Democratizing & Accelerating AI Through Automated Machine Learning

AI Platform Team

Microsoft

Agenda

- Welcome Video
- Why Automated Machine Learning
- Automated ML Capabilities
- How to Get Started

Welcome Video

Why Automated Machine Learning

Machine Learning Process

SQL DB

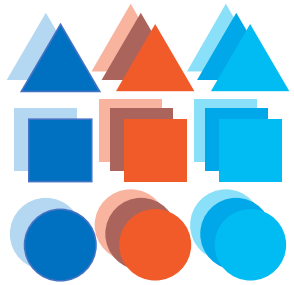
Cosmos DB

Datawarehouse

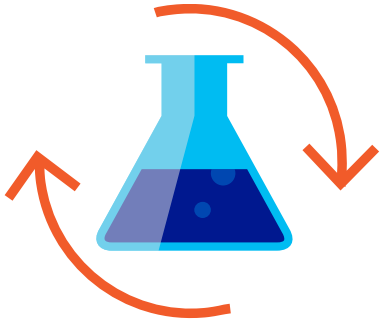
Data lake

Blob storage

…



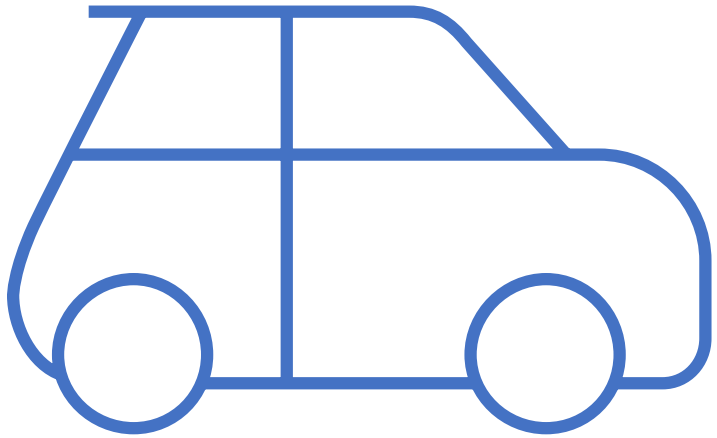
Prepare Data



Build & Train

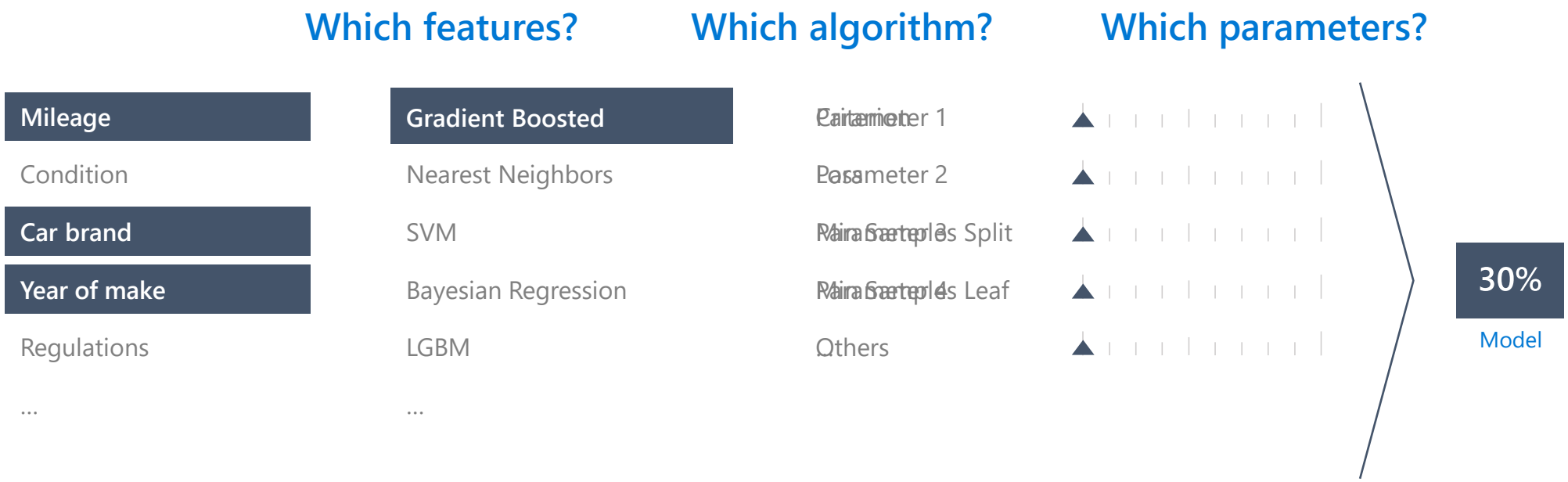
Deploy

Machine Learning Problem Example



How much is this car worth?

Model Creation Is Typically Time-Consuming



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?

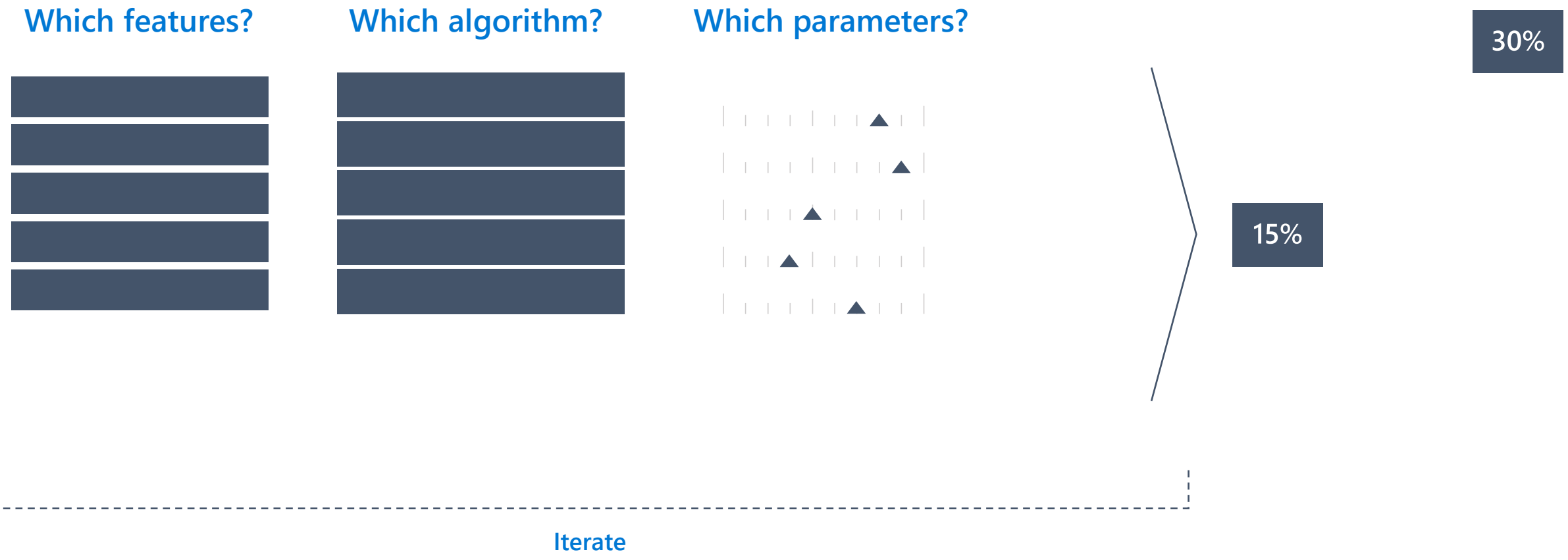
- Number of Neighbors
- Weights
- Min Samples Split
- Min Samples Leaf
- Others

30%

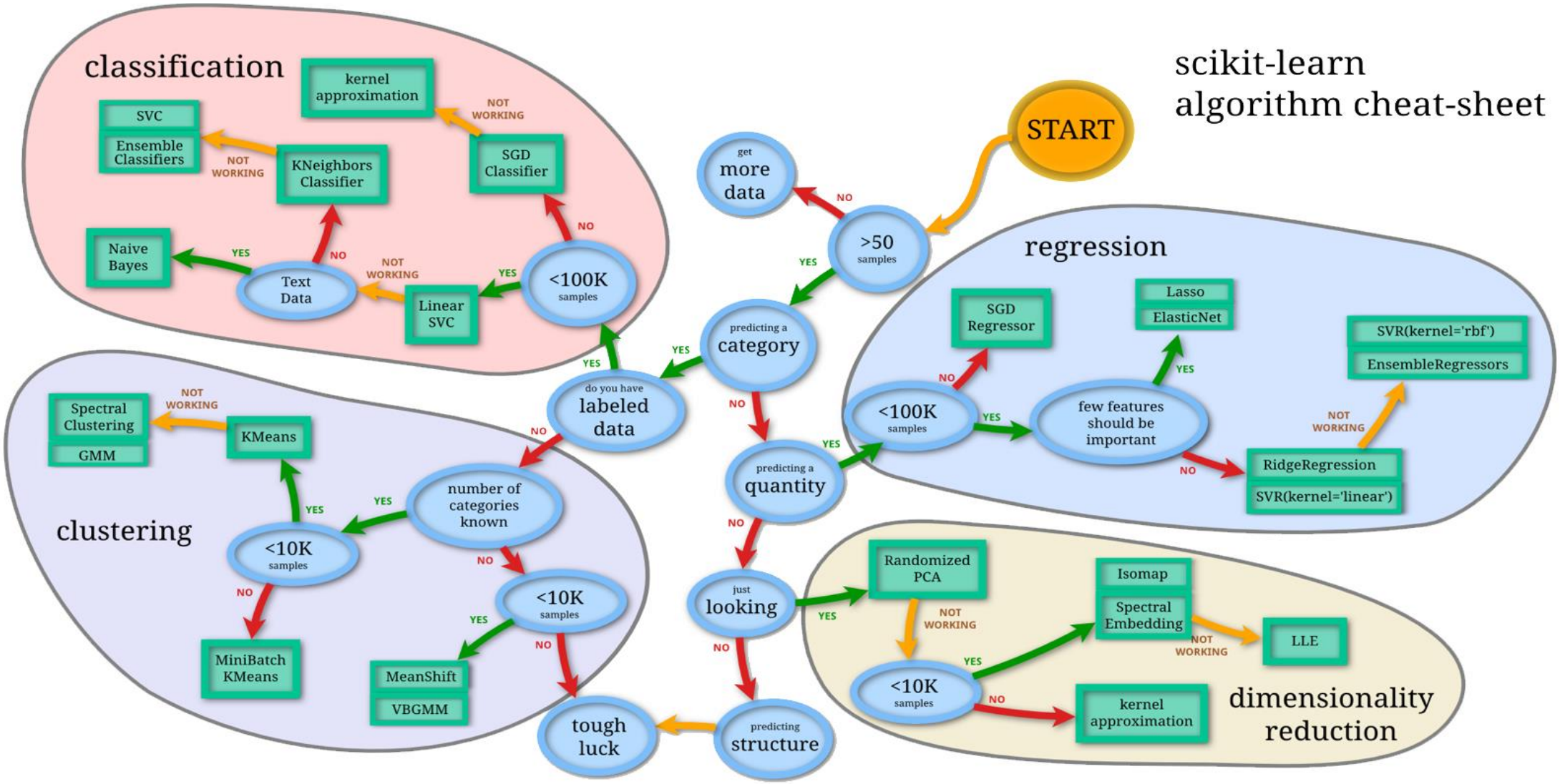
Model

Iterate

Model Creation Is Typically Time-Consuming

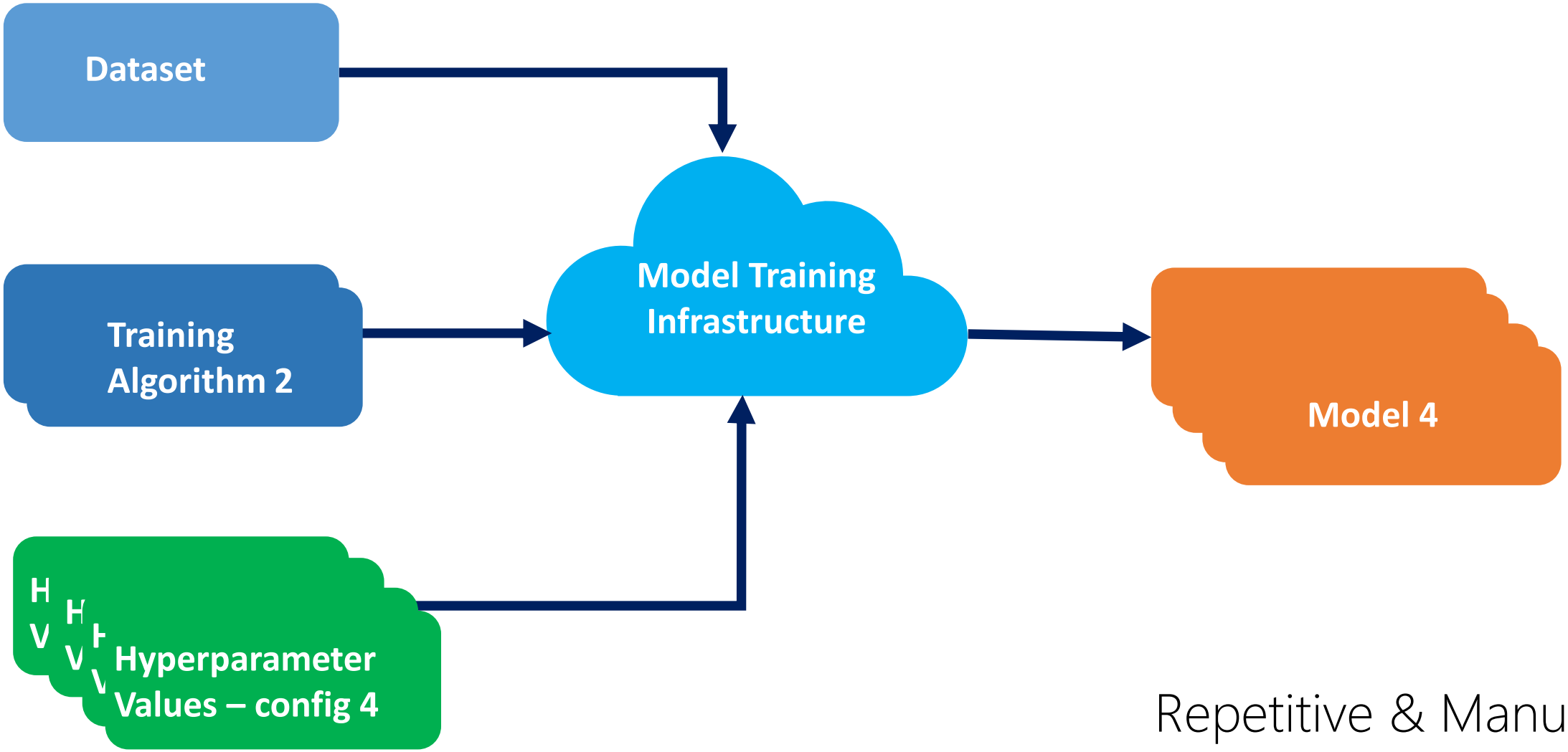


Machine Learning Complexity



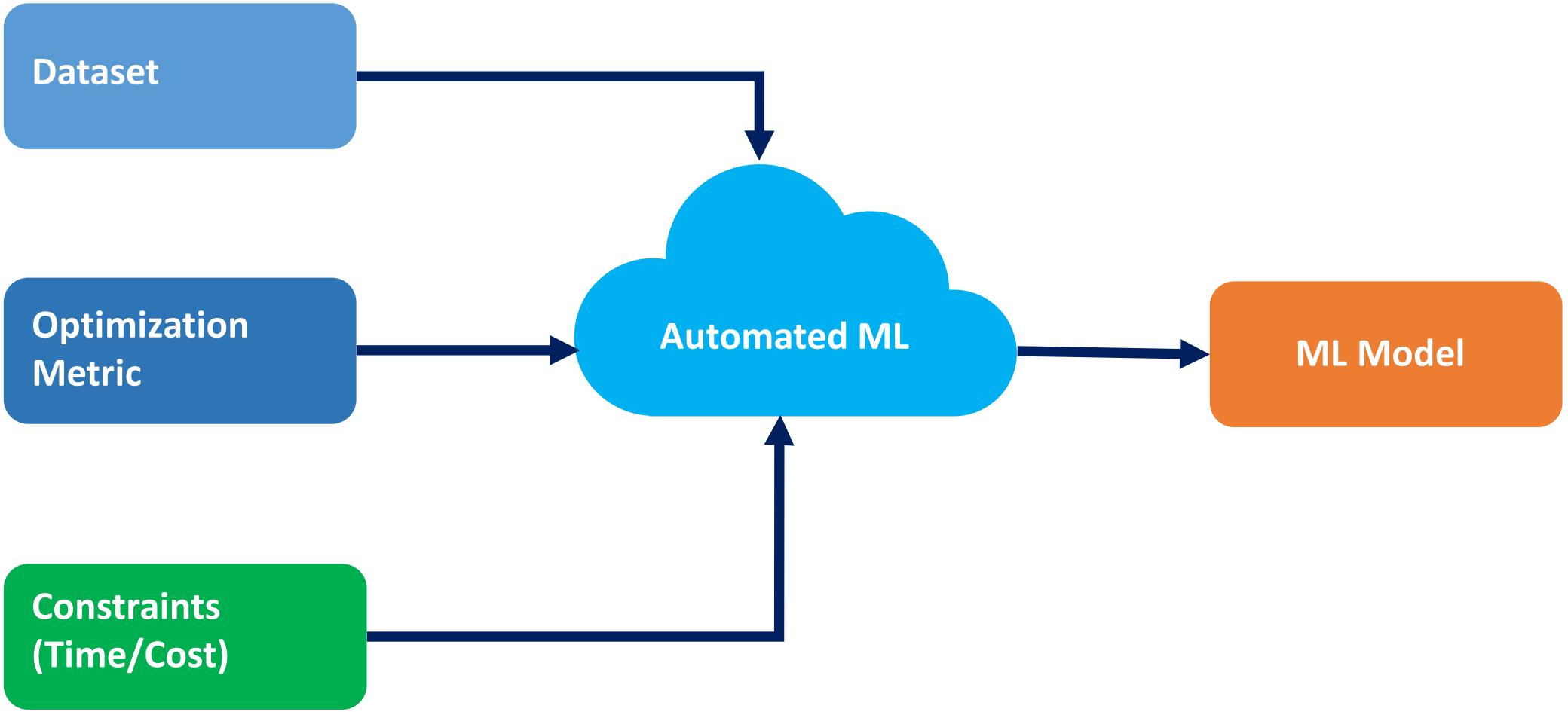
Source: http://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Model Selection & Hyperparameter Tuning



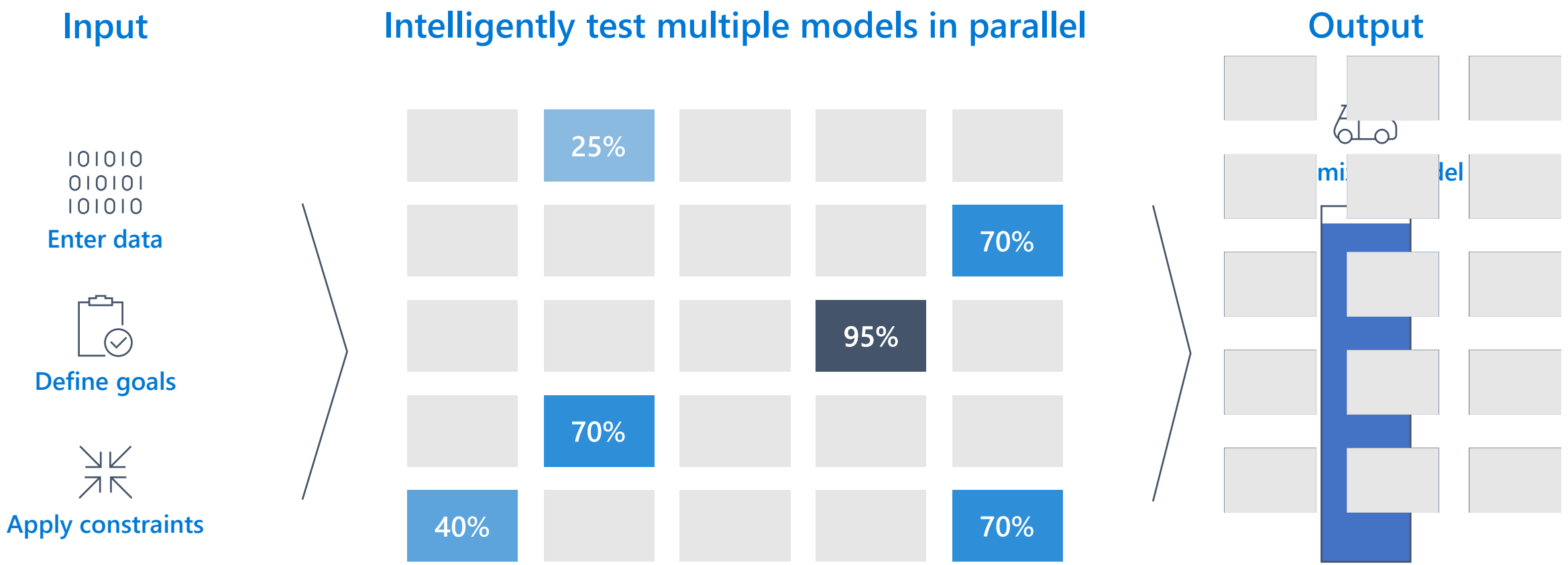
Repetitive & Manual

Introducing Automated Machine Learning



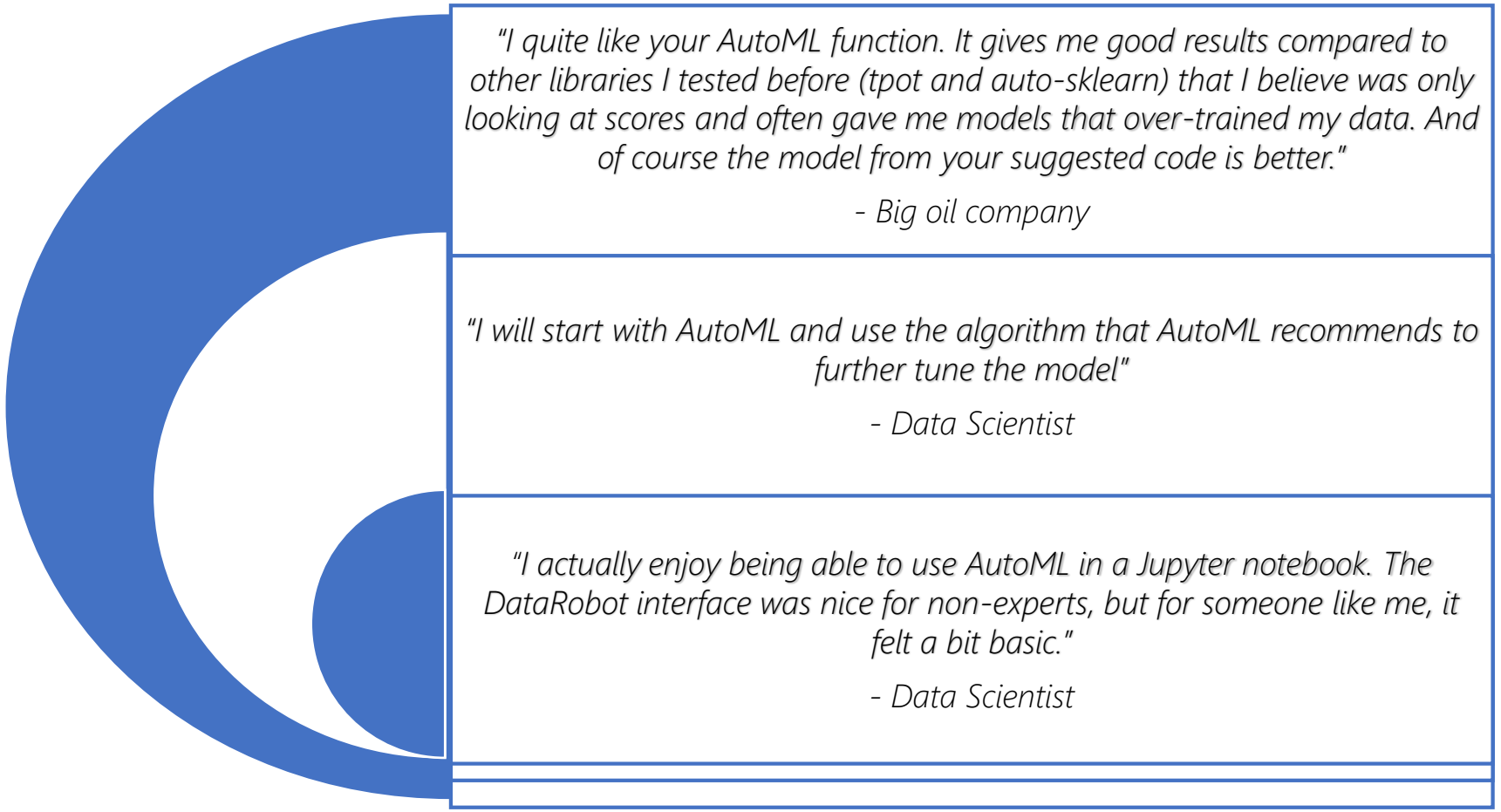
Accessible & Faster

Automated ML Accelerates Model Development



Automated ML Customer Testimonials

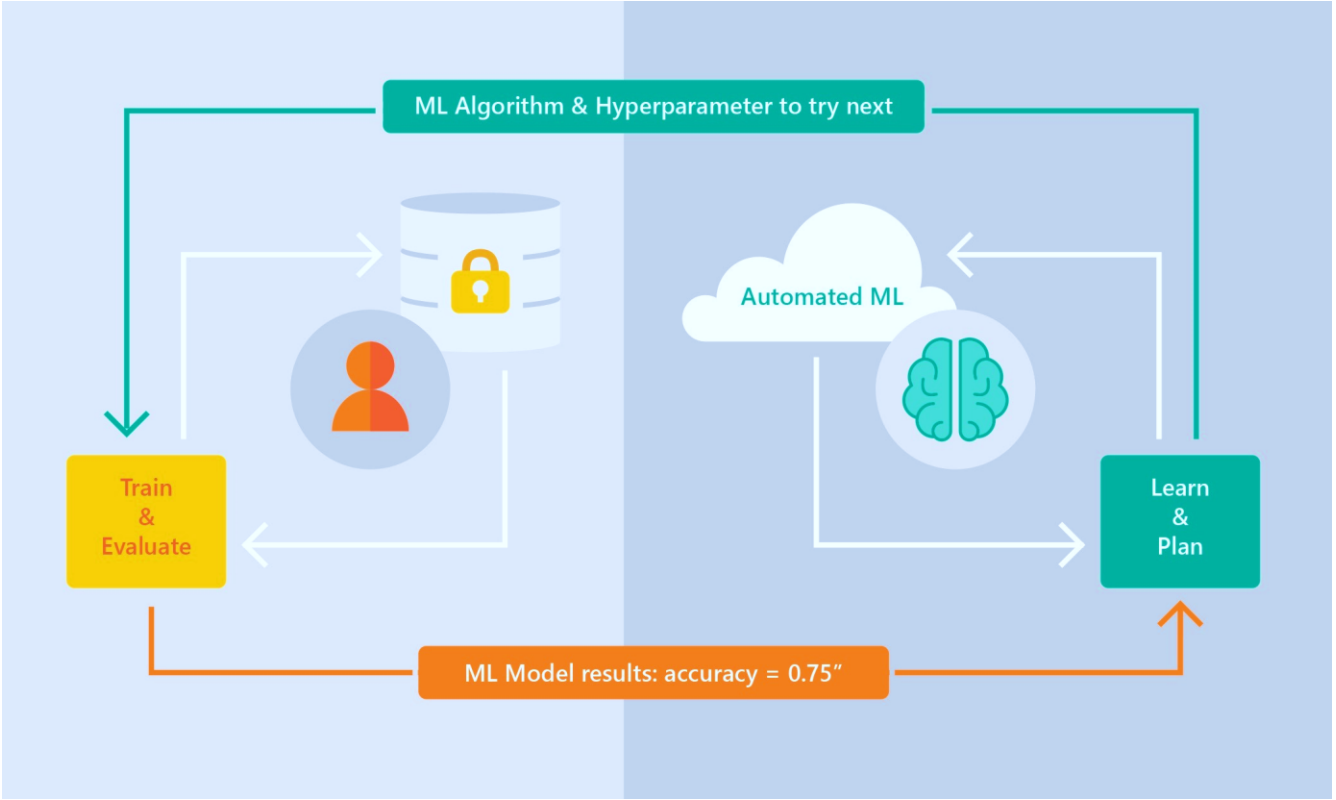
- Press-coverage from public preview:
  - CNET
  - VentureBeat
  - PRNewswire



Automated ML Capabilities

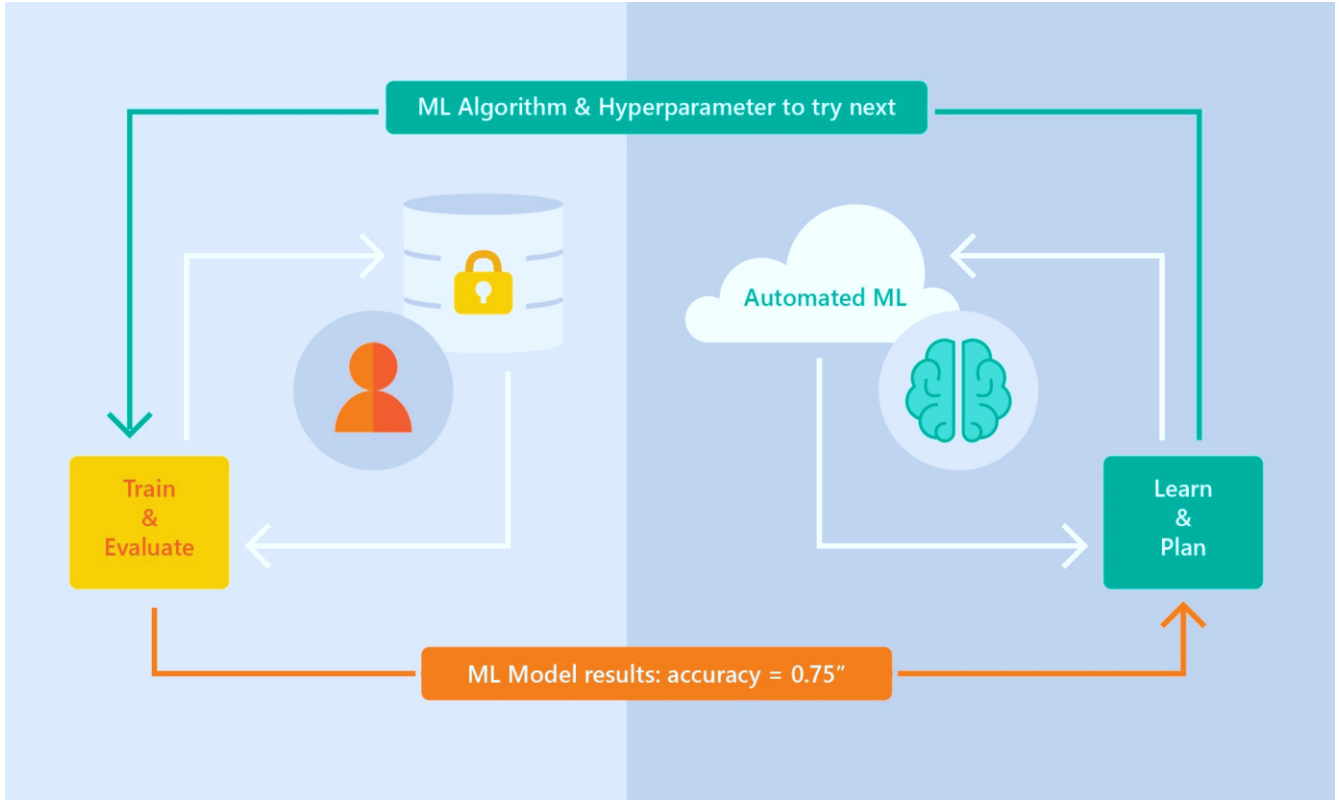
Automated ML Capabilities

- Based on Microsoft Research
- Brain trained with several million experiments
- Collaborative filtering and Bayesian optimization
- Privacy preserving: No need to “see” the data



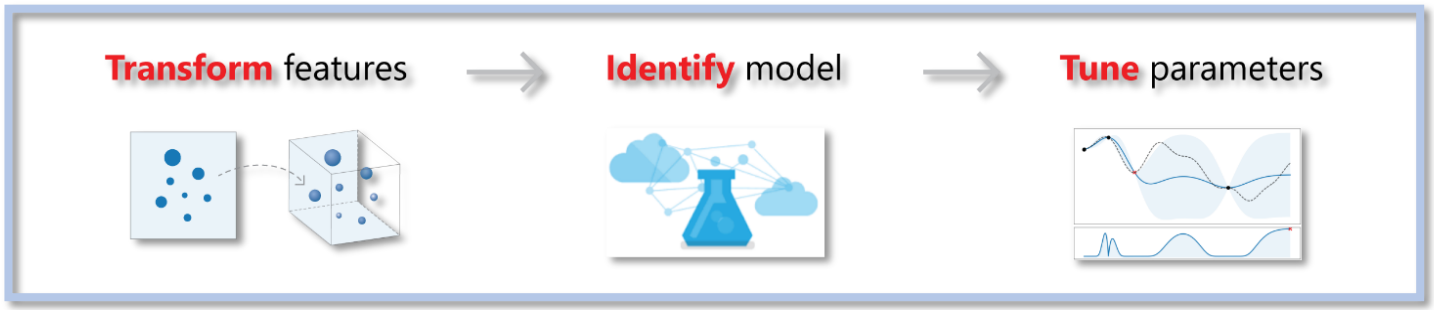
Automated ML Capabilities

- ML Scenarios: Classification & Regression, Forecasting
- Integration: Azure Machine Learning, Azure Notebooks, Jupyter Notebooks
- Data Type: Numeric, Text
- Languages: Python SDK for deployment and hosting for inference
- Training Compute: Local Machine, Remote Azure DSVM (Linux), Azure Batch AI, Databricks
- Transparency: View run history, model metrics
- Scale: Faster model training using multiple cores and parallel experiments



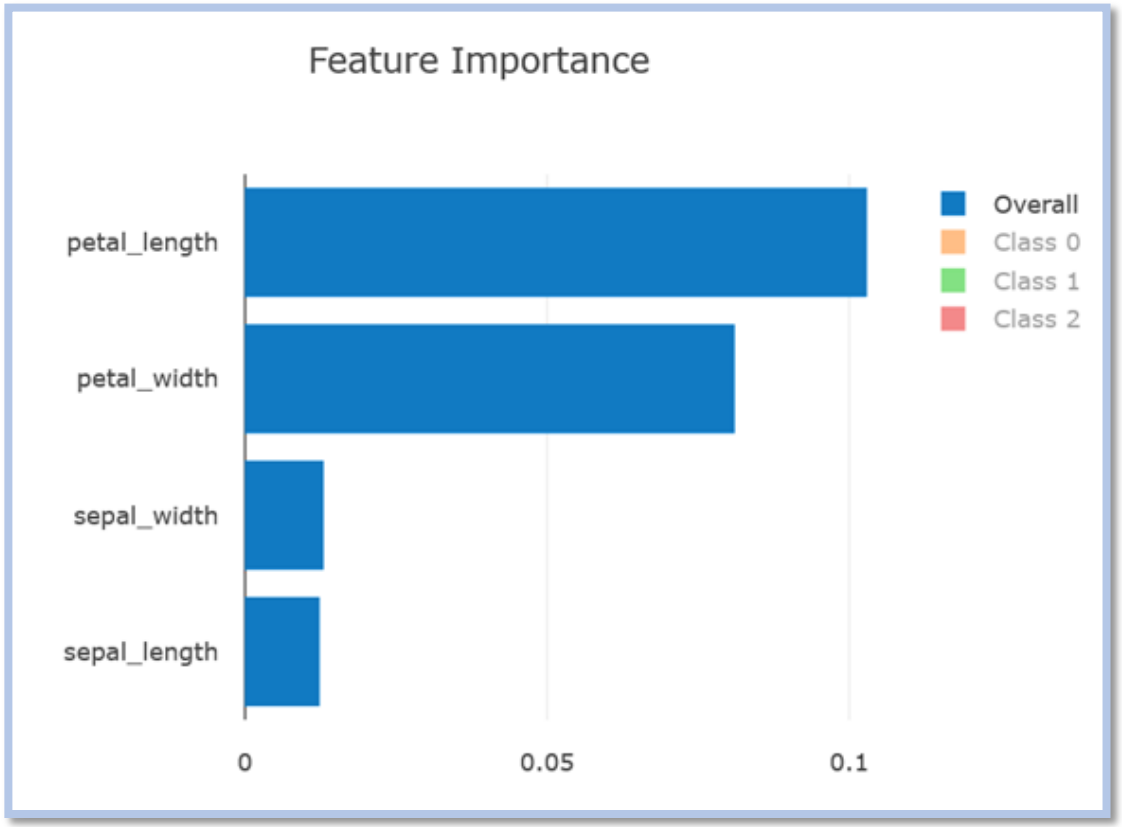
Feature Engineering

- Dropping high cardinality or no variance features
- Missing value imputation
- Generating additional features
- Transformations and encodings



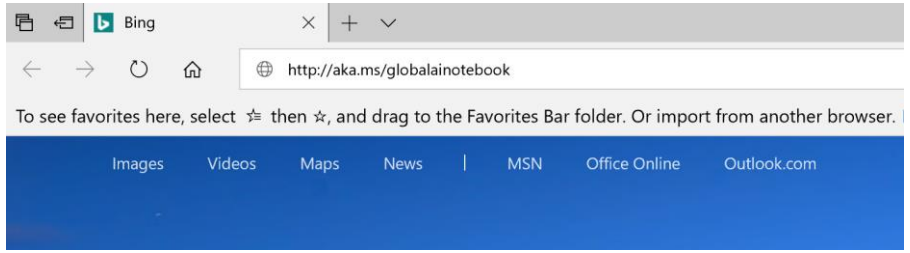
Model Explain-ability

- Feature importance as part of training
- Local feature importance for a given sample

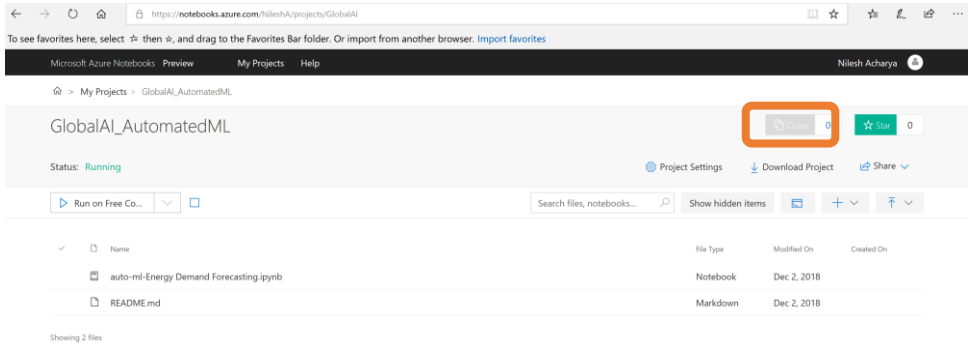


Workshop

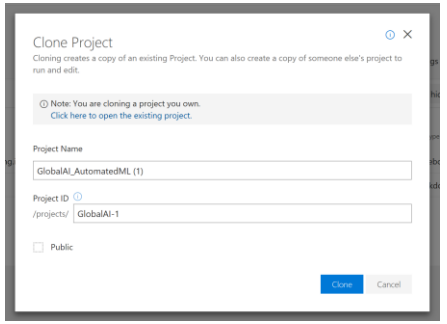
1. Go to aka.ms/GlobalAINotebook



2. Click on Clone



3.Enter Project name, Project ID and uncheck Public tick box



4. Open notebook and follow instructions

Thank You!

AI Platform Team

Resources

- AskAutomatedML@microsoft.com
- https://aka.ms/AutomatedML
- https://aka.ms/AutomatedMLDocs
- https://github.com/Azure/MachineLearningNotebooks