datastore

A datastore is a storage abstraction over an Azure storage account.

The datastore can use either an Azure blob container or an Azure file share as the back-end storage.

Each workspace has a default datastore, and you can register additional datastores.

How AMLS works

Compute target

A compute target is the compute resource that you use to run your training script or host your service deployment.

Supported Target are =>

Compute target	Training	Deployment
Your local computer	✓	
Azure Machine Learning compute	✓	
A Linux VM in Azure (such as the Data Science Virtual Machine)	✓	
Azure Databricks	✓	
Azure Data Lake Analytics	✓	
Apache Spark for HDInsight	✓	
Azure Container Instances		✓
Azure Kubernetes Service		✓
Azure IoT Edge		✓
Project Brainwave (Field-programmable gate array)		✓

Training script

To train a model, you specify the directory that contains the training script and associated files.

You also specify an experiment name, which is used to store information that's gathered during training.

During training, the entire directory is copied to the training environment (compute target), and the script that's specified by the run configuration is started.

Run

A run is a record that contains the following information:

1. Metadata about the run (timestamp, duration, and so on)
2. Metrics that are logged by your script
3. Output files that are autocollected by the experiment or explicitly uploaded by you
4. A snapshot of the directory that contains your scripts, prior to the run

You produce a run when you submit a script to train a model

Snapshot

When you submit a run, Azure Machine Learning compresses the directory that contains the script as a zip file and sends it to the compute target.

The zip file is then extracted, and the script is run there. Azure Machine Learning also stores the zip file as a snapshot as part of the run record.

Activity

An activity represents a long running operation.

Activities can provide notifications through the SDK or the web UI so that you can easily monitor the progress of these operations.

How AMLS works

Image

Images provide a way to reliably deploy a model, along with all components you need to use the model.

An image contains the following items:

1. A model.
2. A scoring script or application. You use the script to pass input to the model and return the output of the model.
3. The dependencies that are needed by the model or scoring script or application. For example, you might include a Conda environment file that lists Python package dependencies.

Deployment

A deployment is an instantiation of your image into either a web service that can be hosted in the cloud or an IoT module for integrated device deployments.

Deployment

Web service

A deployed web service can use Azure Container Instances, Azure Kubernetes Service, or FPGAs. You create the service from an image that encapsulates your model, script, and associated files. The image has a load-balanced, HTTP endpoint that receives scoring requests that are sent to the web service.

Deployment

IoT module

A deployed IoT module is a Docker container that includes your model and associated script or application and any additional dependencies. You deploy these modules by using Azure IoT Edge on edge devices.

Pipeline

You use machine learning pipelines to create and manage workflows that stitch together machine learning phases

Logging

When you develop your solution, use the Azure Machine Learning Python SDK in your Python script to log arbitrary metrics.

After the run, query the metrics to determine whether the run has produced the model you want to deploy.

> ...what we want is a machine that can learn from experience

Alan Turing, 1947

Automated Machine Learning

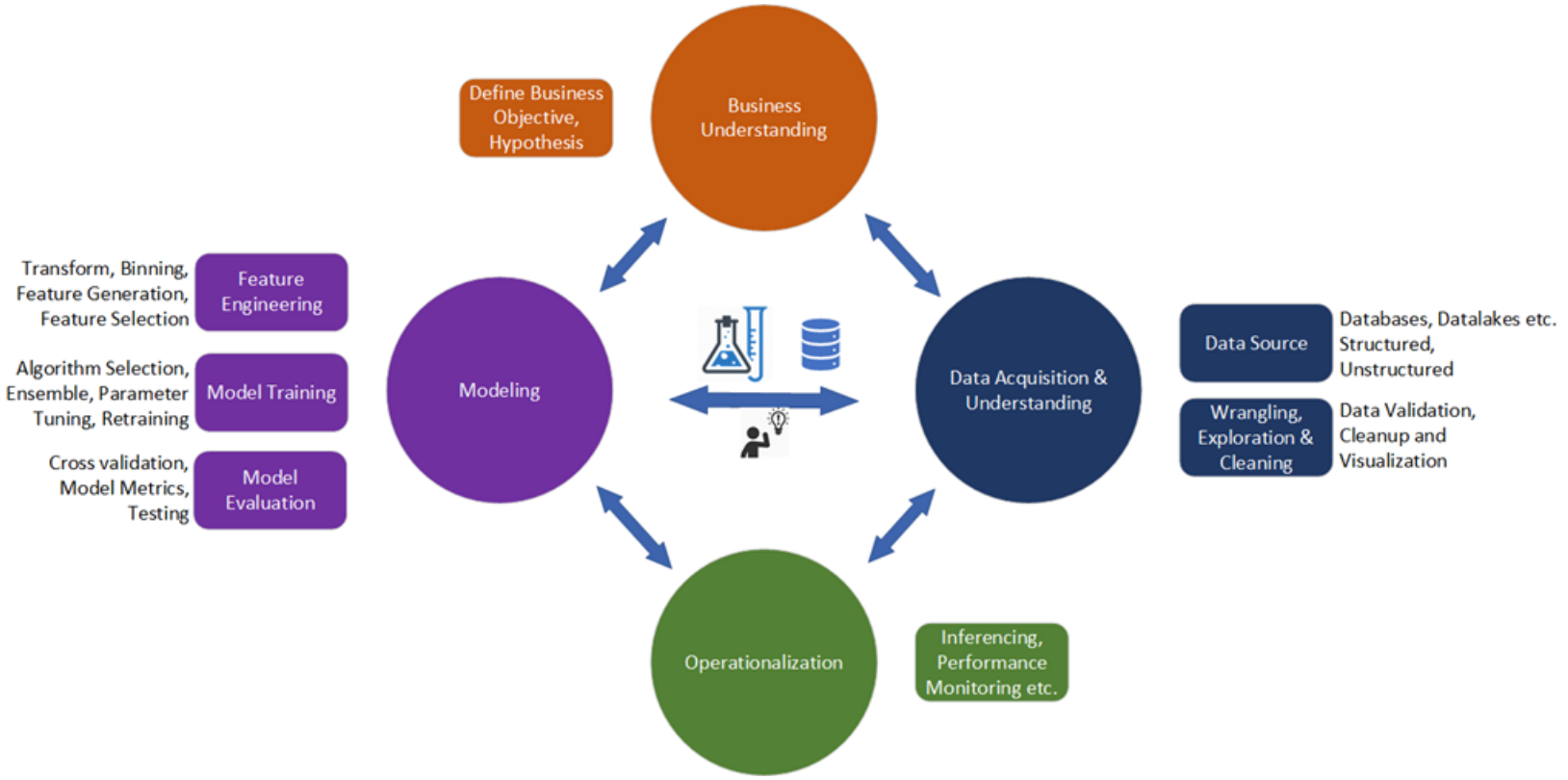
Data scientist and developers

"magic" machine learning solution

Automated Machine Learning

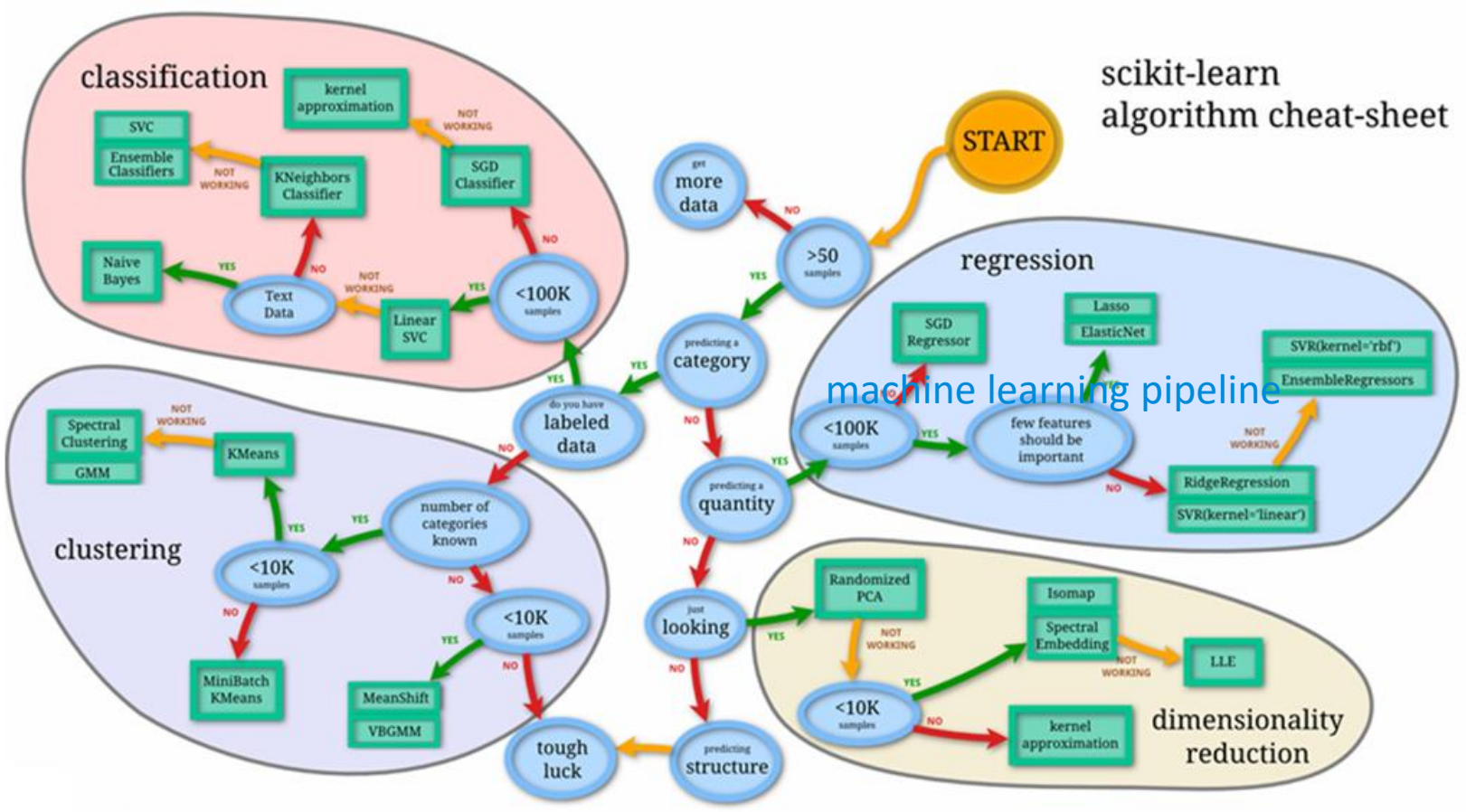
Machine learning lifecycle

1. Business Understanding
2. Data Acquisition
3. Modeling
4. Operationalization



Every Machine Learning solution should start with the business problem you are working to solve followed by acquiring and exploring the data that is needed.

Automated Machine Learning



Decisions

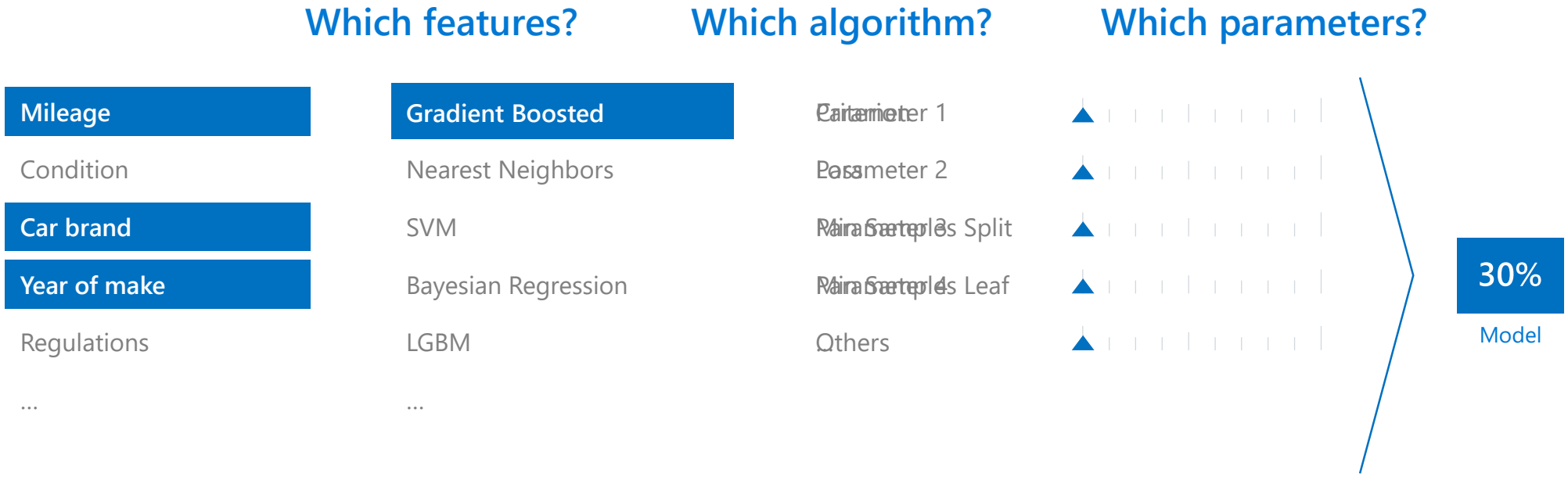
- What ml algorithm would be best?
- What parameter values should they use for the chosen classifier?

And many more…

Source: [scikit-learn machine learning library]

Automated Machine Learning

Model Creation Is Typically Time-Consuming



Automated Machine Learning

Model Creation Is Typically Time-Consuming

Which features?

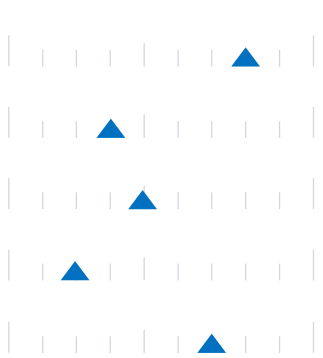
- **Mileage**
- **Condition**
- **Car brand**
- **Year of make**
- Regulations
- ...

Which algorithm?

- **Gradient Boosted**
- **Nearest Neighbors**
- SVM
- Bayesian Regression
- LGBM
- ...

Which parameters?

- Criterion
- Weights
- Min Samples Split
- Min Samples Leaf
- Others



30%

Model

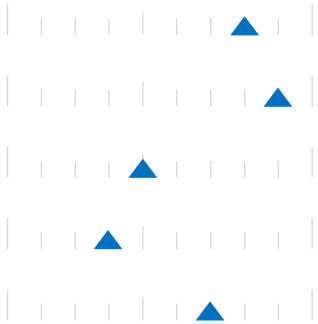
Iterate

Model Creation Is Typically Time-Consuming

Which features?

Which algorithm?

Which parameters?



30%

15%

Iterate

Automated Machine Learning

The combination of data pre-processing steps, learning algorithms, and hyperparameter settings that go into each machine learning solution.

Machine learning pipeline

ACCURACY

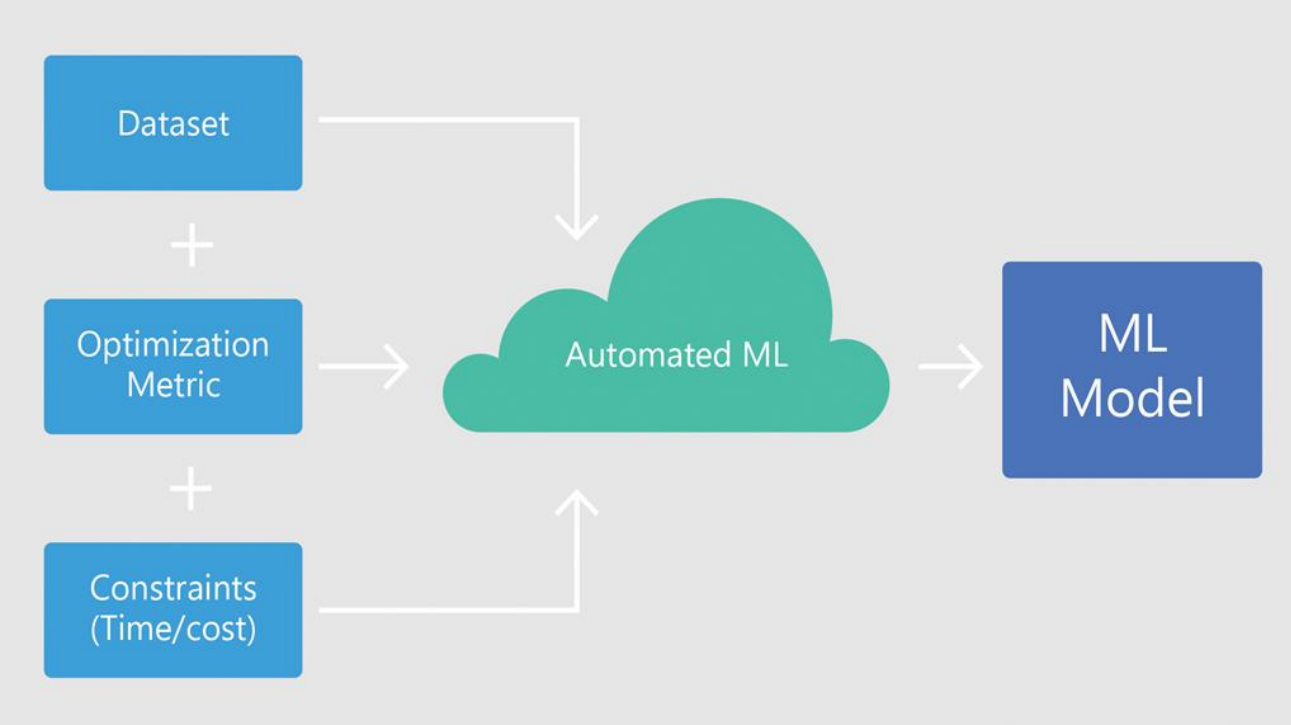
Simplifying machine learning

What if a developer or data scientist could access an automated service that identifies the best machine learning pipelines for their labelled data?

Automated Machine Learning

Automated ML empowers customers, with or without data science expertise, to identify an end-to-end machine learning pipeline for any problem, achieving higher accuracy while spending far less of their time.

And it enables a significantly larger number of experiments to be run, resulting in faster iteration towards production-ready intelligent experiences.



Automated Machine Learning

Automated ML Accelerates Model Development

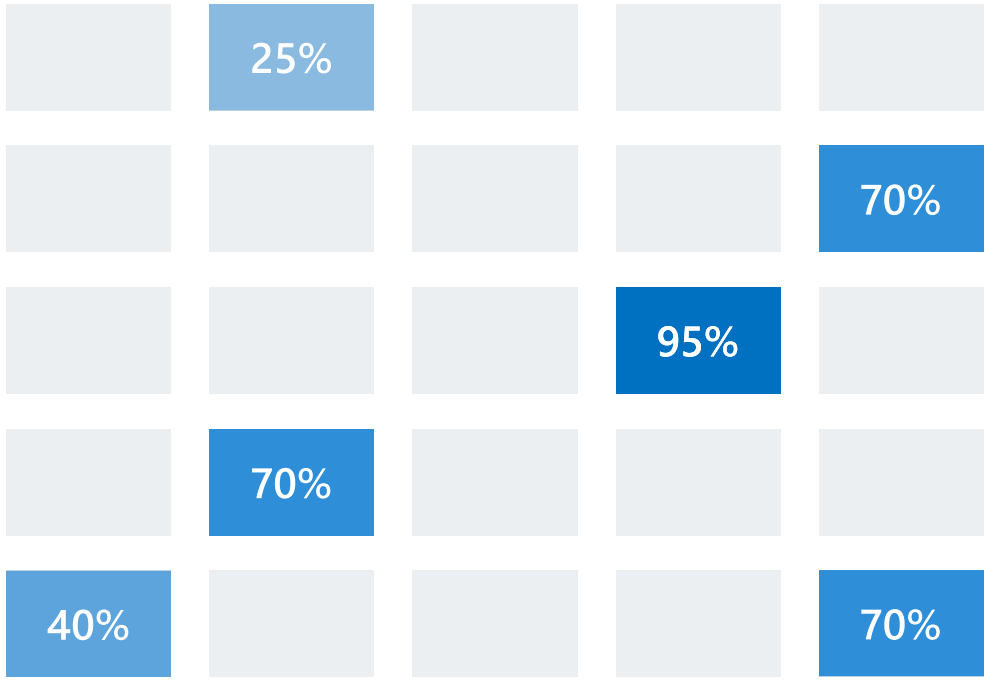
Input

Enter data

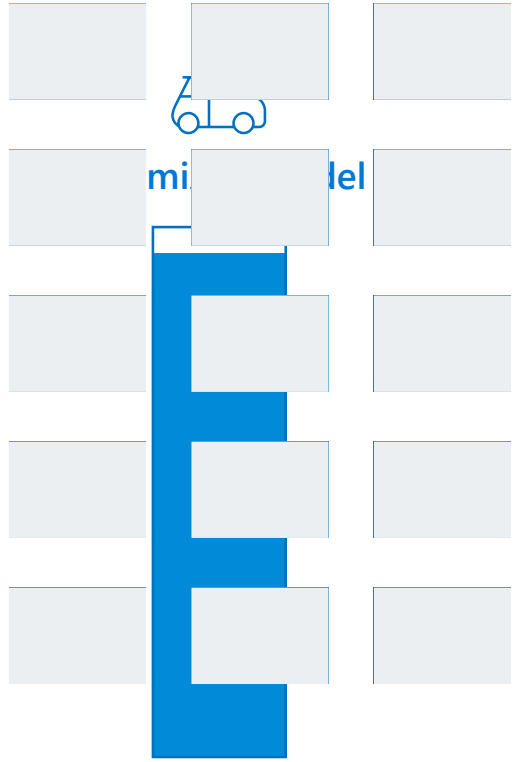
Define goals

Apply constraints

Intelligently test multiple models in parallel

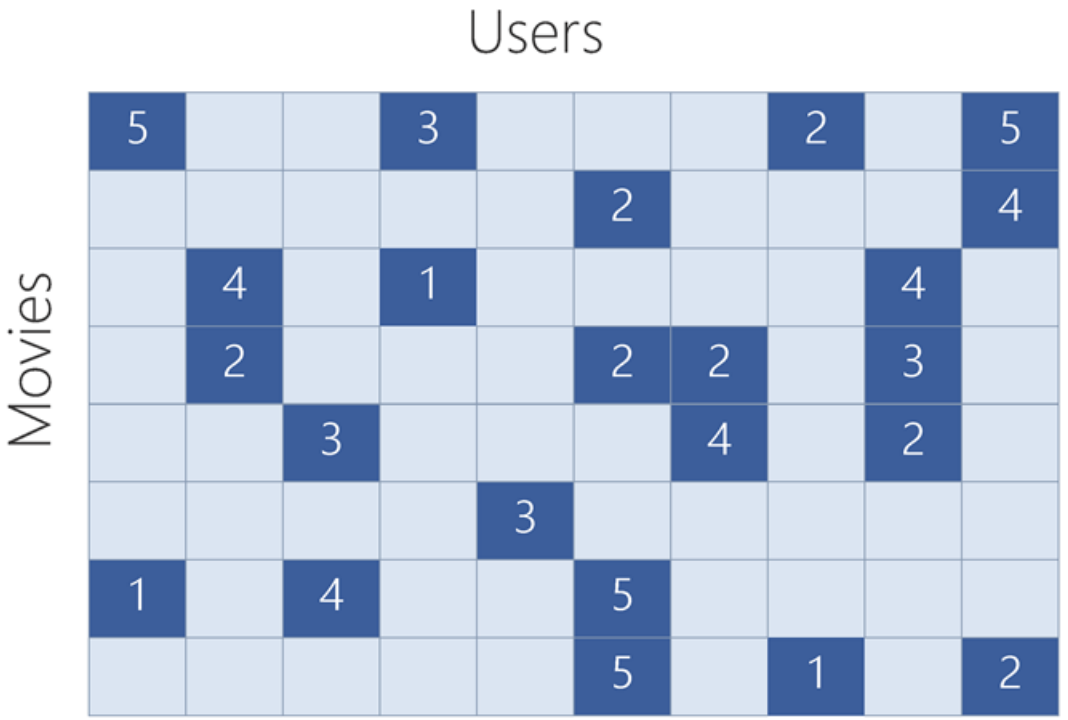


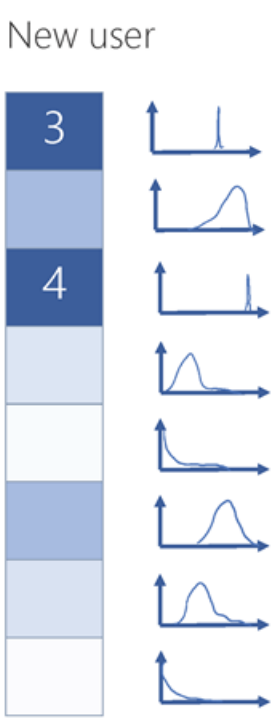
Output



Automated Machine Learning

Automated ML is based on a breakthrough from our Microsoft Research division.

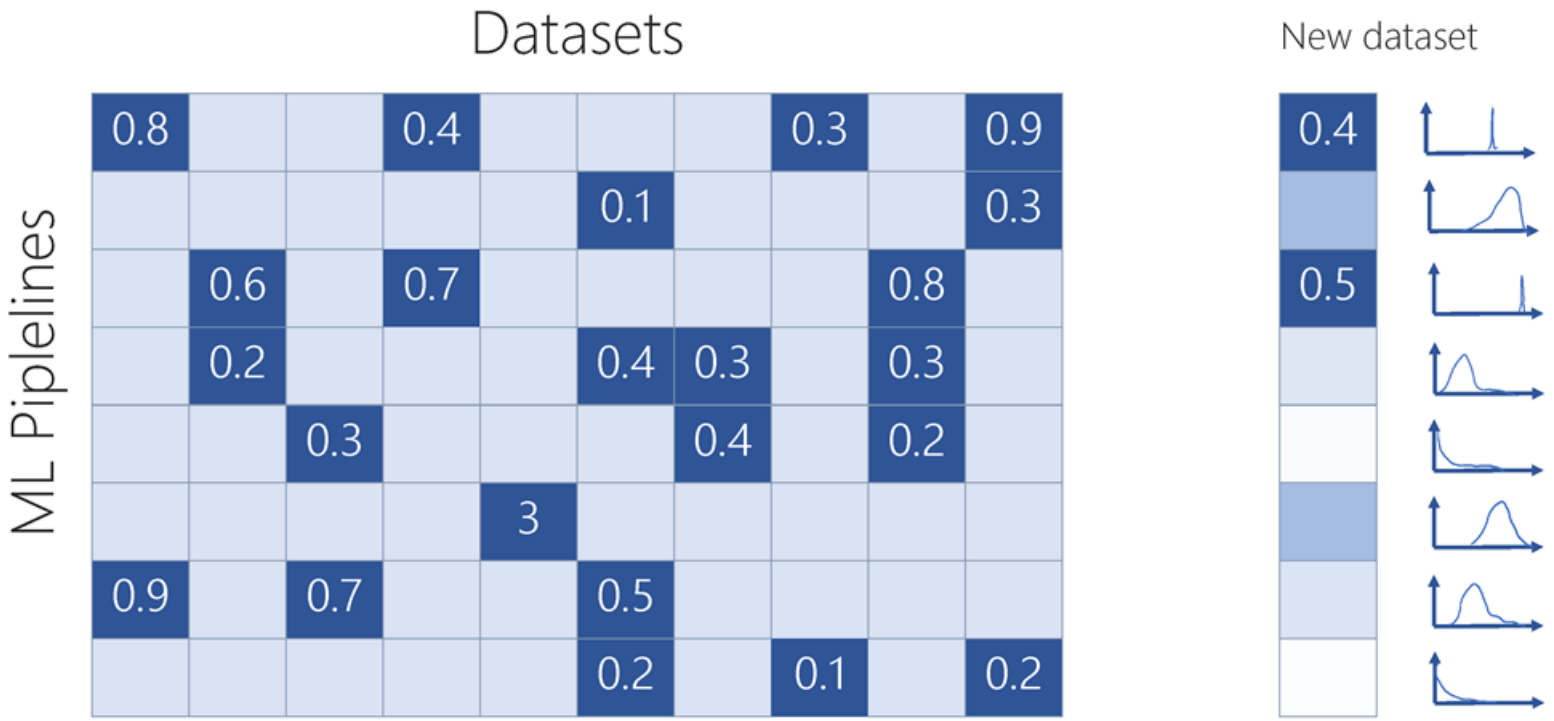




The approach combines ideas from collaborative filtering and Bayesian optimization to search an enormous space of possible machine learning pipelines intelligently and efficiently.

It's essentially a recommender system for machine learning pipelines.
Similar to how streaming services recommend movies for users…

Automated Machine Learning



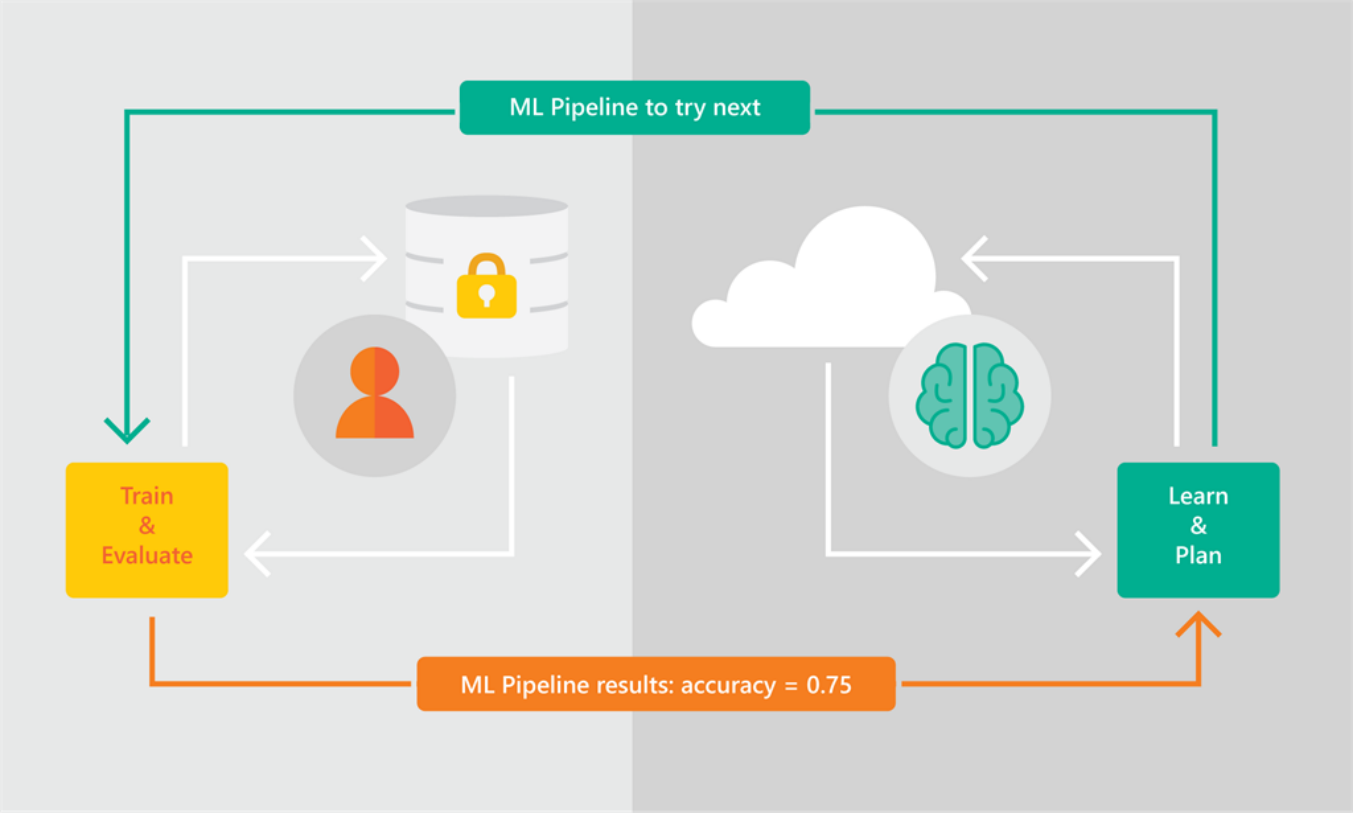
… Automated ML recommends machine learning pipelines for data sets.

No need to “see” the data

Automated ML accomplishes all this without having to see the customer’s data, preserving privacy.
Automated ML is designed to not look at the customer’s data.
Customer data and execution of the machine learning pipeline both live in the customer’s cloud subscription
(or their local machine), which they have complete control of.

Automated Machine Learning

No need to “see” the data



We trained automated ML’s probabilistic model by running hundreds of millions of experiments, each involving evaluation of a pipeline on a data set.

This training now allows the automated ML service to find good solutions quickly for your new problems.

And the model continues to learn and improve as it runs on new ML problems – even though, as mentioned above, it does not see your data.

Thanks

Questions?


MarcoZama


@marco_zamana


marco-zamana