Art of Machine Learning for Beginners: Azure Machine Learning Automated ML

Milwaukee Tech Hub Summer 2021 Pre-Internship program
Azure Machine Learning for Beginners

Balamurugan Balakreshnan – Cloud Solution Architect (Data and AI)


Microsoft
Thank You

Agenda

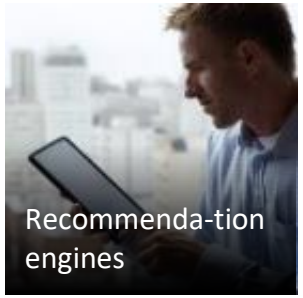
Agenda:

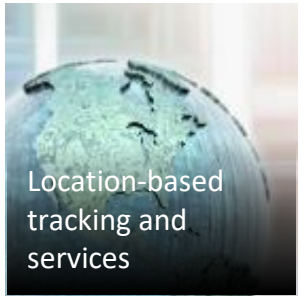
Introduction
Azure Machine learning Intro — 1 hour
OpenHack use case introduction — 15 mins
OpenHack – 4 Hours
Take Break when experimentation starts
Deploy model — 1 hour
Clean up — 15 minutes

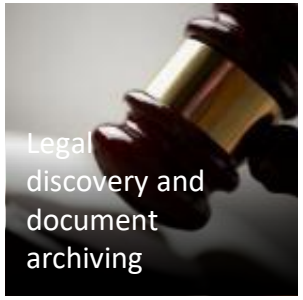
Welcome Video

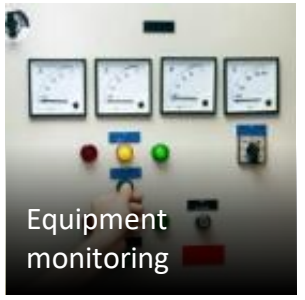
Machine Learning / Predictive Analytics

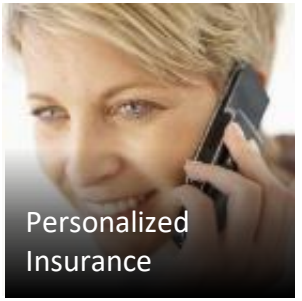
Machine learning & predictive analytics are core capabilities that are needed throughout your business











Machine Learning Overview

- Formal definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" - *Tom M. Mitchell*
- Another definition: "The goal of machine learning is to program computers to use example data or past experience to solve a given problem." – *Introduction to Machine Learning, 2nd Edition, MIT Press*
- ML often involves two primary techniques:
  - Supervised Learning: Finding the mapping between inputs and outputs using correct values to "train" a model
  - Unsupervised Learning: Finding patterns in the input data (similar to *Density Estimates* in Statistics)

Machine Learning

Data:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Rules, or Algorithms:

about, Learning, language – Spelling and sounding builds words

Learning about language. – Words build sentences

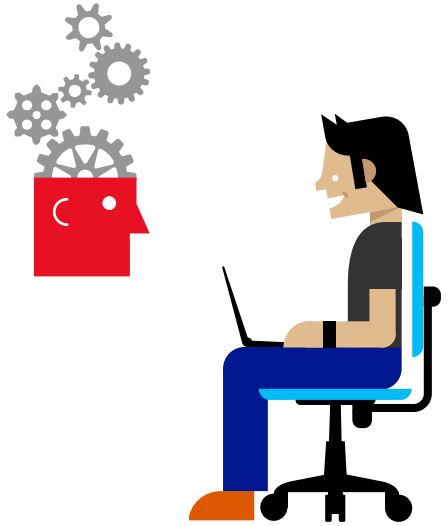
Learning, or Abstraction:

Any new understanding proceeds from previous knowledge.

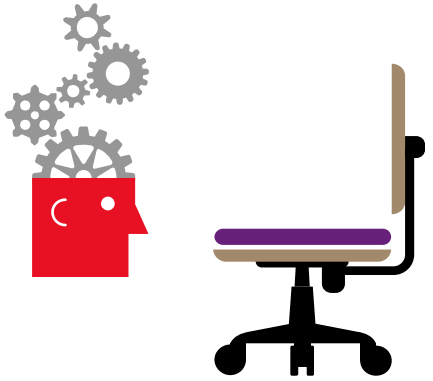
Machine Learning Algorithms

Split into two main categories:

- Supervised learning
  - Predicting the future
  - Learn from known past examples to predict future
  - Labels provided

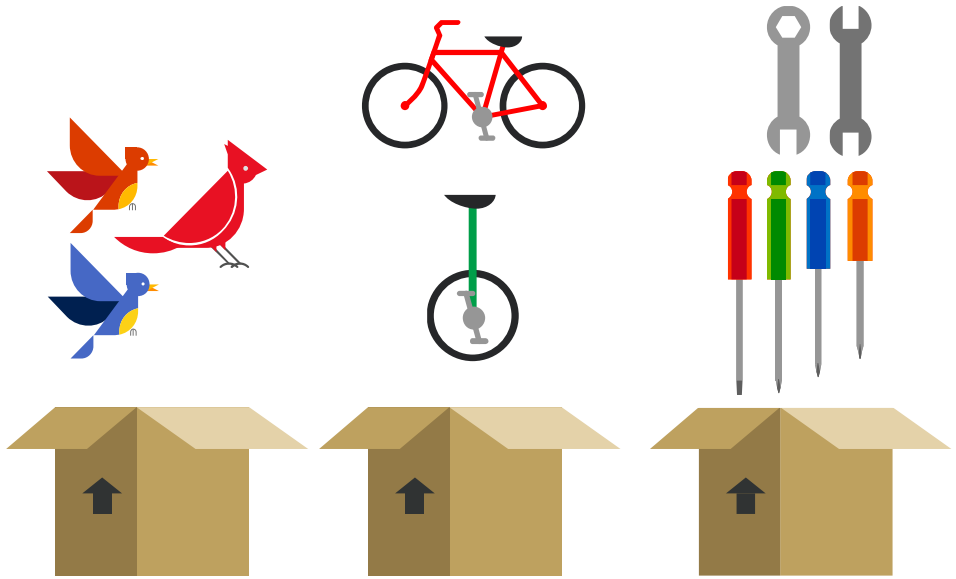


- Unsupervised learning
  - Making sense of data
  - Understanding the past
  - Learning the structure of data
  - Labels no provided



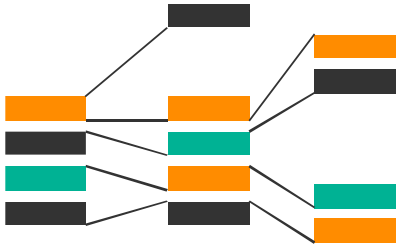
Machine Learning Capabilities

Which category
(Classification)

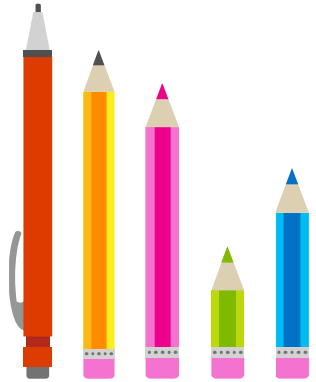


How much/many
(Regression)

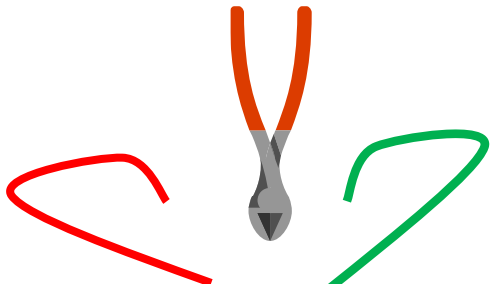
Which group
(Clustering, Recommender)



Is it odd
(Anomaly)



Which action
(Reinforcement Learning)



Supervised Learning

1. Used when you want to predict unknown answers from answers you already have – requires data which shows the answers you can get now
2. Data is divided into two parts: the data you will use to "teach" the system (*data set*), and the data you will use to see if the computer's algorithms are accurate (*test set*)
3. After you select and clean the data, you select data points that show the right relationships in the data. The answers are "labels", the categories/columns/attributes are "features" and the values are…values.
4. Then you select an algorithm to compute the *outcome*. (Often you choose more than one)
5. You run the program on the data set, and check to see if you got the right answer from the test set.
6. Once you perform the experiment, you select the best *model*. This is the final output – the model is then used against more data to get the answers you need

Supervised Learning

1. Car

2. Not Car

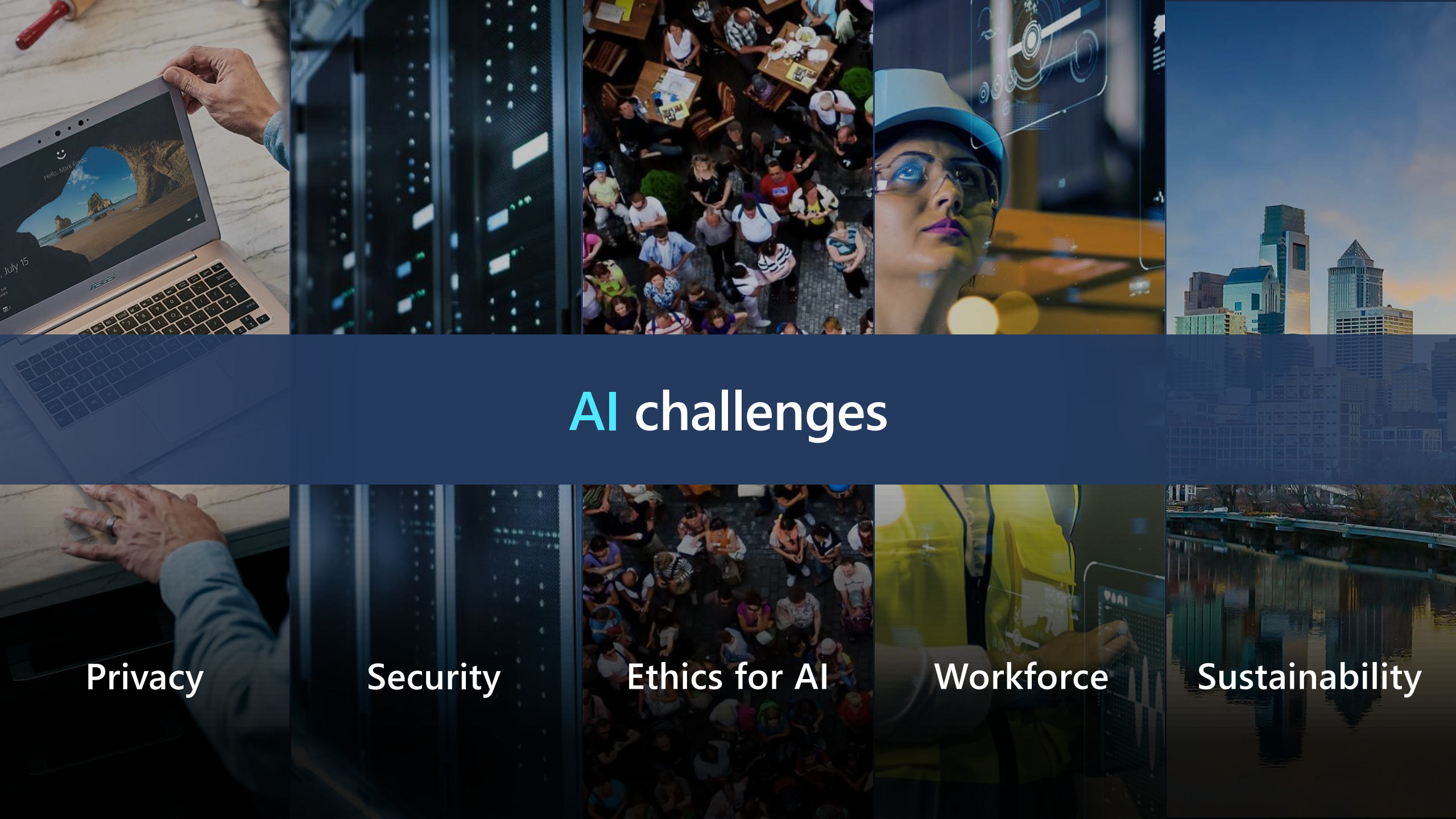
Unsupervised Learning

1. Used when you want to find unknown answers – mostly groupings - directly from data
2. No simple way to evaluate accuracy of what you learn
3. Evaluates more vectors, groups into sets or classifications
4. Start with the data
5. Apply algorithm
6. Evaluate groups

Unsupervised Learning

Example 1 example A Example 2
example B Example 3 example C

example A	example B	example C
Example 1	Example 2	Example 3


AI challenges
Privacy
Security
Ethics for AI
Workforce
Sustainability


Addressing broader challenges
Privacy
Security
Ethics for AI
Workforce
Sustainability

Azure AI

Azure AI

Scenario specific services

Bot Service

Cognitive Search

Form Recognizer

Video Indexer

Customizable AI models

Vision, Speech, Language, Decision

Cognitive Services

Machine Learning service

Azure Machine Learning

Data & Analytics

App Development & Tools

Compute

61 Regions

90+ Compliance Offerings

95% Fortune 500 use Azure

$1B Security investment per year

Azure

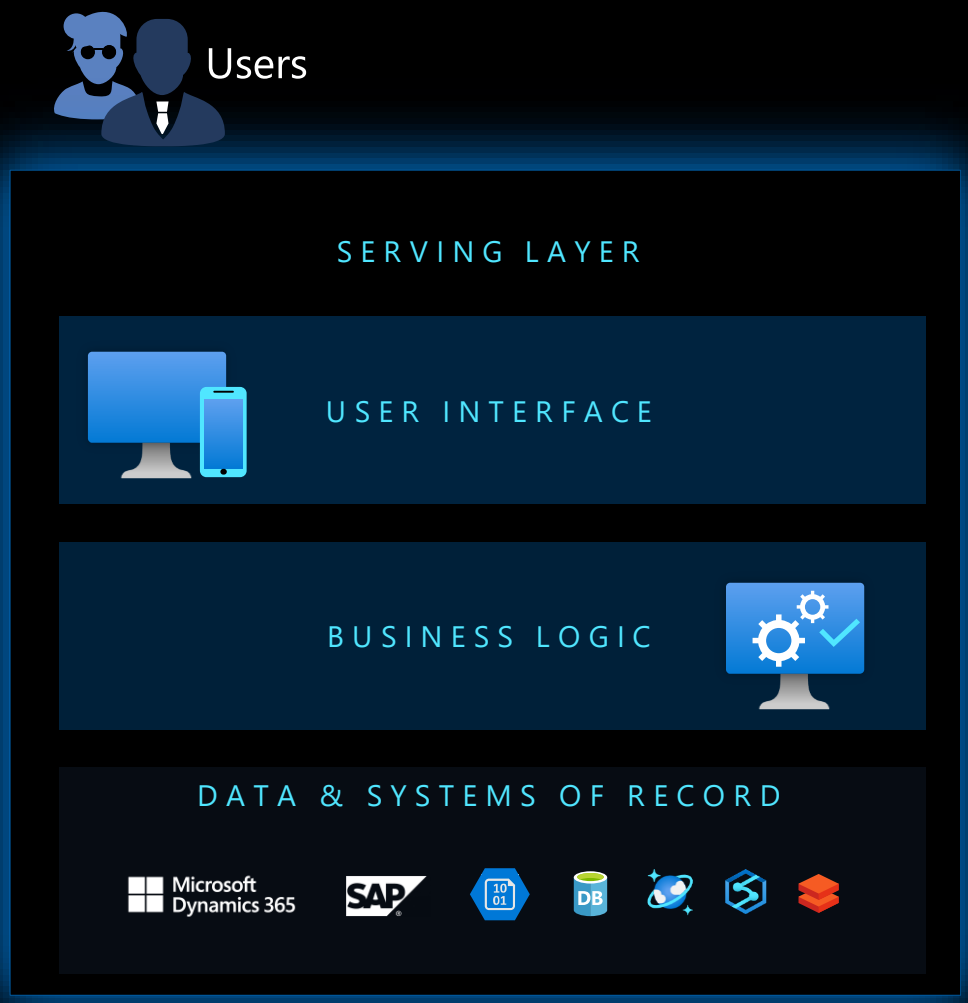
Common AI Patterns

Machine Learning

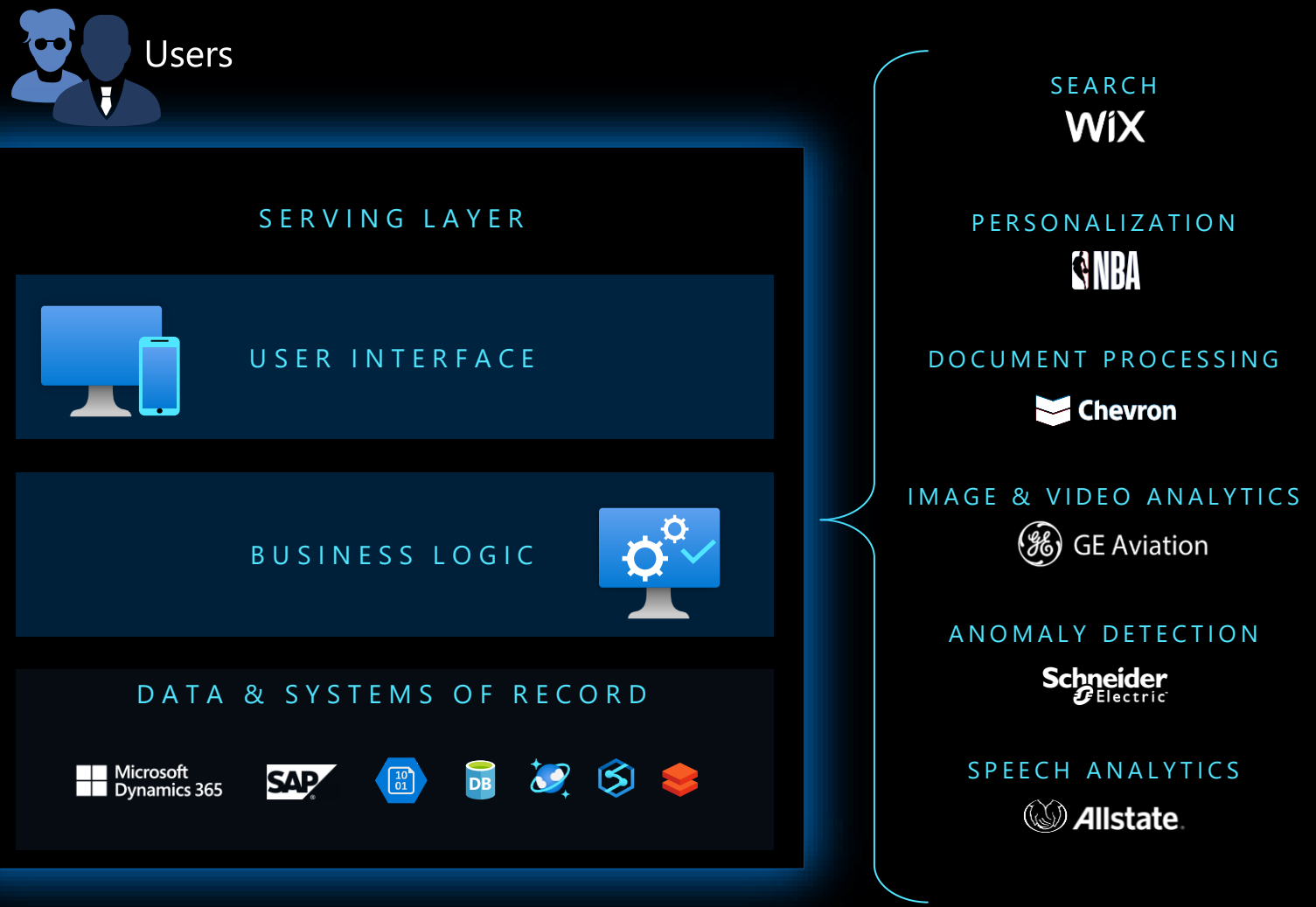
AI-powered Apps

Knowledge Mining

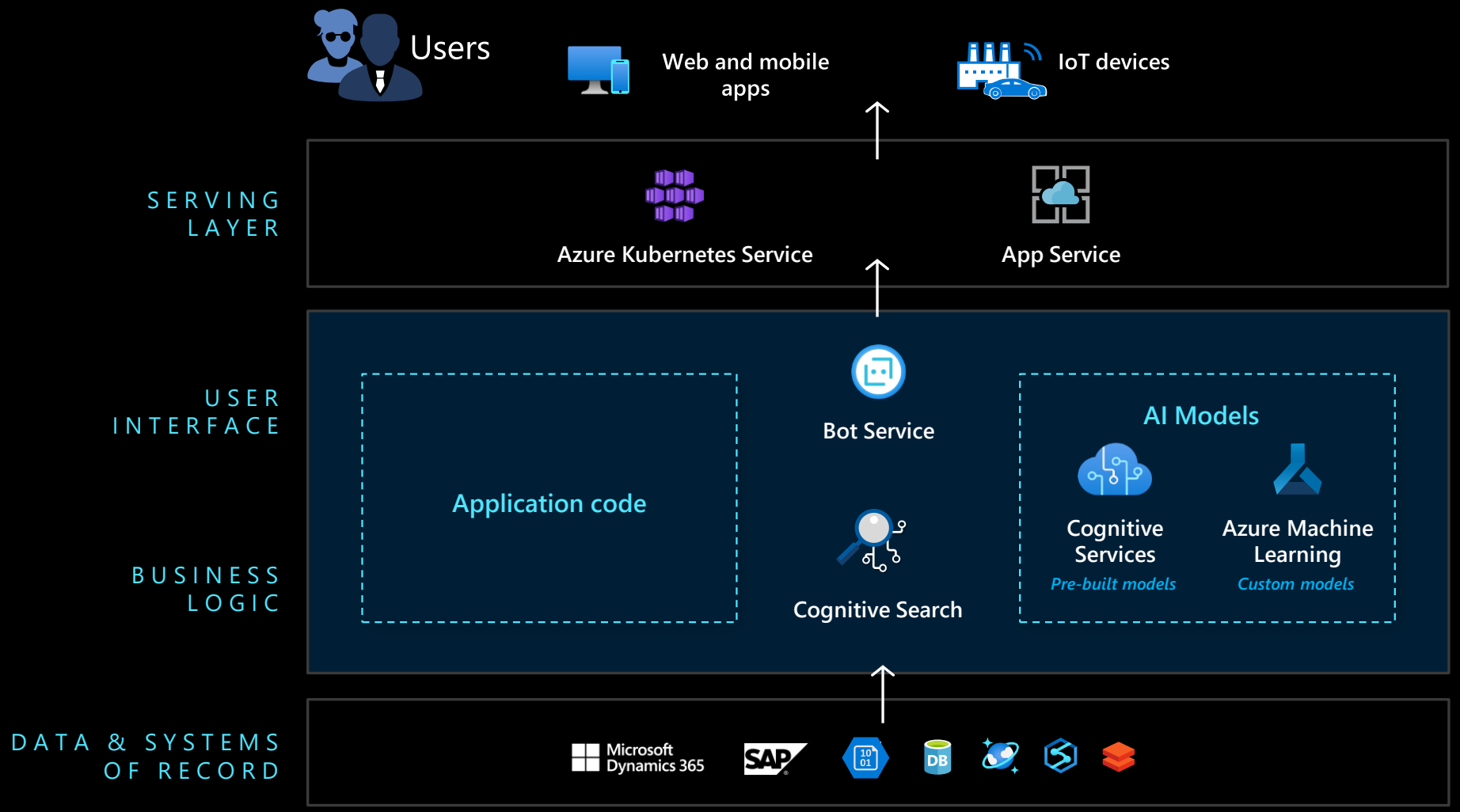
What are AI-powered apps?



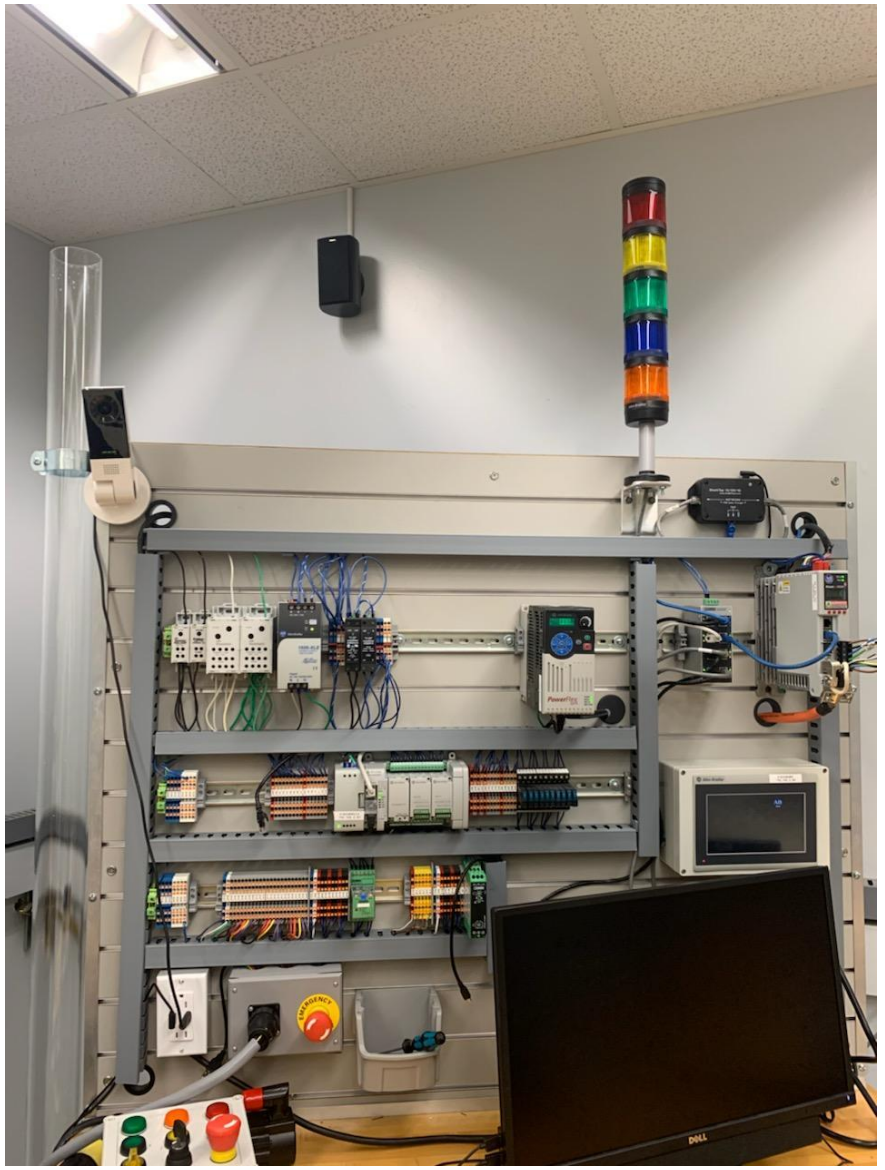
AI-powered apps



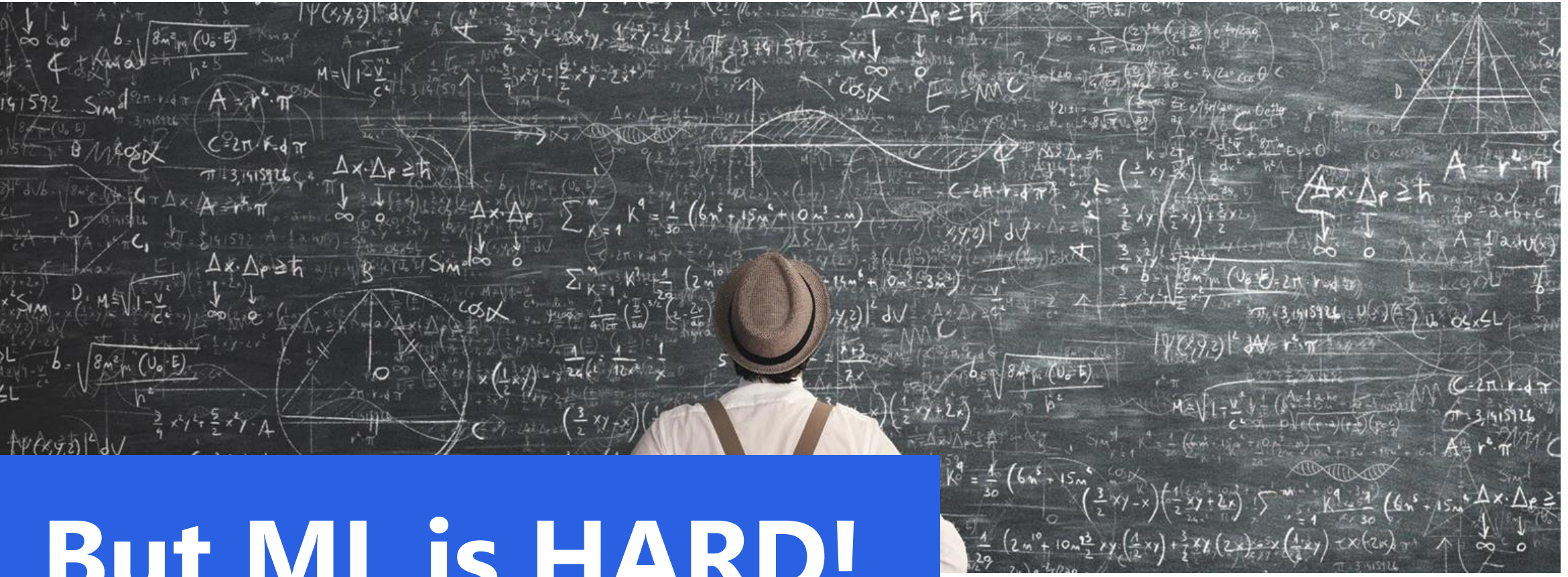
Building AI-powered apps



AI in Action: AI for Good – Workplace Safety



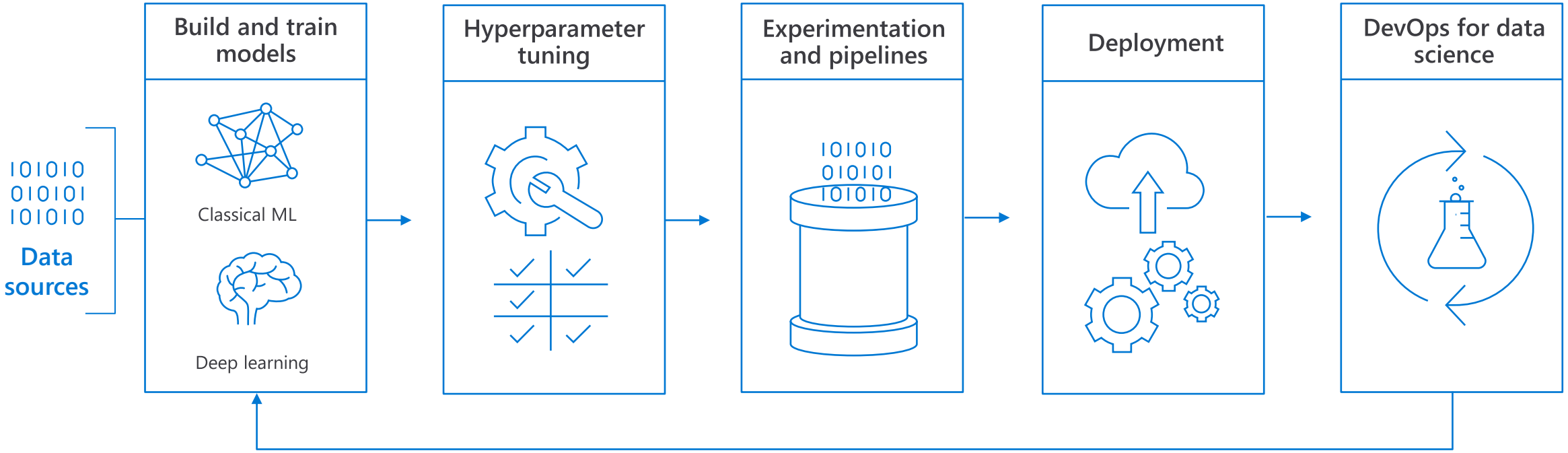
Workplace Safety

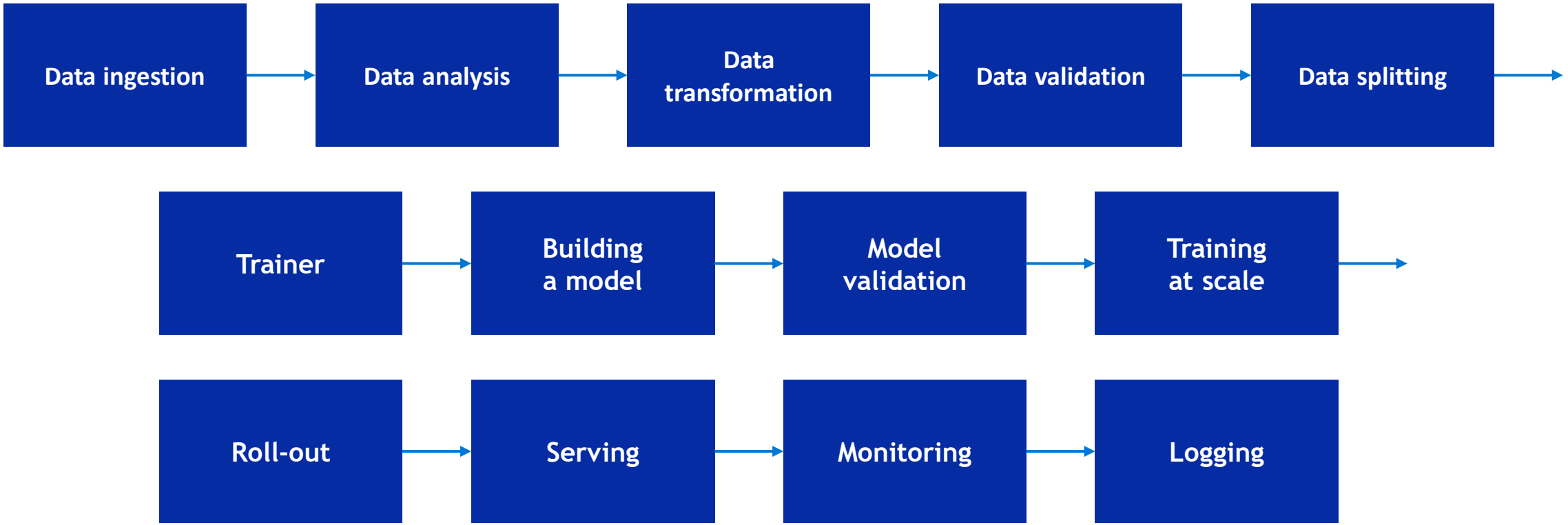


But ML is HARD!

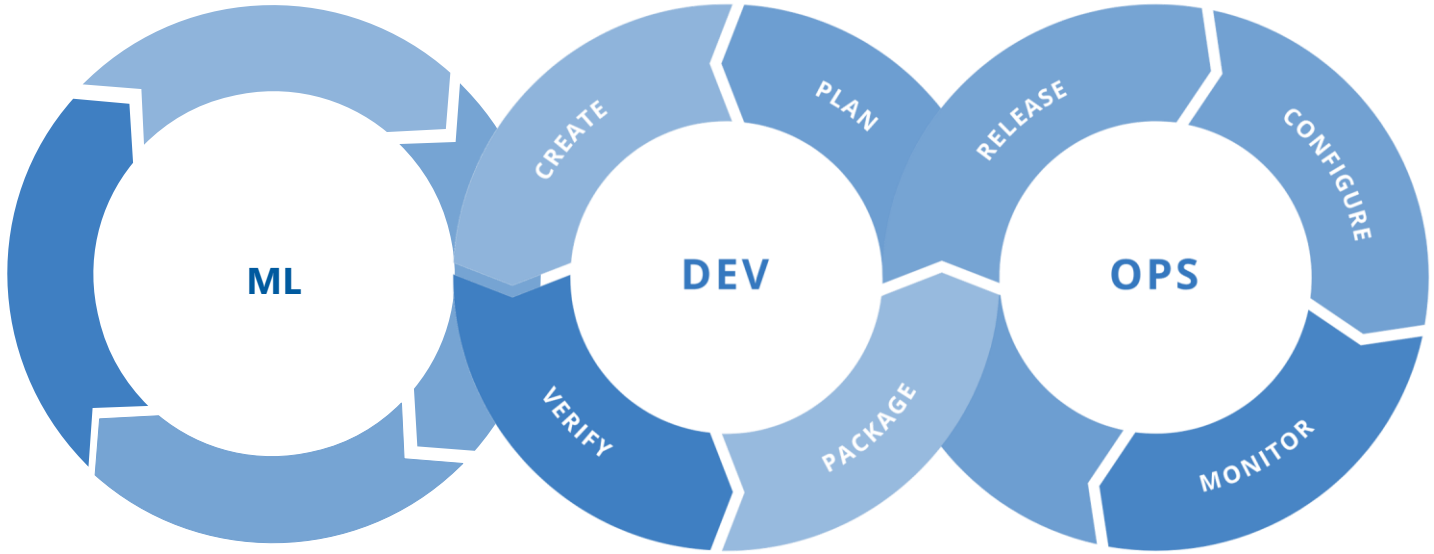
Building a model

Building blocks for a Data Science Project




Data ingestion
Data analysis
Data transformation
Data validation
Data splitting
Trainer
Building a model
Model validation
Training at scale
Roll-out
Serving
Monitoring
Logging

MLOps = ML + DEV + OPS



Experiment
Data Acquisition
Business Understanding
Initial Modeling

Develop
Modeling + Testing
Continuous Integration
Continuous Deployment

Operate
Continuous Delivery
Data Feedback Loop
System + Model Monitoring

Machine Learning on Azure

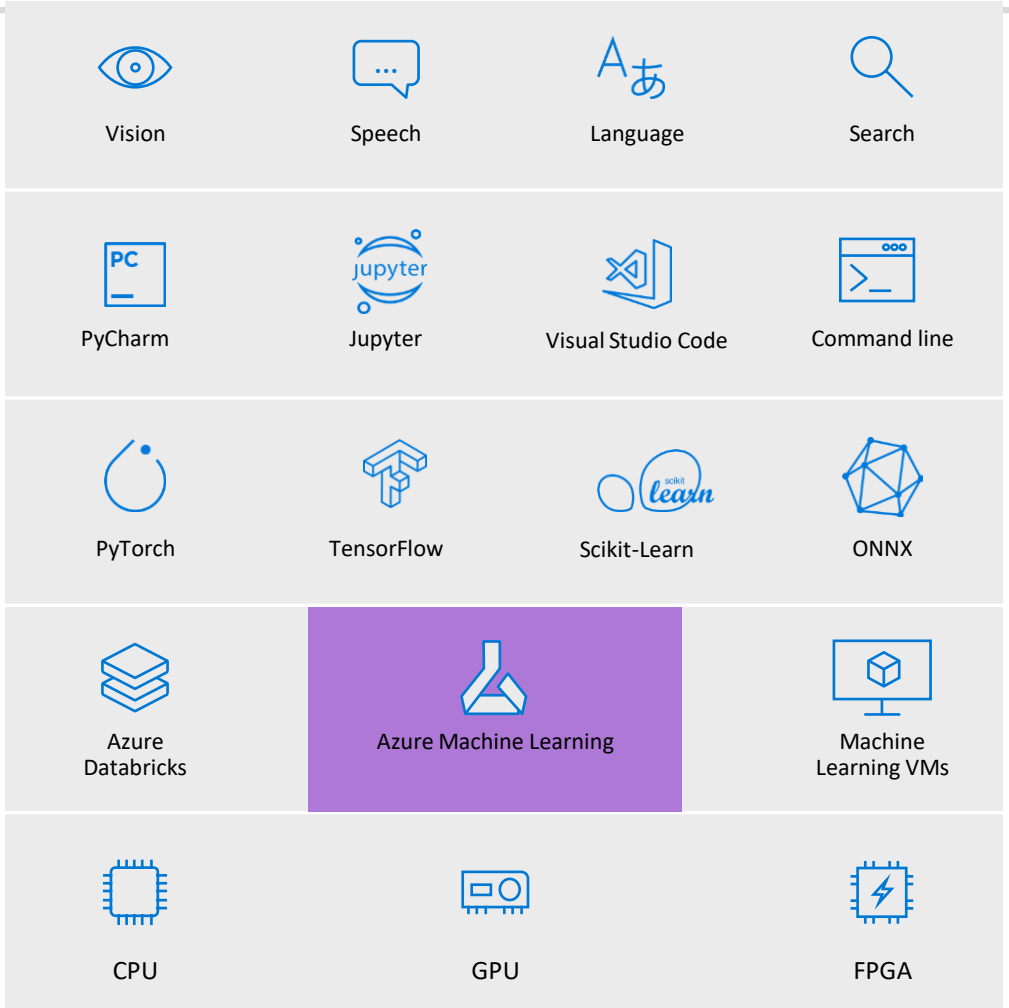
Domain Specific Pretrained Models
To reduce time to market

Familiar Data Science Tools
To simplify model development

Popular Frameworks
To build machine learning and deep learning solutions

Productive Services
To empower data science and development teams

Powerful Hardware
To accelerate deep learning



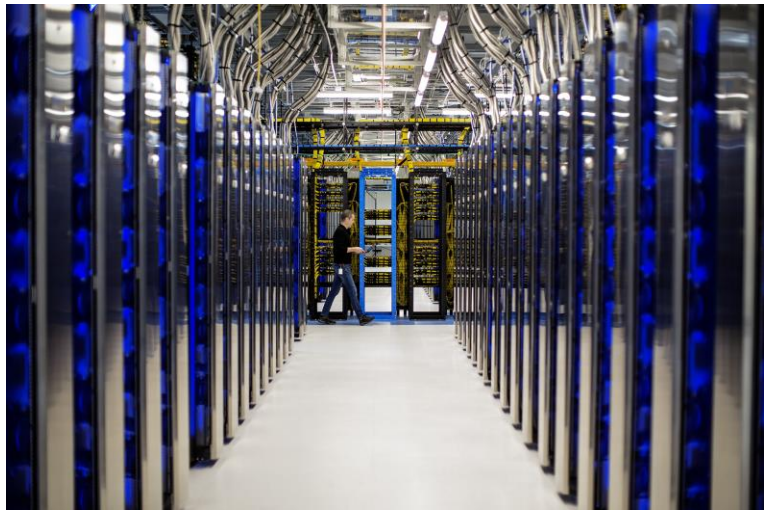
From the Intelligent Cloud to the Intelligent Edge

What is automated machine learning?

Automated machine learning (automated ML) automates feature engineering, algorithm and hyperparameter selection to find the best model for your data.

Automated ML Mission

Enable automated building of machine learning with the goal of accelerating, democratizing and scaling AI



Democratize AI

Enable Domain Experts & Developers to get rapidly build AI solutions

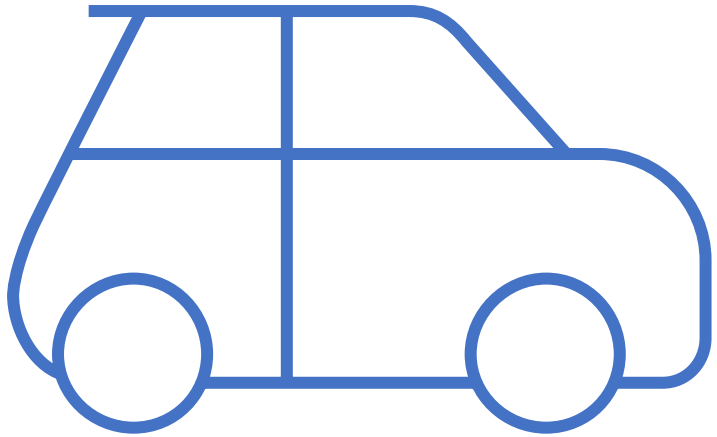
Accelerate AI

Improve Productivity for Data Scientists, Citizen Data Scientists, App Developers & Analysts

Scale AI

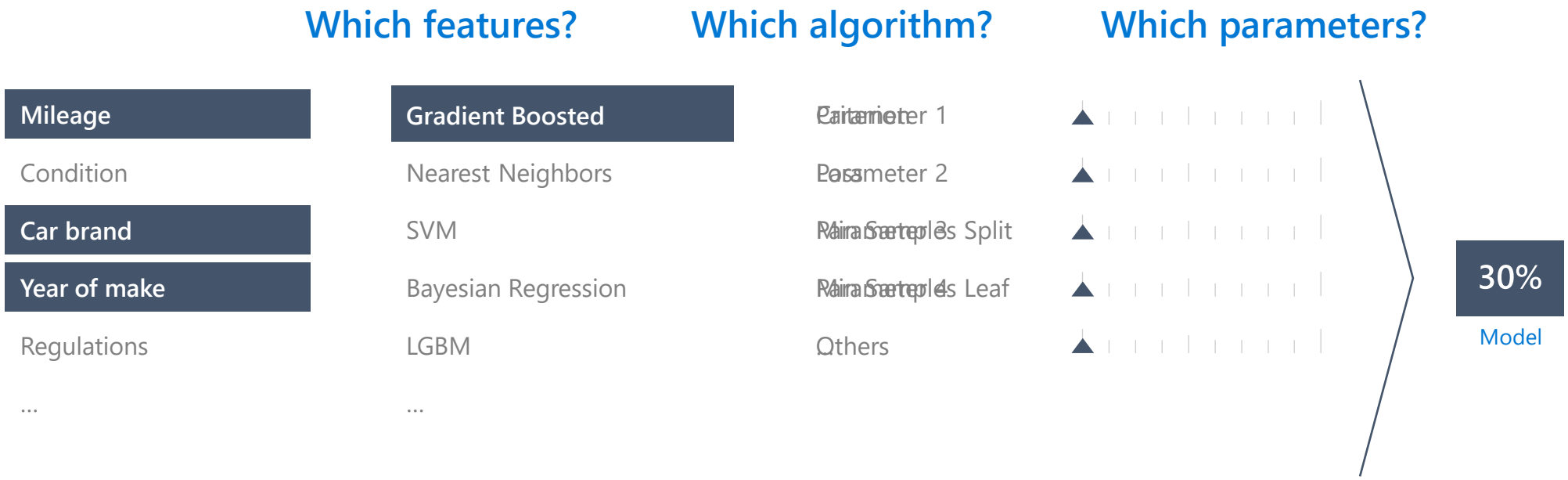
Build AI solutions at scale in an automated fashion

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?

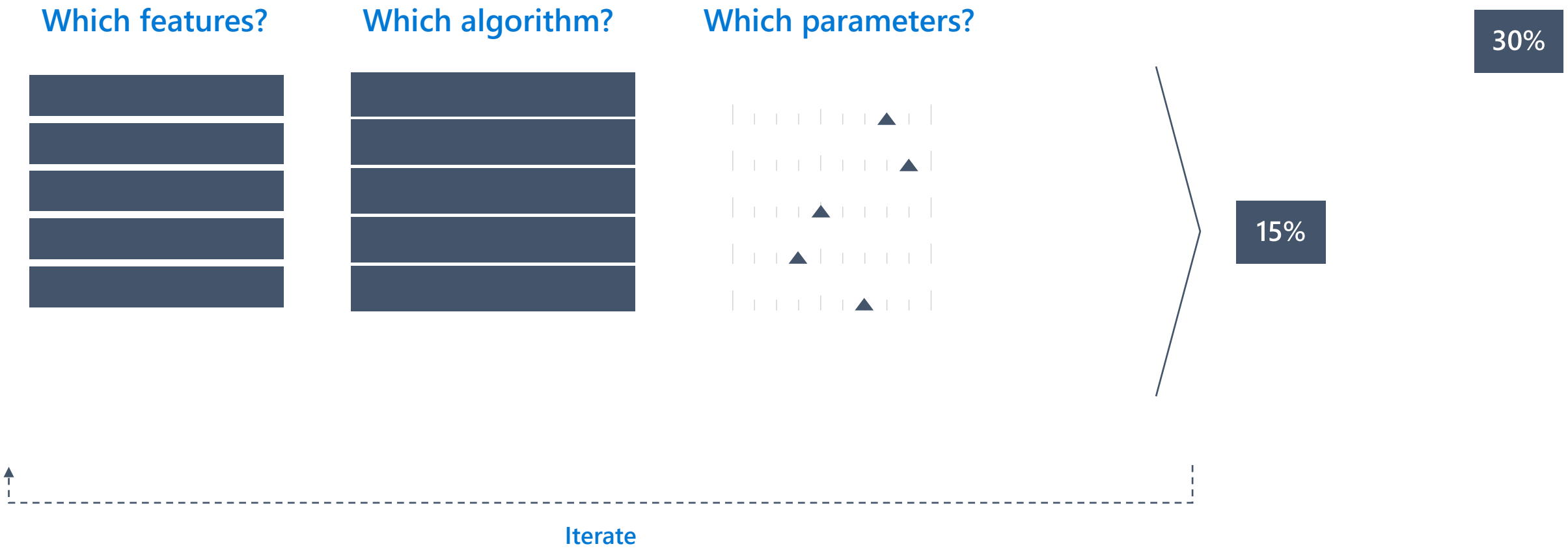
- Criterion
- Weights
- Min Samples Split
- Min Samples Leaf
- Others

30%

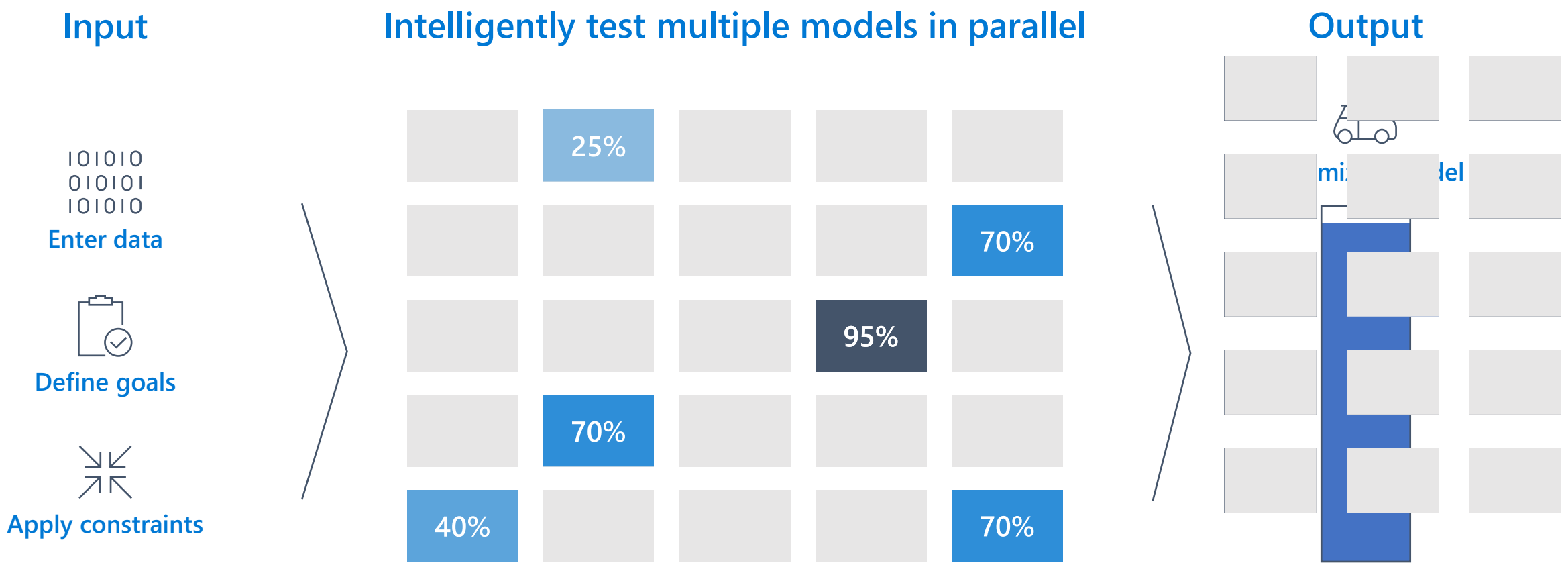
Model

Iterate

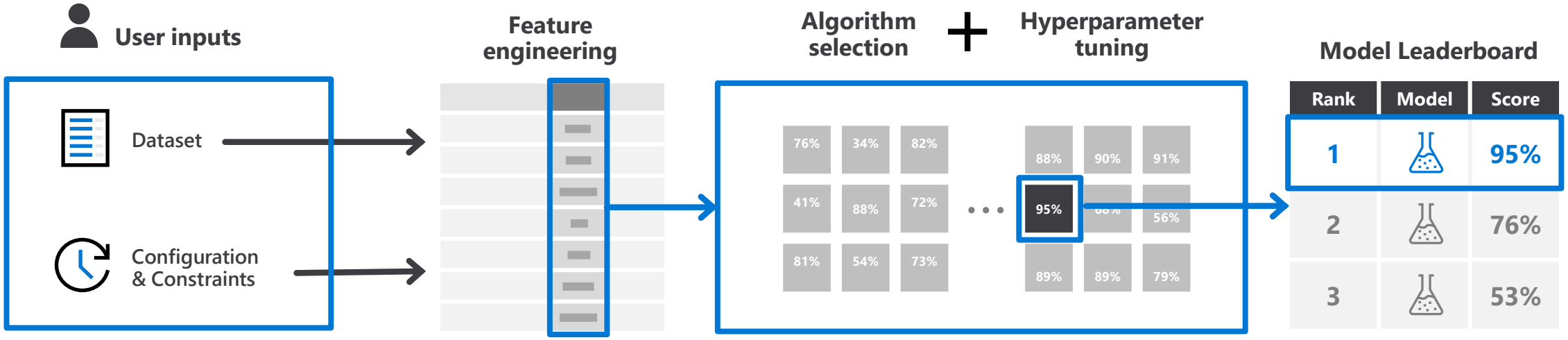
Model Creation Is Typically Time-Consuming



Automated ML Accelerates Model Development



Automated Machine Learning Under the Hood



Automated Machine Learning Capabilities

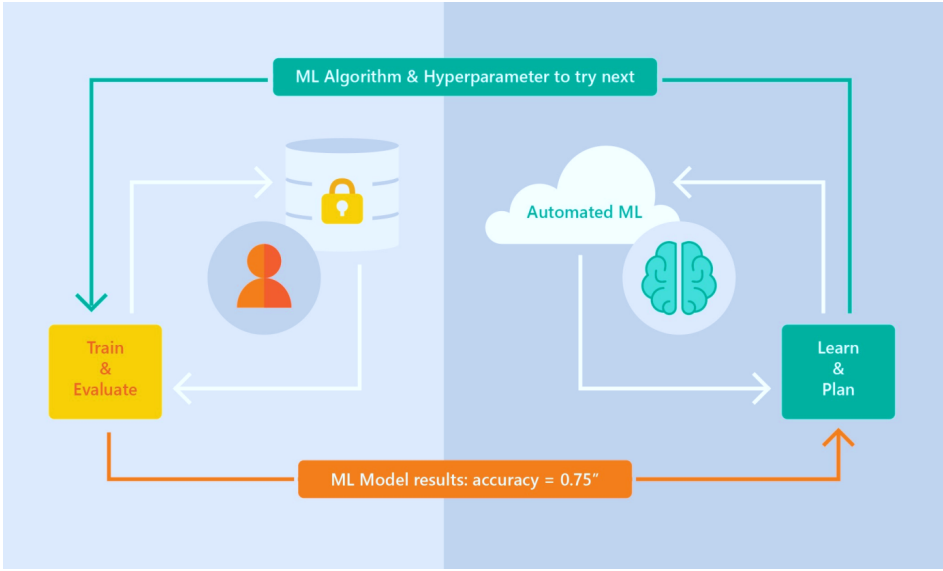
Based on Microsoft Research

Brain trained with several million experiments

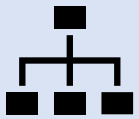
Personalized recommendation approach:

Collaborative filtering and Bayesian optimization

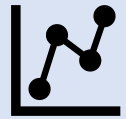
Privacy preserving: No need to “see” the data



Supervised Learning



Classification

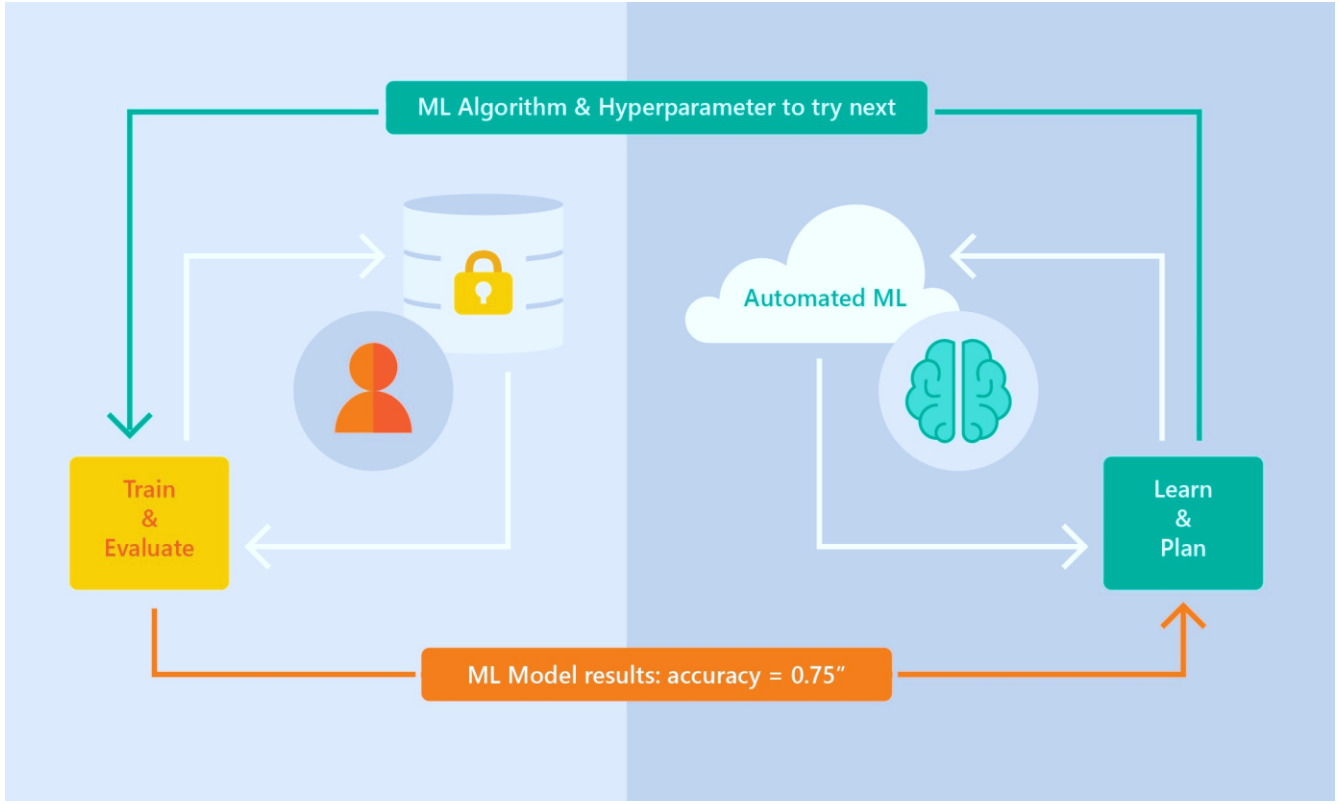


Regression

Time Series Forecasting

Automated ML Capabilities

- ML Scenarios: Classification & Regression, Forecasting
- Languages: Python SDK for deployment and hosting for inference – Jupyter notebooks
- Training Compute: Local Machine, AML Compute, Data Science Virtual Machine (DSVM), Azure Databricks*
- Transparency: View run history, model metrics, explainability*
- Scale: Faster model training using multiple cores and parallel experiments



* In Preview

Guardrails: Detection and auto-correction of data issues

Class imbalance

Train-Test split, CV, rolling CV

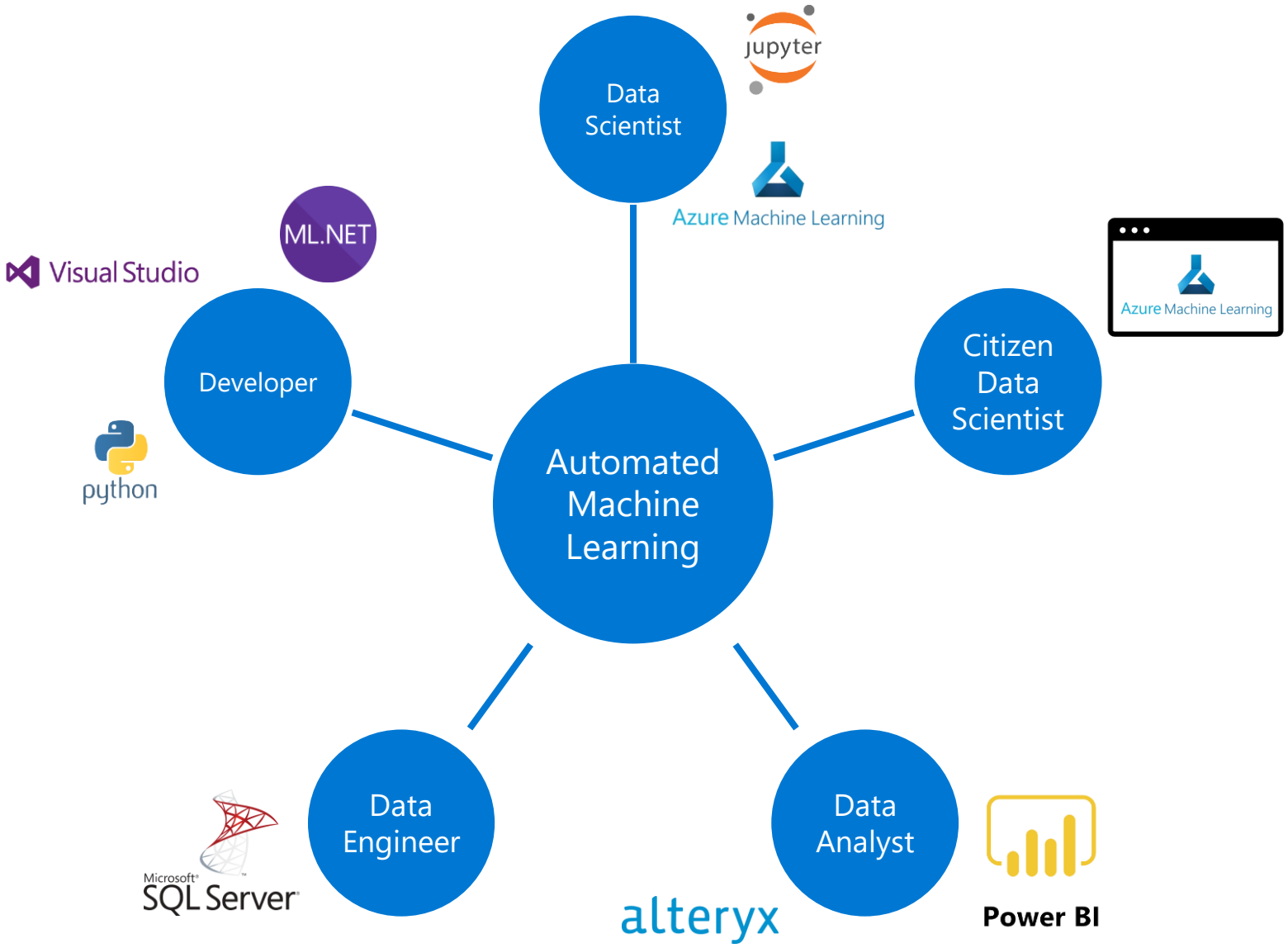
Missing value imputation

Detect high cardinality features

Detect leaky features

Detect overfitting

Machine Learning for Everyone



Customer Testimonials

With one line of code, it runs through different algorithms within the prediction family and the different parameter (or variable) combos that previously were manually tested by the scientists. The power of the cloud comes in here. The results are comparable to what the data scientists produced.

Manish Naik, BP, Digital Innovation

Auto ML's execution of different models was an impressive that enabled data scientists to work iteratively on machine learning experiments to increase auction sales by 10% and optimize the time auction cars are kept in the showroom to less than 30 days.

Farika Maharani, PT. Serasi Autoraya, Data Platform Supervisor

The CBRE AI and Data Engineering Team have successfully deployed a complete Azure Machine Learning model to their new API gateway leveraging the Azure AutoML solution in Azure Databricks. The API Gateway plus the model deployment goes into production this March.

Francis Dogbey, Microsoft CSA

In evaluating Azure Automated ML we discovered real potential in shortening the time to market for producing predictive models. The availability of the Automated ML UI also holds the promise of opening the ML space to non data science trained resources which in turn allows the democratizing of the predictive work without the pain of hiring expensive/ hard to retain staff.

Bogdan Rosca, Senior Director, Principal Information Architect

We see advantages moving over to Azure AutoML because we think we will be able to increase our speed to create models significantly and do more with less in terms of labor hours.

Dan Metzendorf, Data Science Manager, The Sherwin Williams Company

AutoML resulted in a significant improvement in model performance (1) Consistently produced better models than other automated ml libraries (TPOT) (2) Also outperformed hand-tuned models. AutoML explored a solution space larger than what was plausible to do manually

David Robinson, Devon Energy Data Scientist

The reason we see the sharp uplift in sales is the customers are getting content that really connects with them, and they're getting offers for things that are truly relevant and relevant at that moment in time.... Microsoft—they are really wanting to be our partners and were really going to help us on this journey, which was very differentiating

Daniel Humble, Chief Data and Analytics Officer, Walgreens Boots Alliance

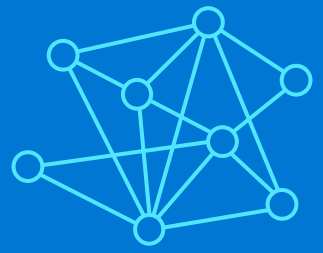
If I have 200 models to train—I can just do this all at once. It can be farmed out to a huge computer cluster, and it can be done in minutes so I'm not waiting for days or setting experiments to run over the weekend anymore.

Dean Riddlesden, Senior Data Scientist, Walgreens Boots Alliance

We tried AutoML for aspect ratio model and pleased to see AutoML produced much better model than our baseline. We need to build almost 50 models and are looking forward to the productivity boost we will get by not hand tuning each one of them!

Saurabh Naik, Sr. Software Engineer

Teams uses machine learning to analyze, gain insights and improve the quality of calls. We use AutoML to significantly scale up the application of ML solutions by semi-automating model train tasks

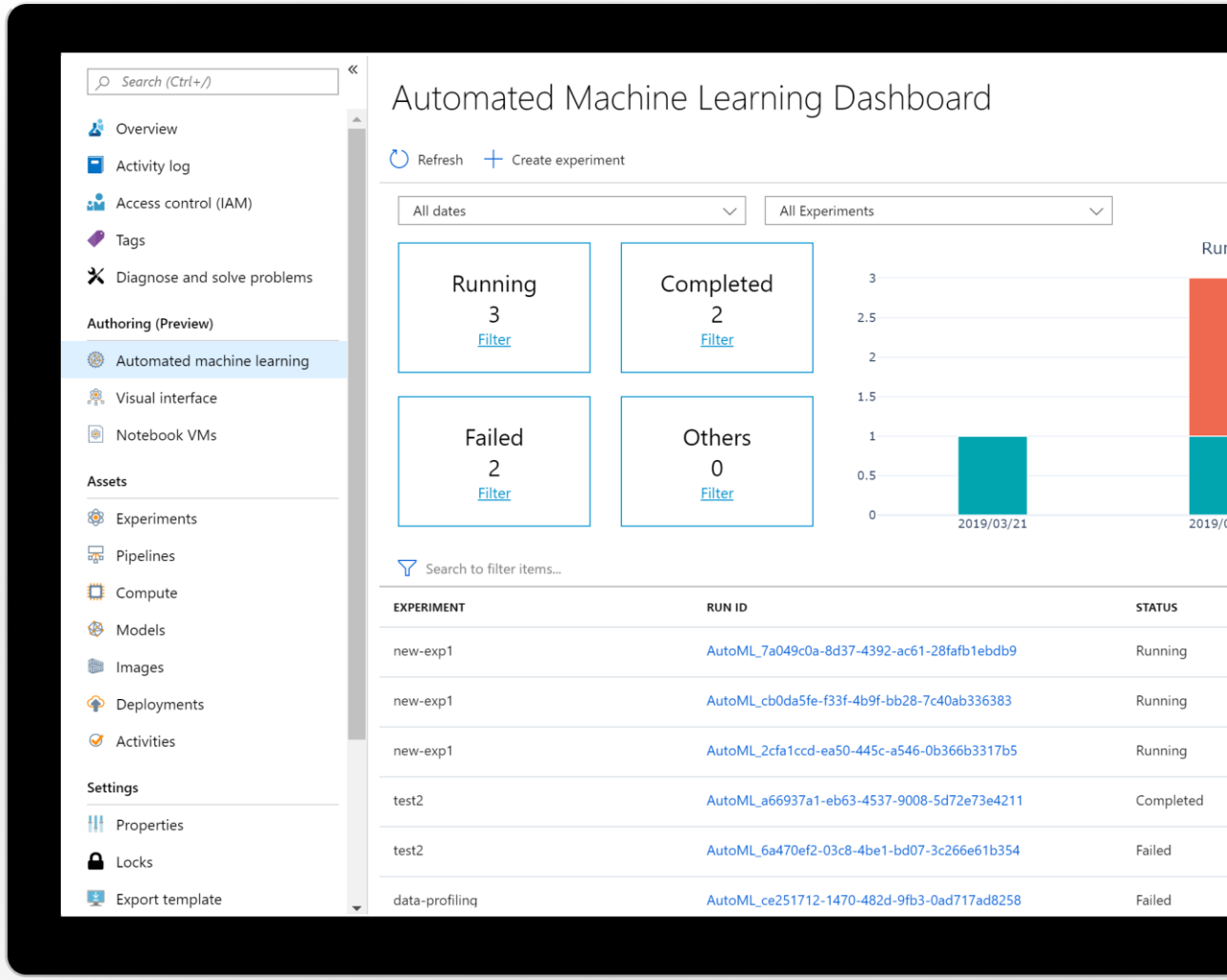


What's new?

Latest announcements (Blog post with all the announcements)

Automated ML Updates

- Link
- Managed Endpoint



Workshop

Prerequisites
1. Azure Subscription
2. Azure ML workspace

Workshop Lab Instructions

1. Go to Medium location: OpenHack
2. Git hub link

Agenda:

Introduction
Azure Machine learning Intro — 1 hour
OpenHack use case introduction — 15 mins
OpenHack – 4 Hours
Take Break when experimentation starts
Deploy model — 1 hour
Clean up — 15 minutes

Note: When experiment is submitted will take a long time to run since automated machine learning runs about 50 or more algorithm to provide us results.
Deploying to REST API as ACI or AKS will also take about 15 to 30 minutes. Some time delete and recreate will be needed but as the instructor for details

Resources

http://aka.ms/amlfree

Learn more : https://aka.ms/automatedmldocs
Notebook Samples : https://aka.ms/automatedmlsamples
Blog Post : https://aka.ms/AutomatedML
Product Feedback : AskAutomatedML@microsoft.com
Workplace Safety:
https://github.com/balakreshnan/WorkplaceSafety
Covid 19 Social Distancing:
sdd/covidproject.md at master · balakreshnan/sdd (github.com)
AI and Big Data:
AzurePercept/synapseint.md at main · balakreshnan/AzurePercept (github.com)

Microsoft

Resources

PPE Compliance Detection using Artificial Intelligence in Learning Factories

A Collaboratively Developed Platform to Introduce Fundamentals of IoT and IIoT

Microsoft


Microsoft
Thank You