Microsoft

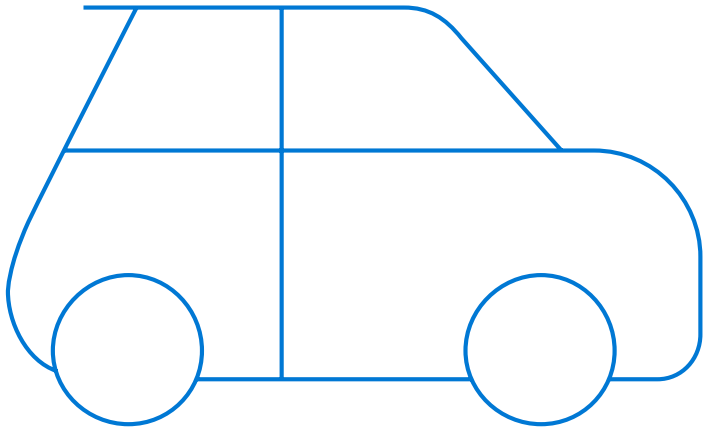
Automatically Generate ML Models with Automated ML in the Azure Machine Learning Service

Heather Grandy

Technical Specialist, Azure Data & AI

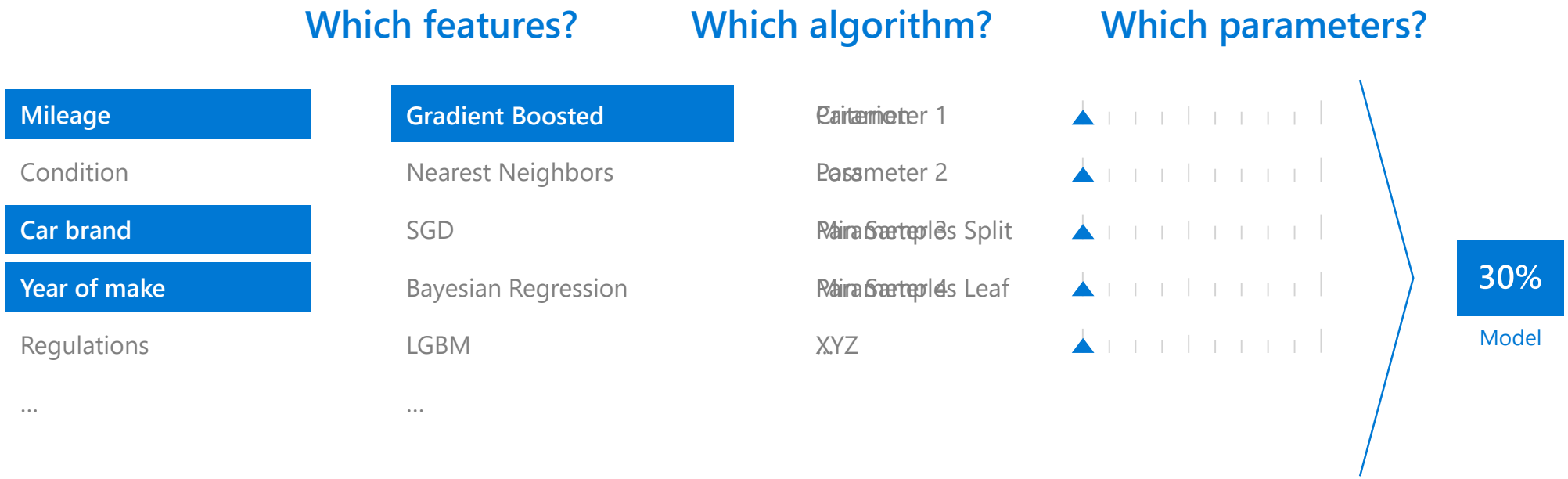
Azure Machine Learning

Automated Machine Learning



How much is this car worth?

Model Creation: A Time-Consuming Process



Model Creation: A Time-Consuming Process

Which features?

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

Which algorithm?

- Gradient Boosted
- Nearest Neighbors
- SGD
- Bayesian Regression
- LGBM
- ...

Which parameters?

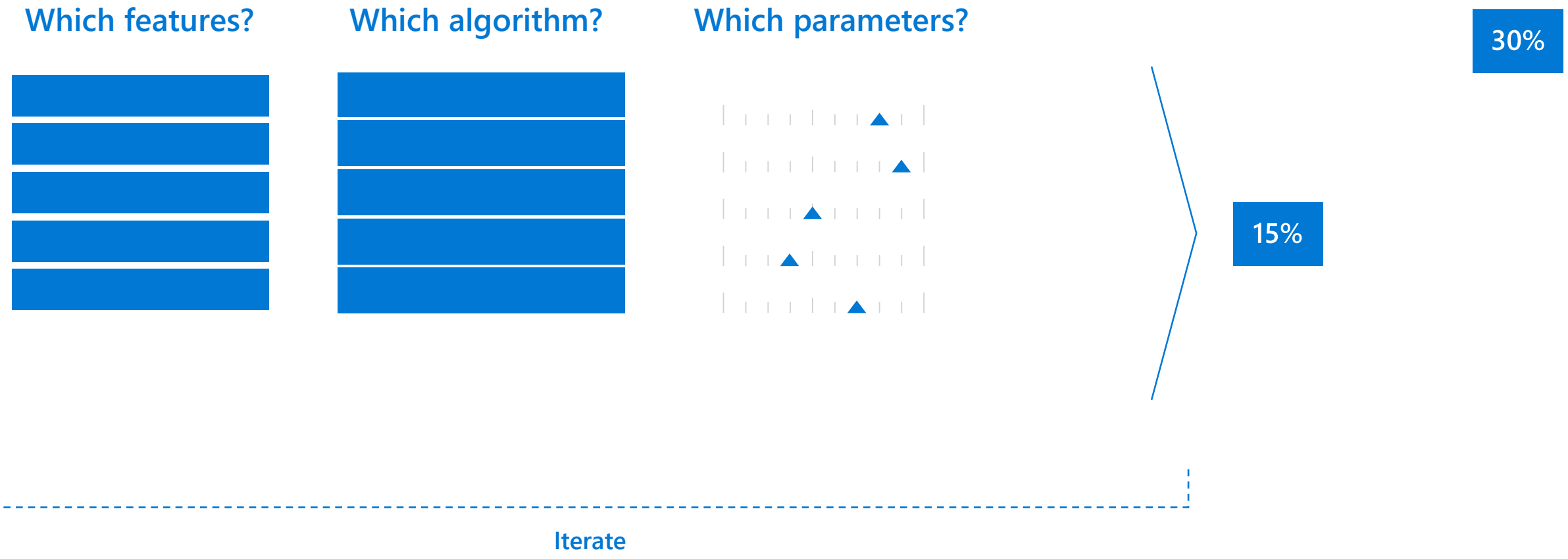
- Criterion
- Weights
- Min Samples Split
- Min Samples Leaf
- XYZ

30%

Model

Iterate

Model Creation: A Time-Consuming Process



Azure Machine Learning Accelerates Model Development

with Automated Machine Learning

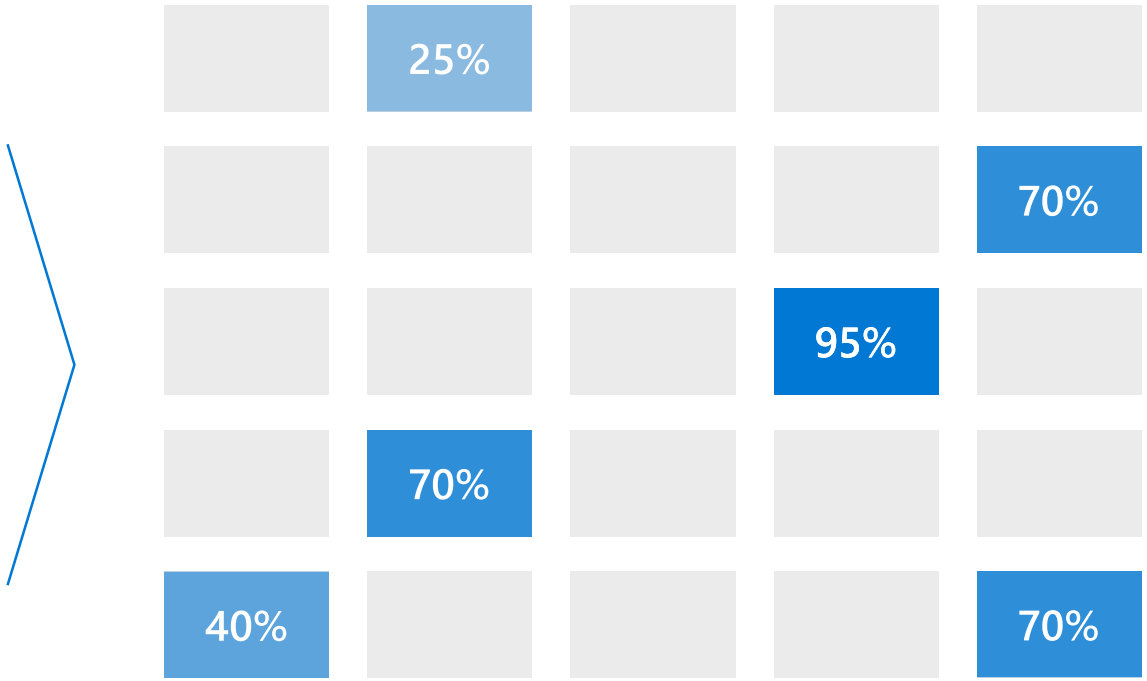
Input

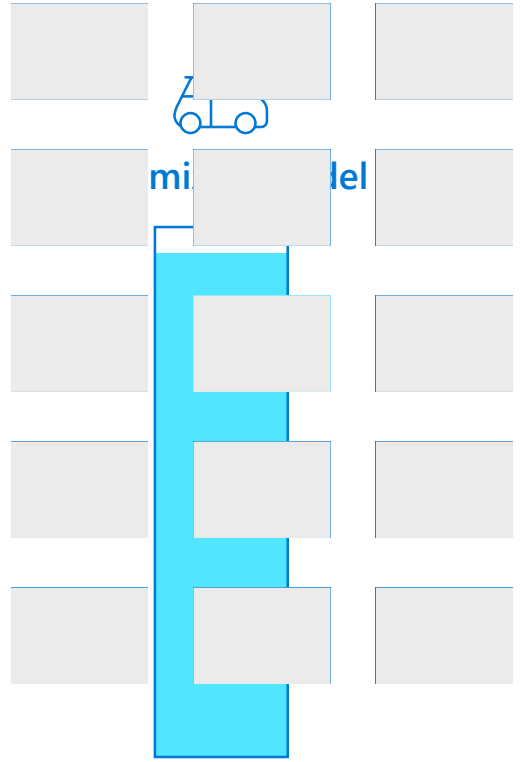
Enter data

Define goals

Apply constraints

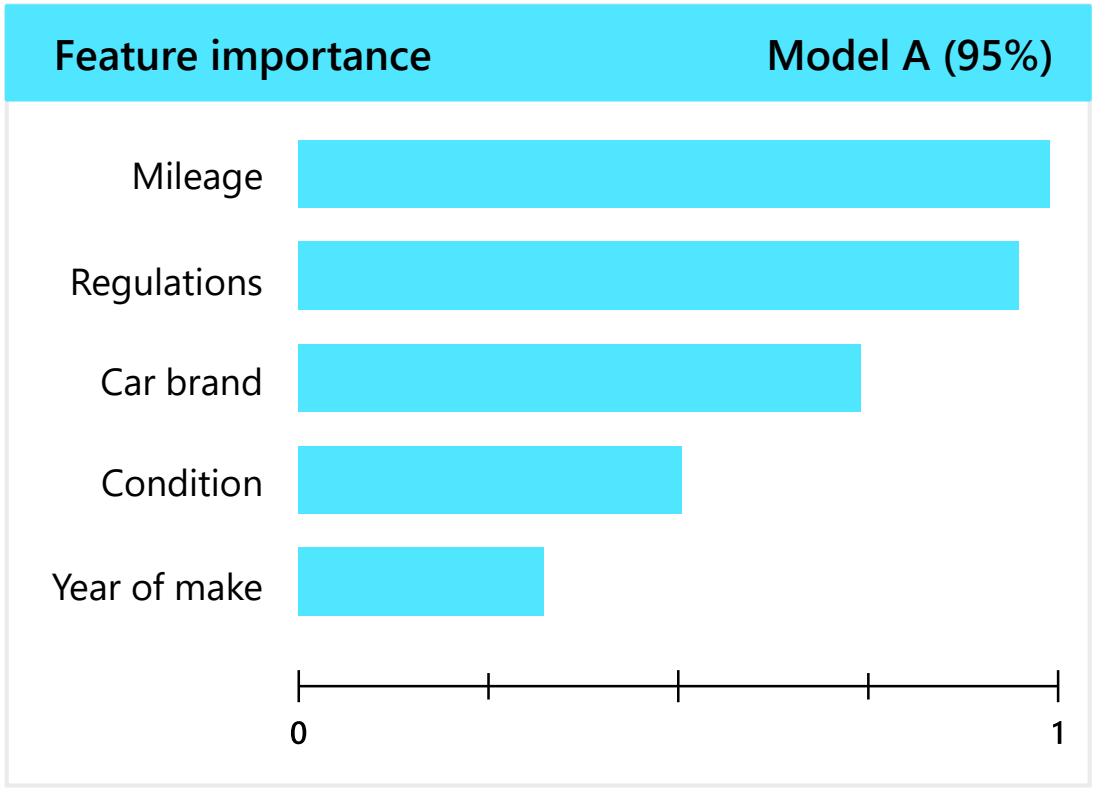
Intelligently test multiple models in parallel

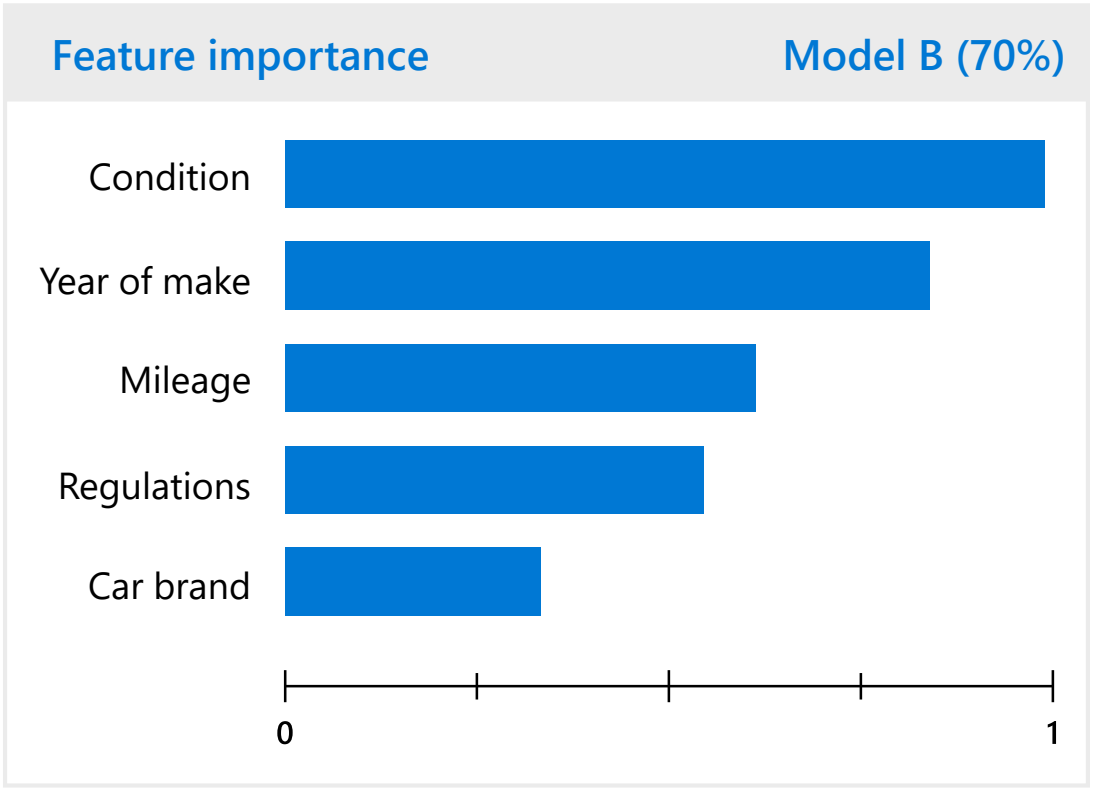




Azure Machine Learning Accelerates Model Selection

with Model Explainability





Microsoft

Example: Movie Recommender

AI to build AI: Movie recommendation

Viewers like you watch *Avengers Endgame*.

AI to build AI: Movie recommendation

Datasets ~~Viewers~~ like ~~you watch *Avengers Endgame*~~.

yours perform well with X ML learner

AI to build AI: Movie recommendation

Datasets ~~Viewers~~ like ~~you watch *Avengers Endgame*~~.

yours perform well with X ML learner

Probabilistic Matrix Factorization for Automated Machine Learning

Nicolo Fusi [1] Rishit Sheth [1 2] Melih Huseyn Elibol [1]

Abstract

In order to achieve state-of-the-art performance, modern machine learning techniques require careful data pre-processing and hyperparameter tuning. Moreover, given the ever increasing number of machine learning models being developed, model selection is becoming increasingly important. Automating the selection and tuning of machine learning pipelines, consisting of data preprocessing methods and machine learning models, has long been one of the goals of the machine learning community. In this paper, we propose to solve this meta-learning task by combining ideas from collaborative filtering and Bayesian optimization. Specifically, we exploit experiments performed on hundreds of different datasets via probabilistic matrix factorization and then use an acquisition function to guide the exploration of the space of possible pipelines. In our experiments, we show that our approach quickly identifies high-performing pipelines across a wide range

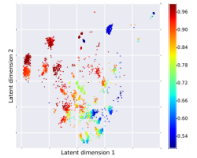
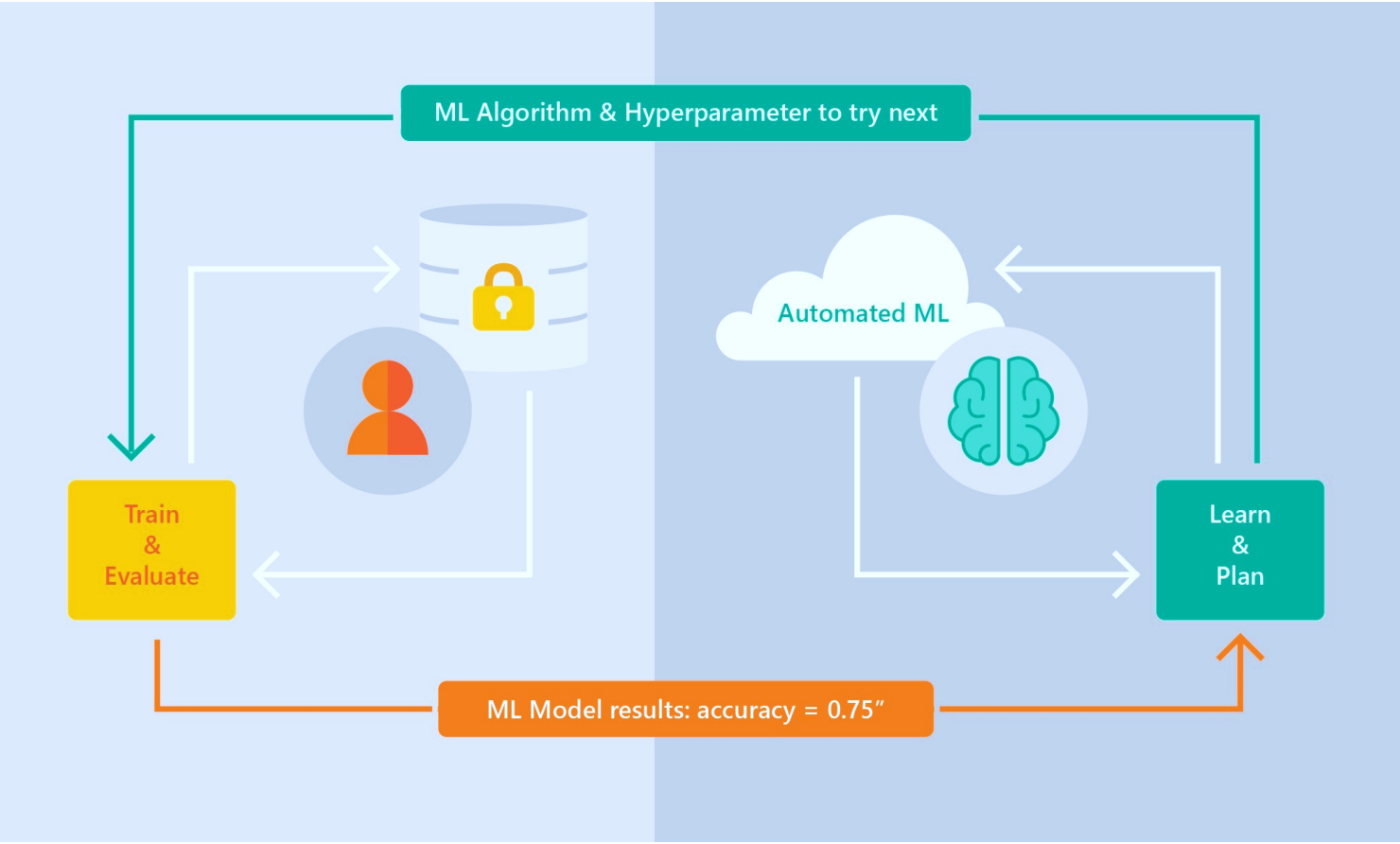


Figure 1: Two-dimensional embedding of 5,000 ML pipelines across 576 OpenML datasets. Each point corresponds to a pipeline and is colored by the AUROC obtained by that pipeline in one of the OpenML datasets (OpenML dataset id 943).

https://arxiv.org/abs/1705.05355

ML Learner Recommendation Process



Automated ML recommender never "sees" your data

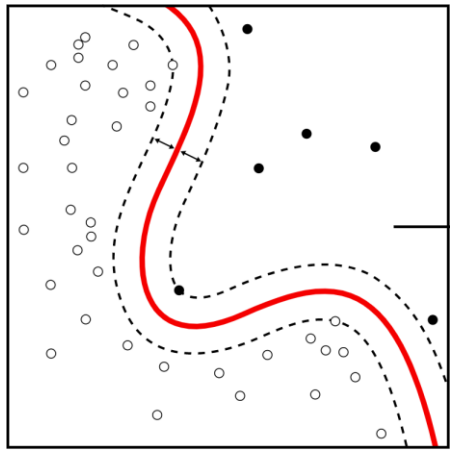
Example: Sales Funnel Optimization

Product

Training Data

user_id	user_device	user_os	user_age	zipcode	last_message	time	target
23433	ios	ios_11	30	92505	what ligtweight options do you have?	2018-12-15	13.5in
5223423	android	android_7	65	75240	How do increase default fntsize	2019-01-15	15in
343433	android	android_9		98004		2018-08-01	15in

AutoML



- Auto-featurizes your data
- End-to-end training on Azure
- Security: AutoML never "sees" your data

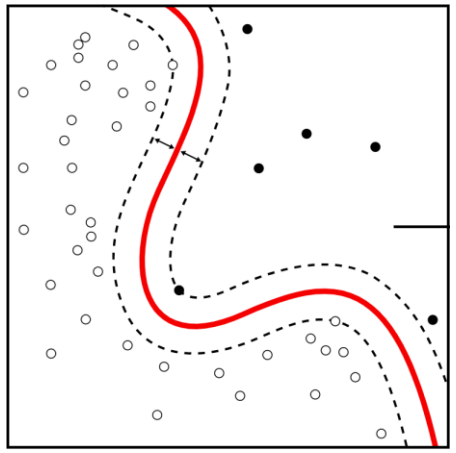
Example: Sales Funnel Optimization

Product

Training Data

user_id	user_device	user_os	user_age	zipcode	last_message	time	target
23433	ios	ios_11	30	92505	what ligtweight options do you have?	2018-12-15	13.5in
5223423	android	android_7	65	75240	How do increase default fntsize	2019-01-15	15in
343433	android	android_9		98004		2018-08-01	15in

AutoML



- Auto-featurizes your data
- End-to-end training on Azure
- Security: AutoML never "sees" your data

Deployment

Deployment Options

- Microservice deployment (AKS)
- IOT Edge
- Different run-time: ONNX
  - How do you deploy inferencing/prediction code in different run times?
  - With ONNX you can
    - scikit-learn/tensorflow trained model → prediction code directly in C# or Java applications
    - AutoML can generate ONNX formatted models

Alibaba Group 阿里巴巴集团 · AMD · arm · aws · Baidu 百度 · BITMAIN · CEVA · Facebook Open Source · habana · Hewlett Packard Enterprise · HUAWEI · IBM · intel AI · MathWorks · MEDIATEK · Microsoft · NVIDIA · NXP · Oath · Preferred Networks · QUALCOMM · sas · skymizer · SYNOPSYS · Tencent · mi · unity

Microsoft

Democratizing AutoML

Democratization - AutoML Integrations

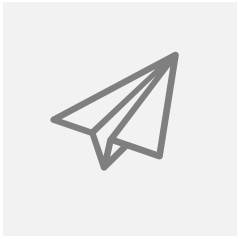
Demo

AutoML – Azure Portal Experience

AutoML with Azure Machine Learning

Democratize AI

Enable domain experts & developers to rapidly build AI solutions



Accelerate AI

Improve productivity for Data scientists

Scale AI

Build AI solutions at scale in an automated fashion

AutoML Resources

- Automated ML Blog: https://aka.ms/AutomatedML
- Automated ML Docs: https://aka.ms/AutomatedMLDocs
- Questions: askautomatedml@microsoft.com
- Demo Notebooks: https://github.com/buildaidemos

Lab Setup

AutoML in Azure Databricks

Microsoft

Thank you!