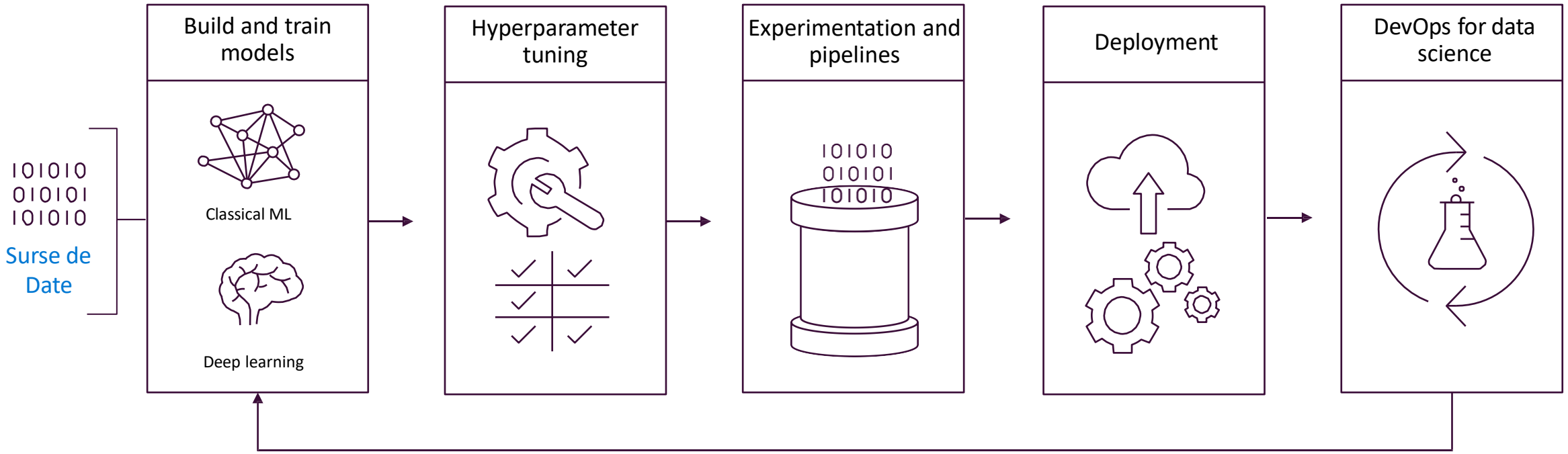
Cristian KEVORCHIAN

Facultatea de Matematică și Informatică

Automatizarea Proiectelor ML

DATA SCIENCE Project Lifecycle

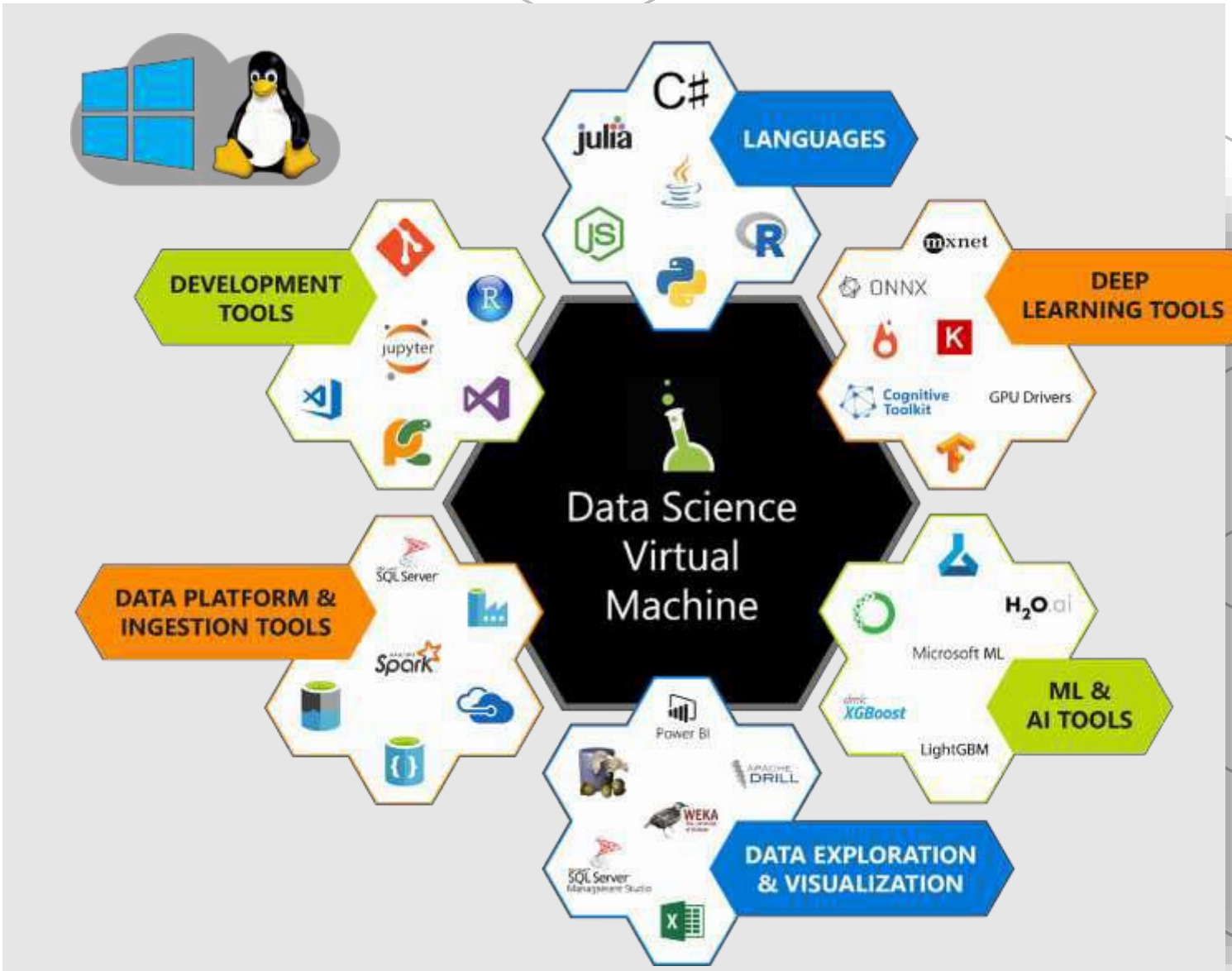


A hyperparameter is a setting that is made prior to the start of the learning process. These variables are adjustable and have a direct impact on how successfully a model trains. Hyperparameters in machine learning include the following:

- Rate of Learning
- Count of Epochs
- Momentum
- Constant of regularization
- A decision tree's number of branches
- A clustering algorithm's number of clusters (like k-means)

Data Science Virtual Machines(DSVM)

DSVMs are pre-installed, configured, and tested Azure Virtual Machine images that are often used for data analytics, machine learning, and AI training..



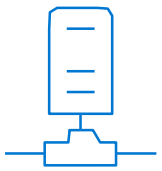
Training infrastructure

Containers and Orchestration
Use computing of AML (cloud managed) infrastructure or through your own infrastructureJob scheduling

Distributed Data
Resource management and sharing in the workplace

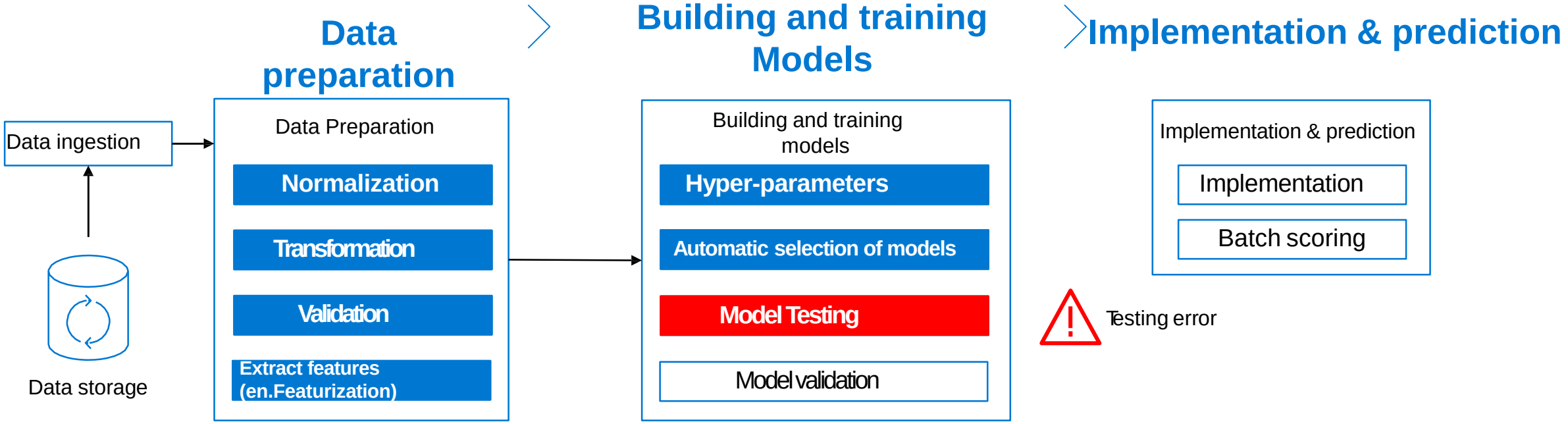
Training with cloud scaling using the chosen framework

Scalling resources
Self-scaling of resources with payment for the execution of the job

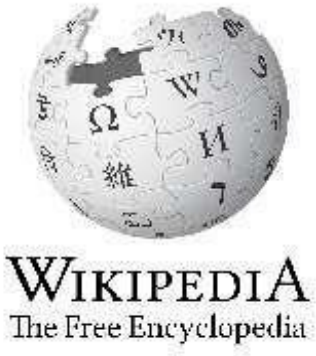


Provision of cluster VM Utilizarea ultimatelor serii de MV NDv2 cu GPU NVIDIA V100

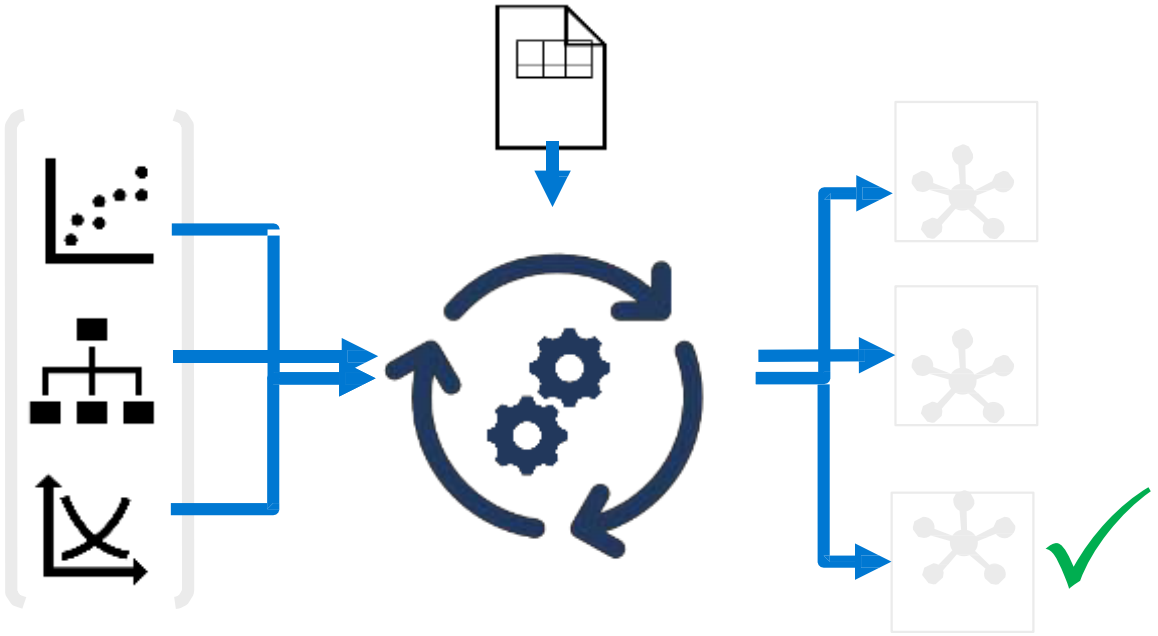
Pipeline-uri Azure ML



ML Automation

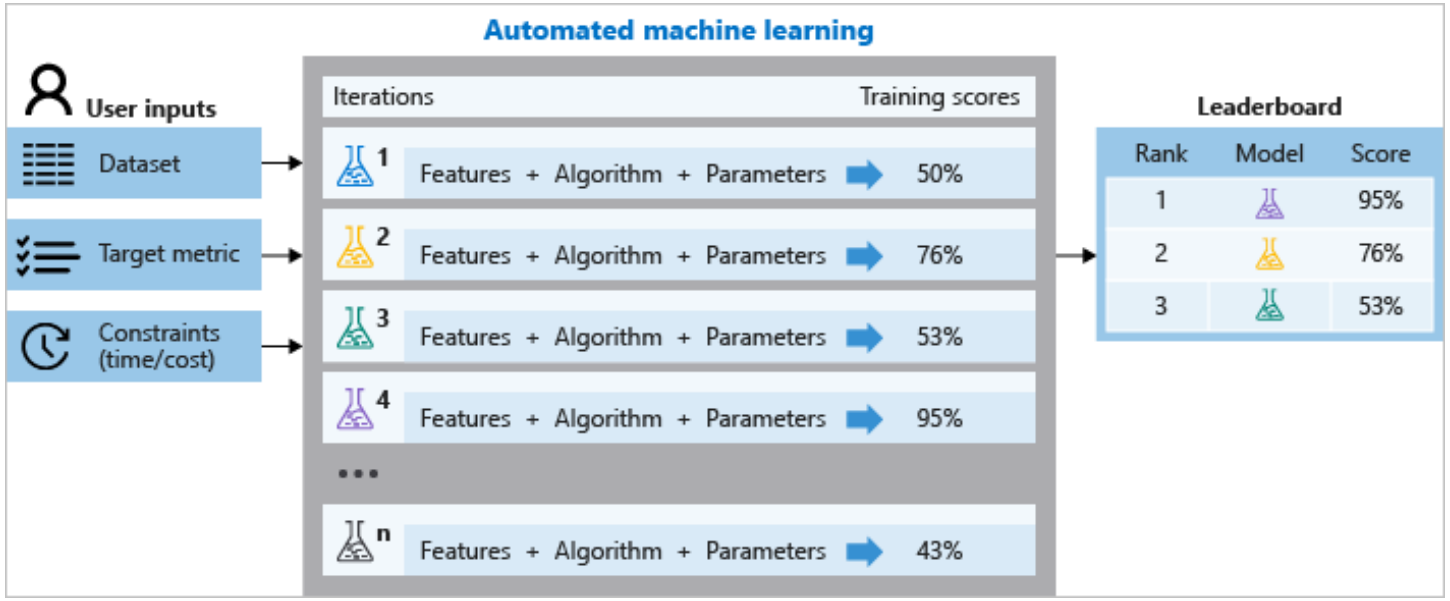


Automated machine learning (AutoML) is the process of automating end-to-end the process of applying machine learning to real-world problems



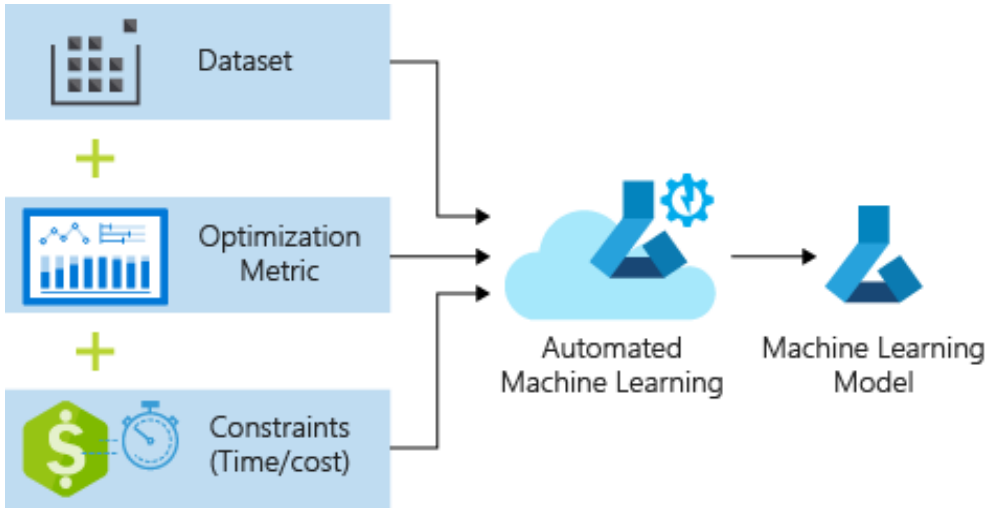
The process of automating iterative and time-consuming activities for the building of machine learning models is known as ML automation, sometimes known as AutoML. Allows DS, analysts, and developers to design high-complexity ML models while maintaining model quality and efficiency.

Automation ML with Azure AutoML

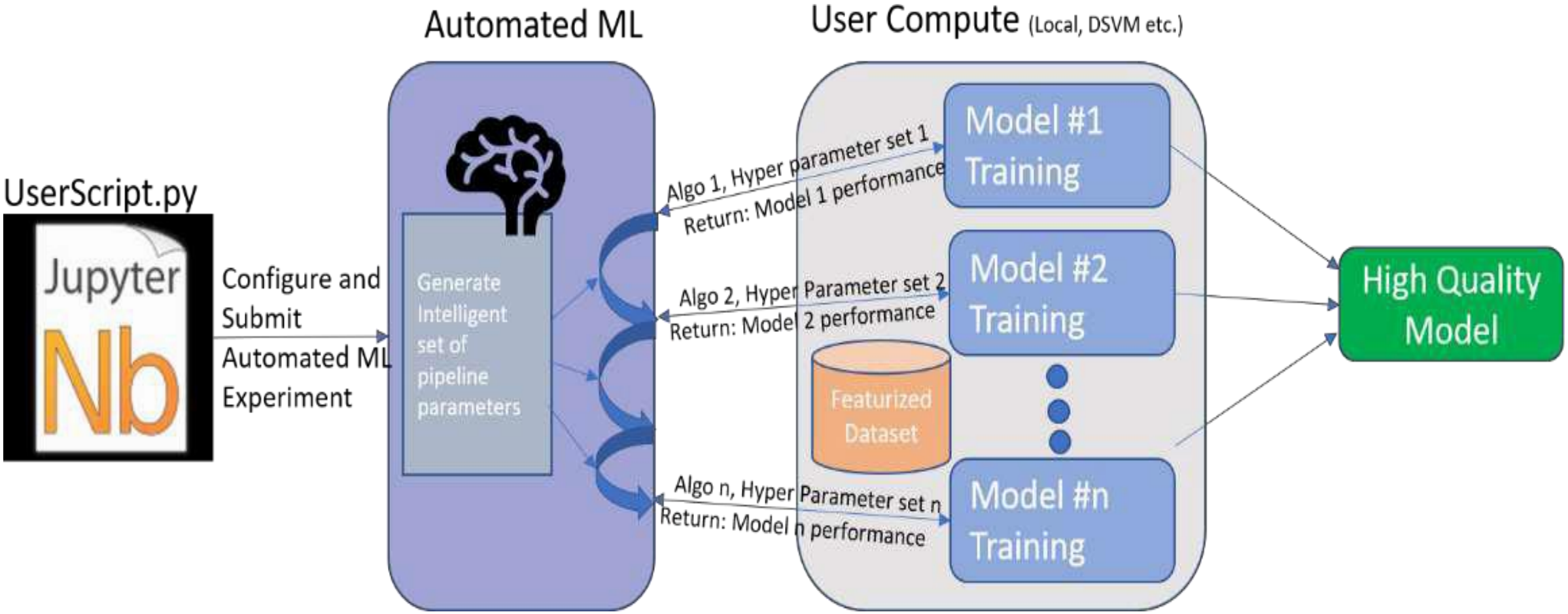


DS-isti, developers as users of AutoML

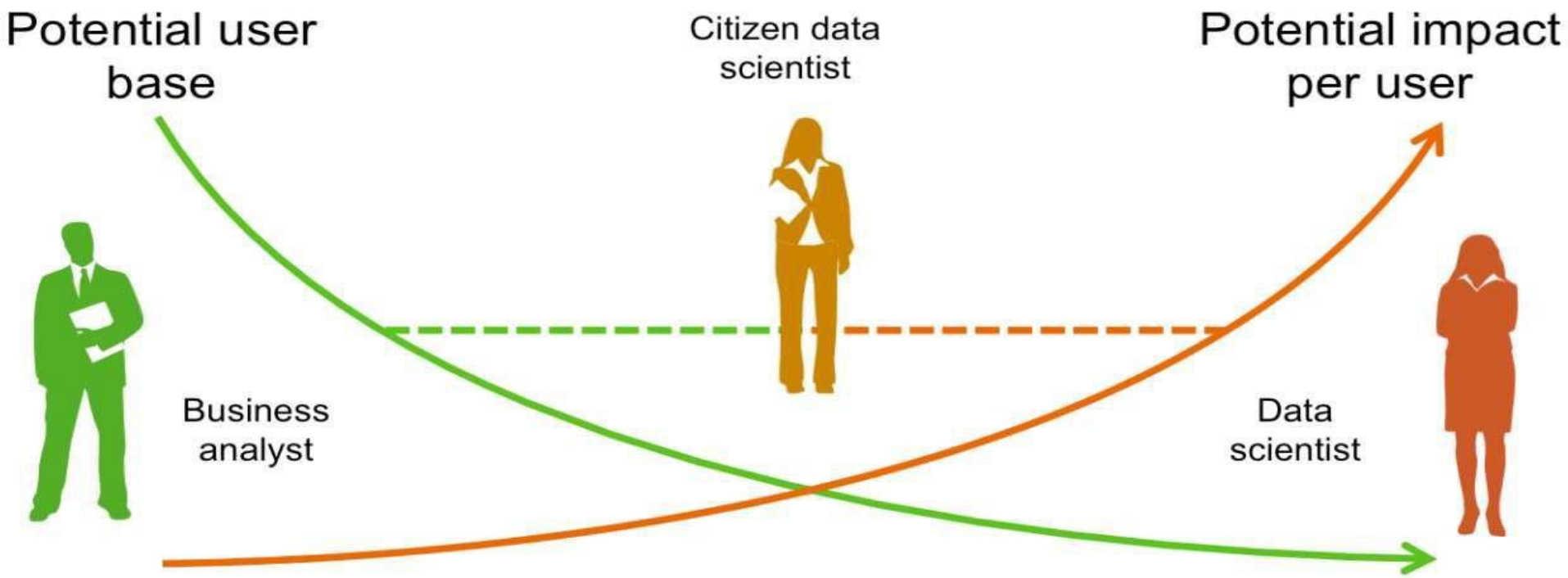
- Utilizing best practices
- Agility in problem solving



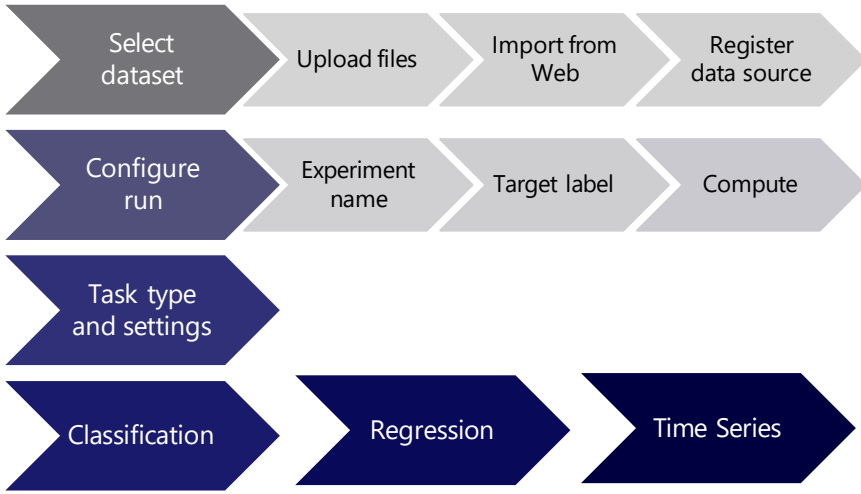
Work with Azure AutoML

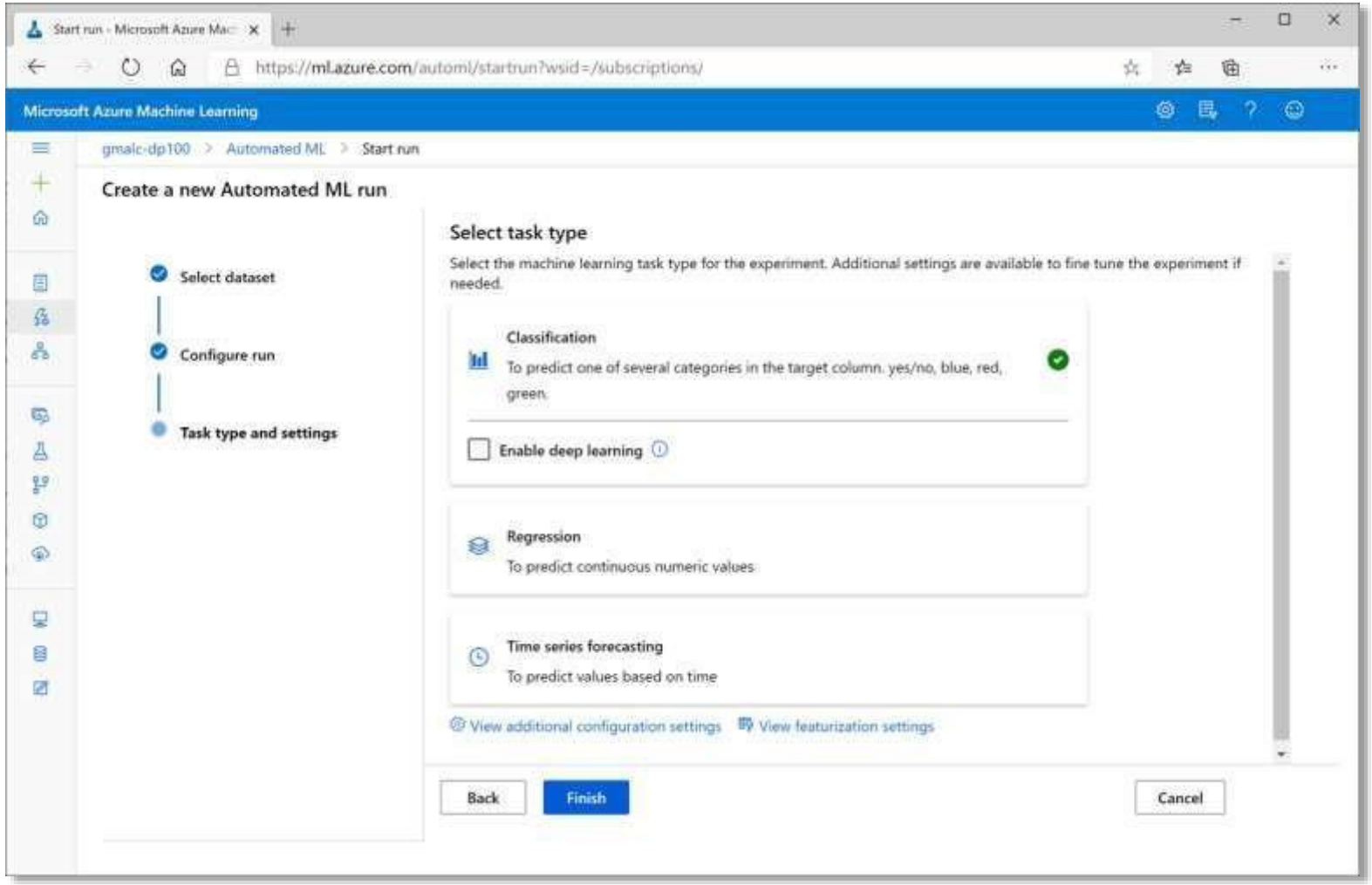


A citizen data scientist, according to Gartner, is "someone who creates or generates models that use predictive or prescriptive analytics, but whose primary job function is outside of the field of statistics and analytics." They serve as a bridge between those who perform self-service analytics as business users and those who perform advanced analytics as data scientists.



Automated ML în Azure Machine Learning Studio





Supported Algorithms

Logistic Regression	Elastic Net	Elastic Net
Stochastic Gradient Descent (SGD)	Light GBM	Light GBM
Naive Bayes	Gradient Boosting	Gradient Boosting
C-Support Vector Classification (SVC)	Decision Tree	Decision Tree
Linear SVC	KNearest Neighbors	KNearest Neighbors
KNearest Neighbors	LARS Lasso	LARS Lasso
Decision Tree	Stochastic Gradient Descent (SGD)	Stochastic Gradient Descent (SGD)
Random Forest	Random Forest	Random Forest
Extremely Randomized Trees	Extremely Randomized Trees	Extremely Randomized Trees
Gradient Boosting		
Light GBM		

Demo