



NT Conference 2018

Andrey Vykhodtsev,
Cloud Solution Architect
Central & Eastern Europe



#ntk18



Azure Machine Learning Services: Real life examples



What is Machine Learning ?

Using **known data**, develop a **model** to predict **unknown data**.


Known Data: Big enough archive, previous observations, past data

Unknown Data: Missing, Unseen, not existing, future data

Model: Known data + Algorithms (ML algorithms)

Why Machine Learning?

Is this where you work? If so, I can give you better directions and suggestions.



Yes

No

How'd you guess?

ask me anything

Here's the forecast.

New York, NY

68°

TUE 5/22
Low 58°
Overcast

30%

70°

3:00 PM
Cloudy

20%

71°

4:00 PM
Cloudy

10%

73°

5:00 PM
Partly Cloudy

10%

73°

6:00 PM
Partly Cloudy

10%

71°

7:00 PM
Partly Cloudy

20%

68°

8:00 PM
Partly Cloudy

20%

Data from iMap Weather

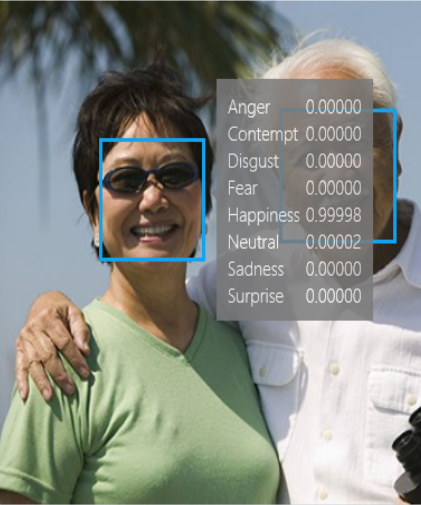
Search for 'what's the weather today'

try Will it be warm next Tuesday?

Emotion Recognition

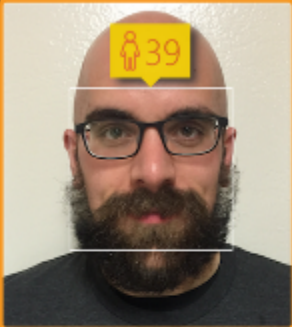
Identify emotions communicated by the facial expressions in techniques to provide these results. You can also click the op

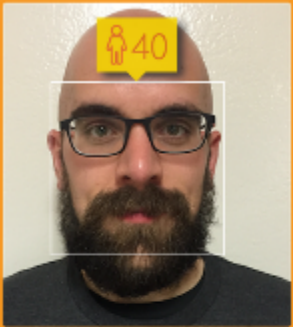
- Image resolution $\geq 36 \times 36$ pixels and the file size $< 4M$
- The frontal and near-frontal faces have the best results
- Recognition is experimental, and not always accurate.

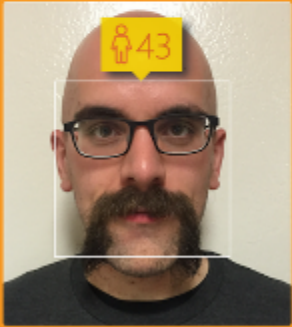



Anger	0.00000
Contempt	0.00000
Disgust	0.00000
Fear	0.00000
Happiness	0.99998
Neutral	0.00002
Sadness	0.00000
Surprise	0.00000

Image URL









Mark Cerqueira @mcerqueira

May 6

How the #HowOldRobot handled my diminishing beard: no facial hair - youth, mustache - old, ha

清炒娃娃菜

Stir fried baby vegetables

清炒油麦菜

Stir fried lettuce

红烧茄子

Braised Eggplant

老厨白菜

Old kitchen Chinese cabbage

西红柿炒鸡蛋

Scrambled egg with tomato

汤

Tap to pause.



getting a morgage in seattle



8,140,000 RESULTS Any time ▾

Ads related to getting a morgage in seattle

[15-Year Mortgage Rates | QuickenLoans.com](#)

www.QuickenLoans.com/Rates

Lock Your Rate. 3.500% (3.92% APR) With America's #1 Online Lender.

[Lendi](#)

[Lending](#)

APR fr

[TILA](#)

[seattle](#)

Meet o

[Pre Q](#)

www.wellstargo.com/mortgage

Estimate how much you can afford

Including results for [getting a mortgage in seattle](#).

Do you want results only for [getting a mortgage in seattle](#)?

[Seattle Mortgage Rates - Find the Best Home Loan | Zillow](#)

www.zillow.com/mortgage-rates/wa/seattle ▾

See up to the minute **Seattle mortgage** rates and find **Seattle** Washington's best, lowest possible quote with Zillow **Mortgage Marketplace**.

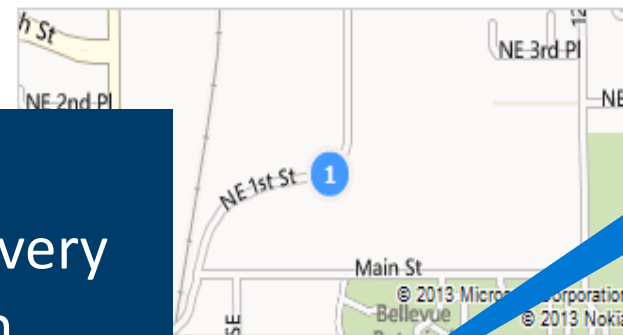
[Seattle's Best Mortgage](#)

www.seattlesbm.com ▾

Get the best mortgage loan for you at **Seattle's Best Mortgage**. (CL#117721) When you decide to buy a home or refinance a **mortgage**, it's a big step.

1 11911 Ne 1st St Ste B306, Bellevue · (425) 228-7000 · [Directions](#) · Bing Local

Seattle's Best Mortgage Inc



St Ste B306 · Bellevue
0 · [Directions](#)

Website: seattlesbm.com

[Report a problem](#)

RELATED SEARCHES

[Getting a First Mortgage](#)

[Getting a Mortgage Self-Employed](#)

[Getting a Mortgage Loan Approved](#)

[Getting a Mortgage On Land](#)

[Getting a Mortgage in 2013](#)

[How to Get a Mortgage License](#)

[How to Get a Mortgage After Bankruptcy](#)

[Mortgage Calculator](#)

Ads related to getting a morgage in seattle

What language?

Which ads to show, and in what order?

Misspelled?

Which links are most likely to get clicked?

What is the probability of a click on each ad?

What is the intent?

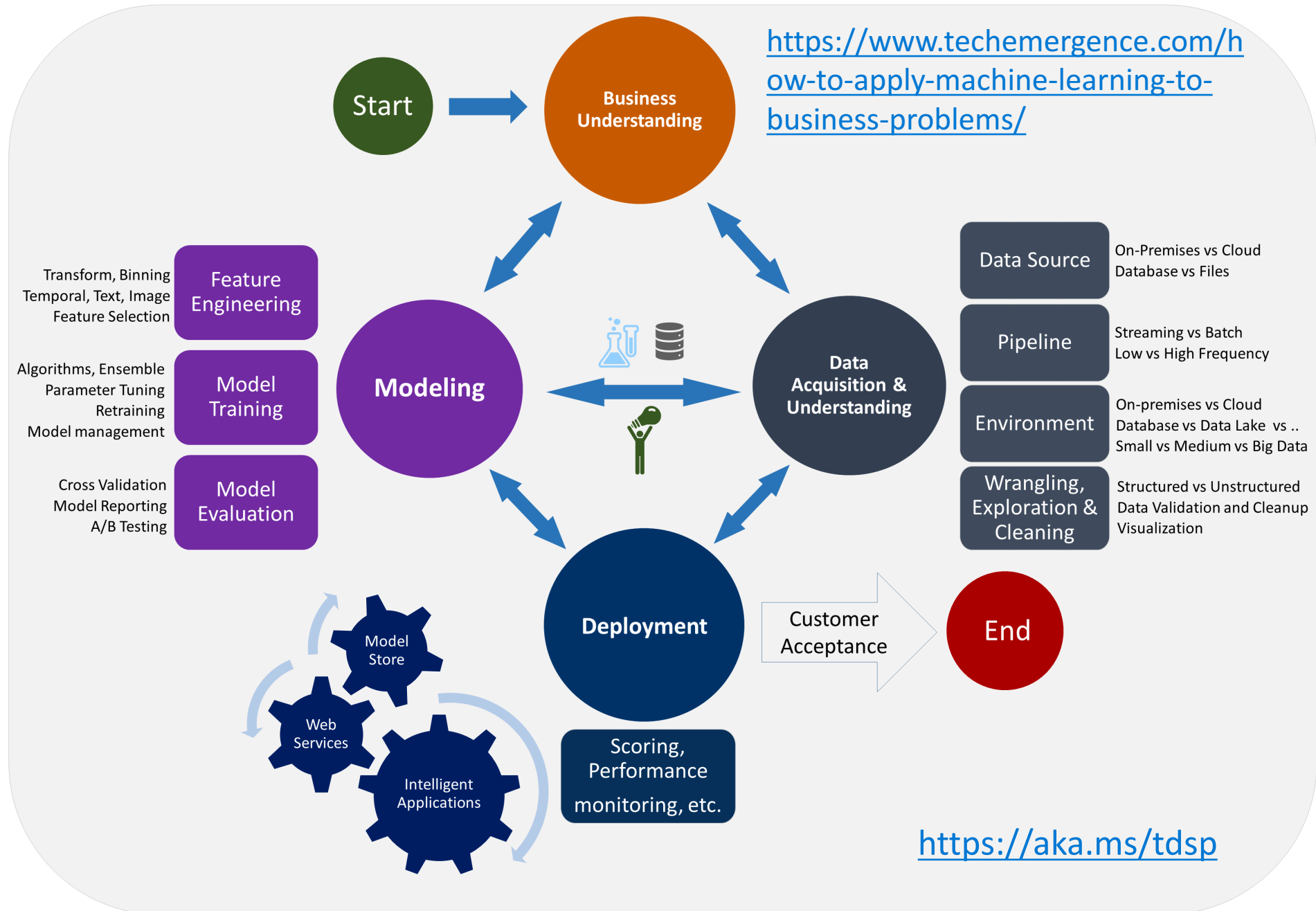
Are any of these pages malicious?

What pages should we index?

What ad pricing will optimize revenue?

Machine learning enables nearly every value proposition of web search.

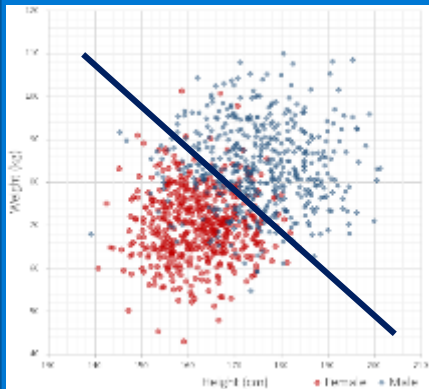
Data Science Lifecycle



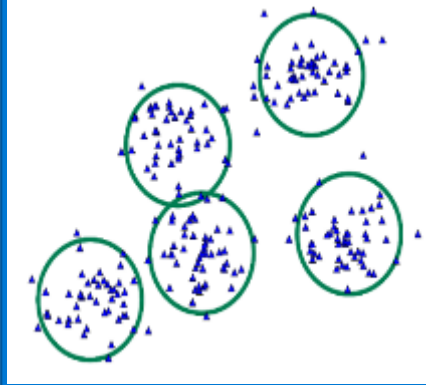
Common Classes of Algorithms

(Supervised|Unsupervised)

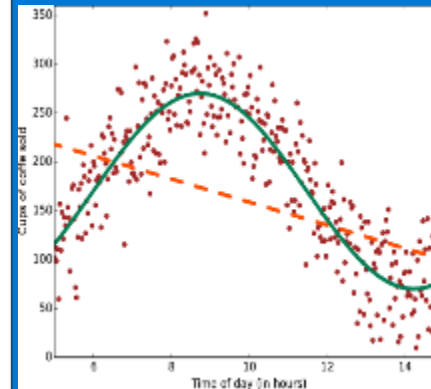
Classification



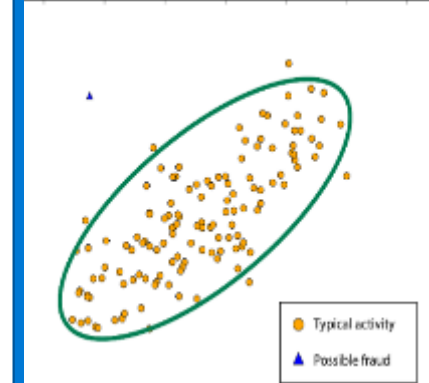
Clustering



Regression

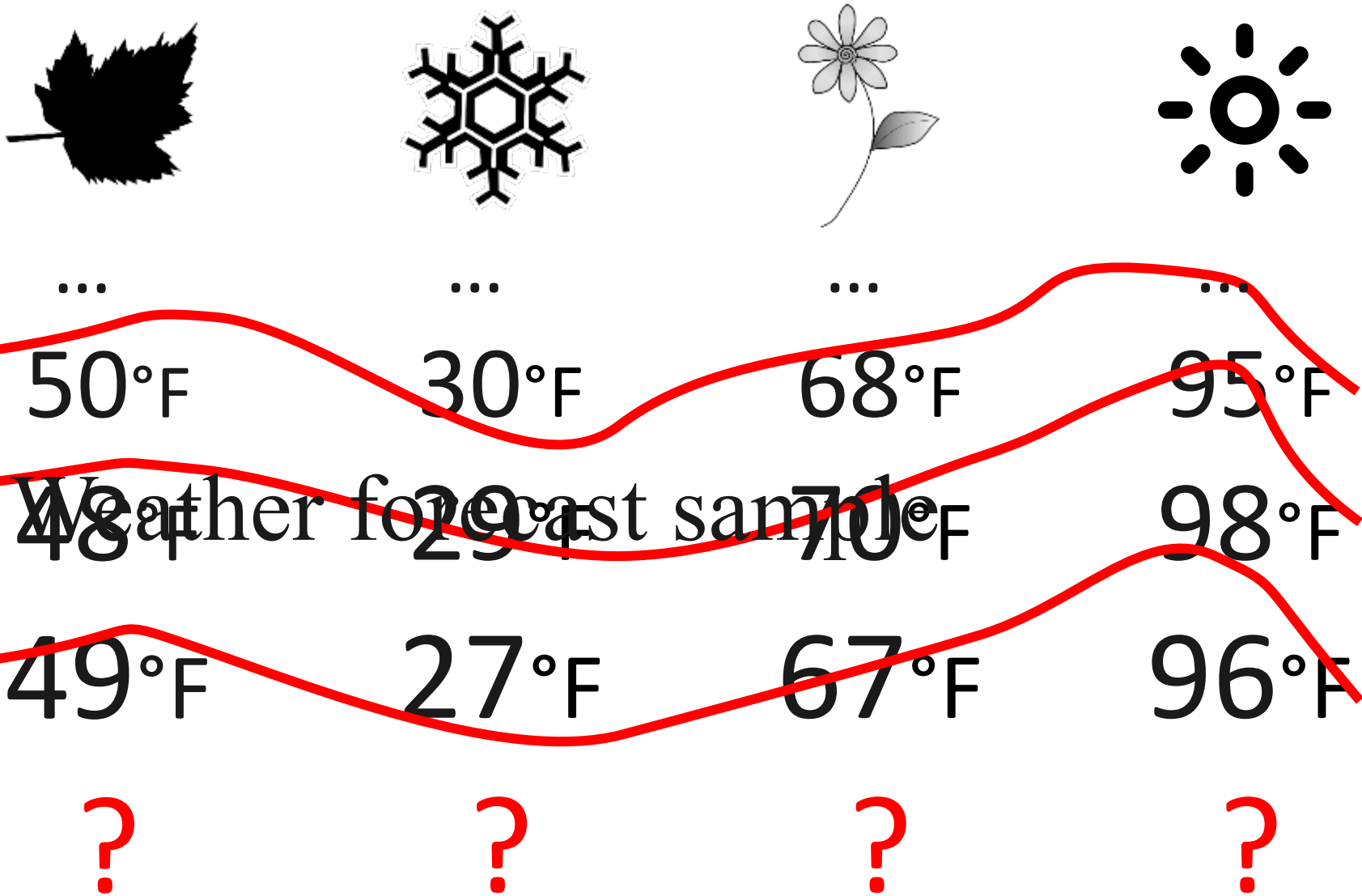


Anomaly Detection



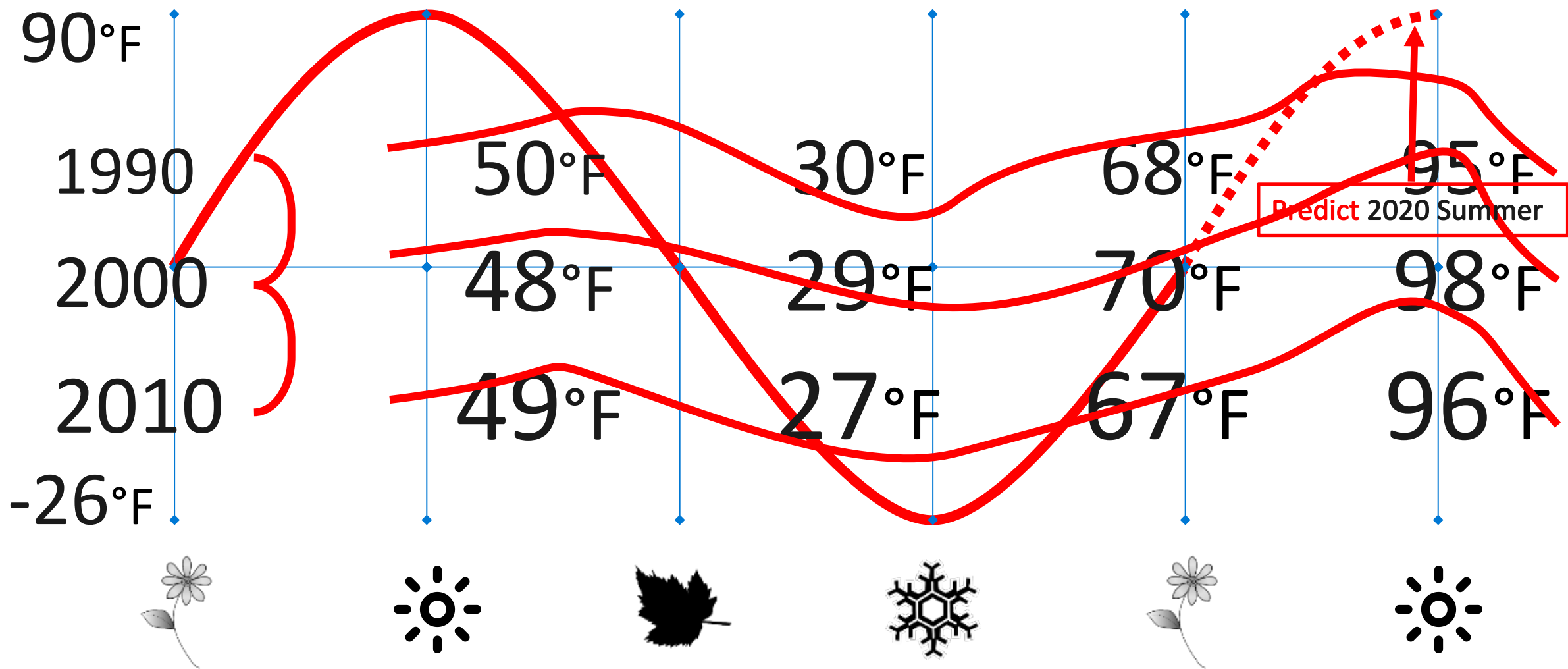
EXAMPLE

Known data
Model
Unknown data



Using **known data**, develop a **model** to predict **unknown data**.

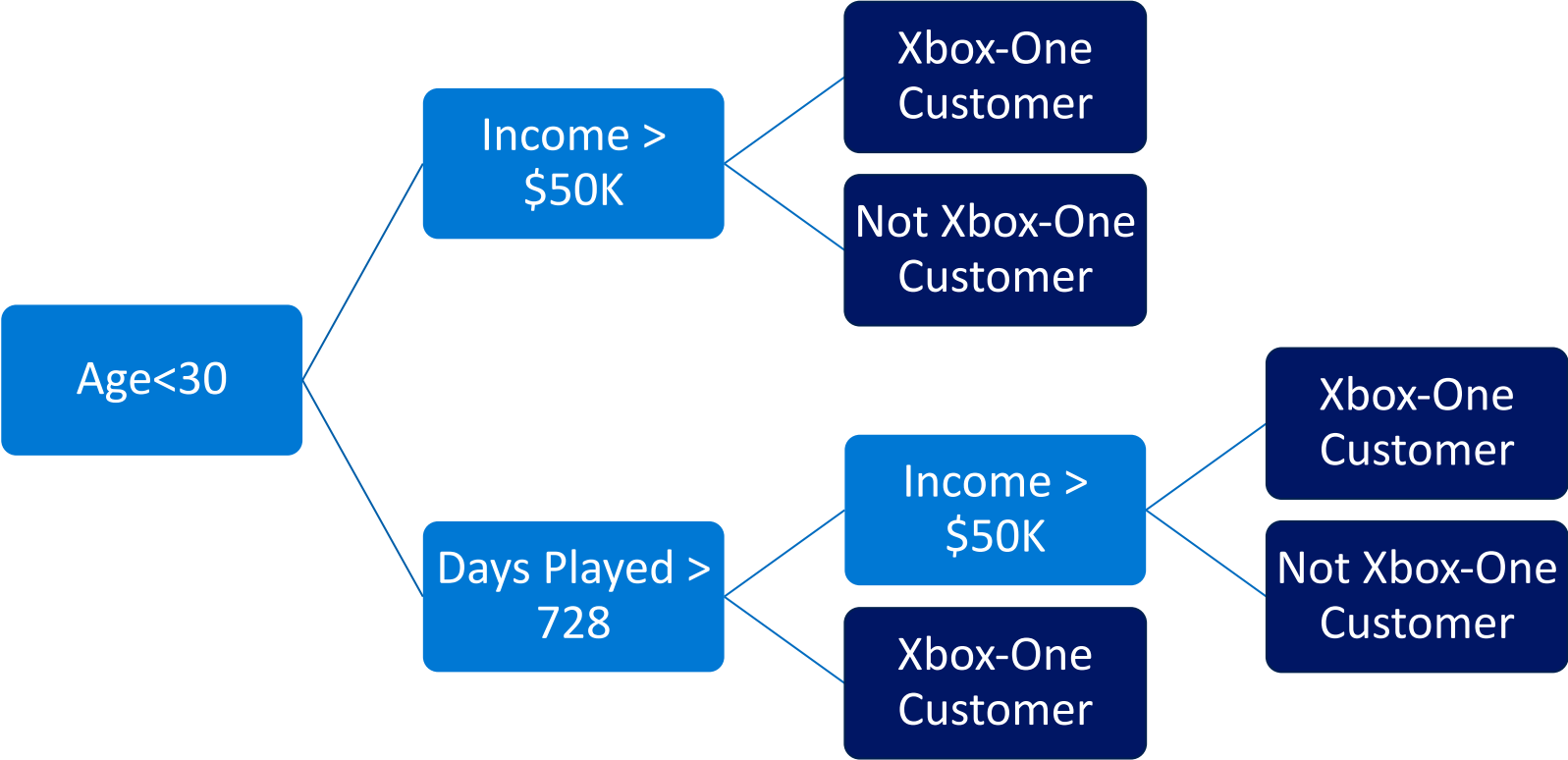
Model (Regression)



Using known data, develop a model to predict unknown data.

EXAMPLE

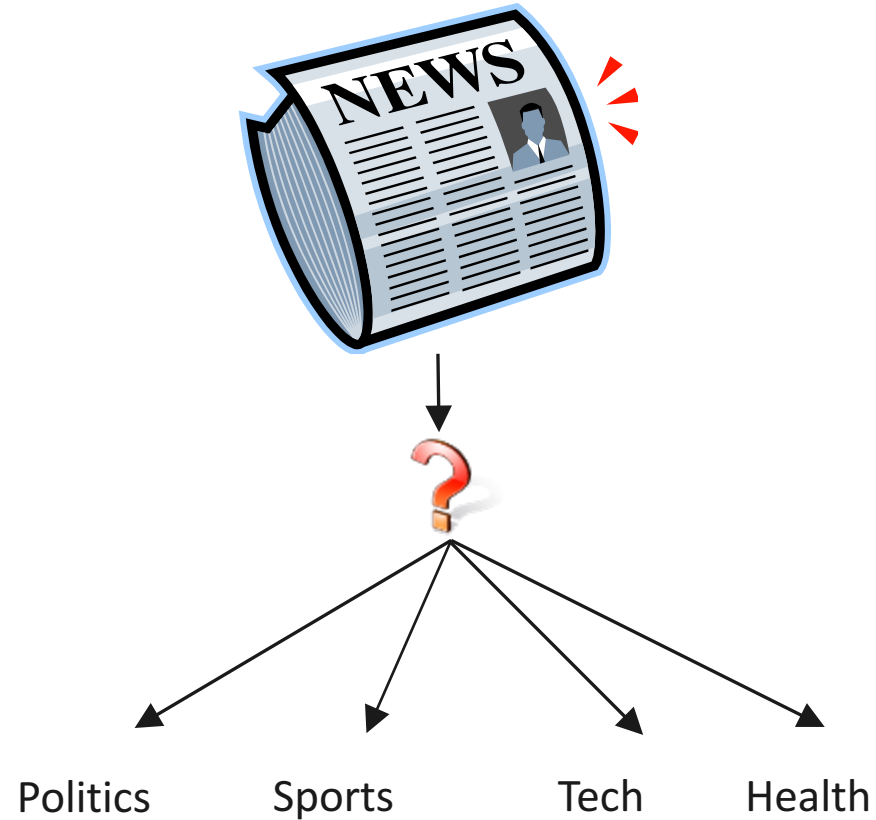
Model (Decision Tree)



EXAMPLE

Model (Classification)

Classify a news article as (politics, sports, technology, health, ...)



Using known data, develop a model to predict unknown data.

Known data (Training data)

Documents

Labels



Tech



Health



Politics



Politics



Sports

Documents consist of unstructured text. Machine learning typically assumes a more structured format of examples

Process the raw data

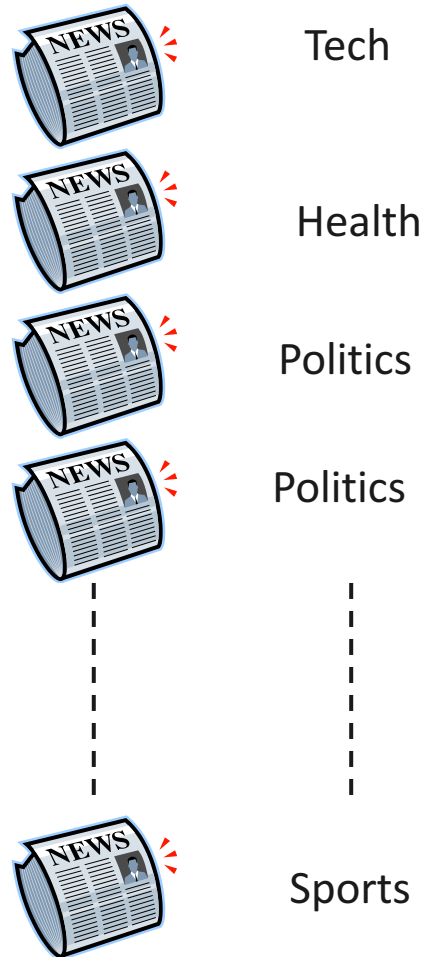
Using known data, develop a model to predict unknown data.

Known data (Training data)

Documents

Labels

Process each data instance to represent it as a feature vector



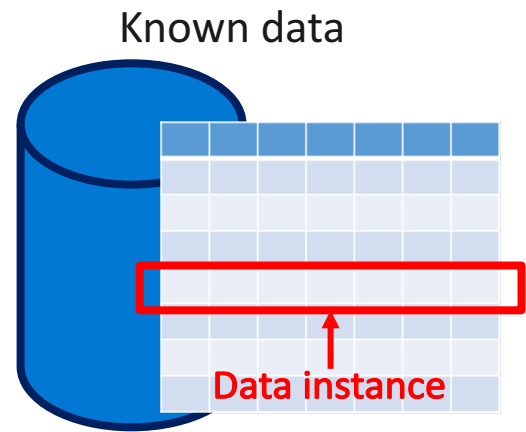
Feature

Documents

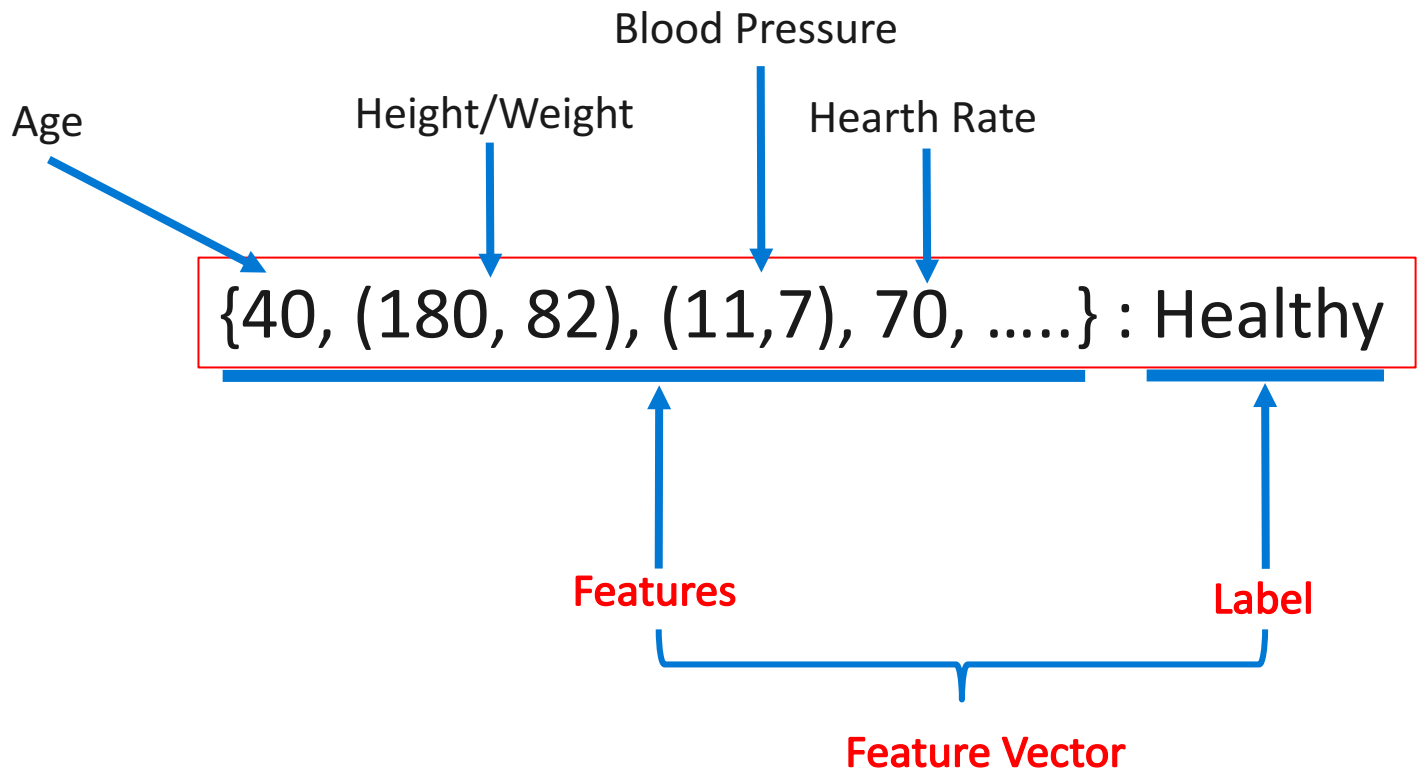
Labels

Using **known data**, develop a **model** to predict **unknown data**.

Feature vector








i.e.



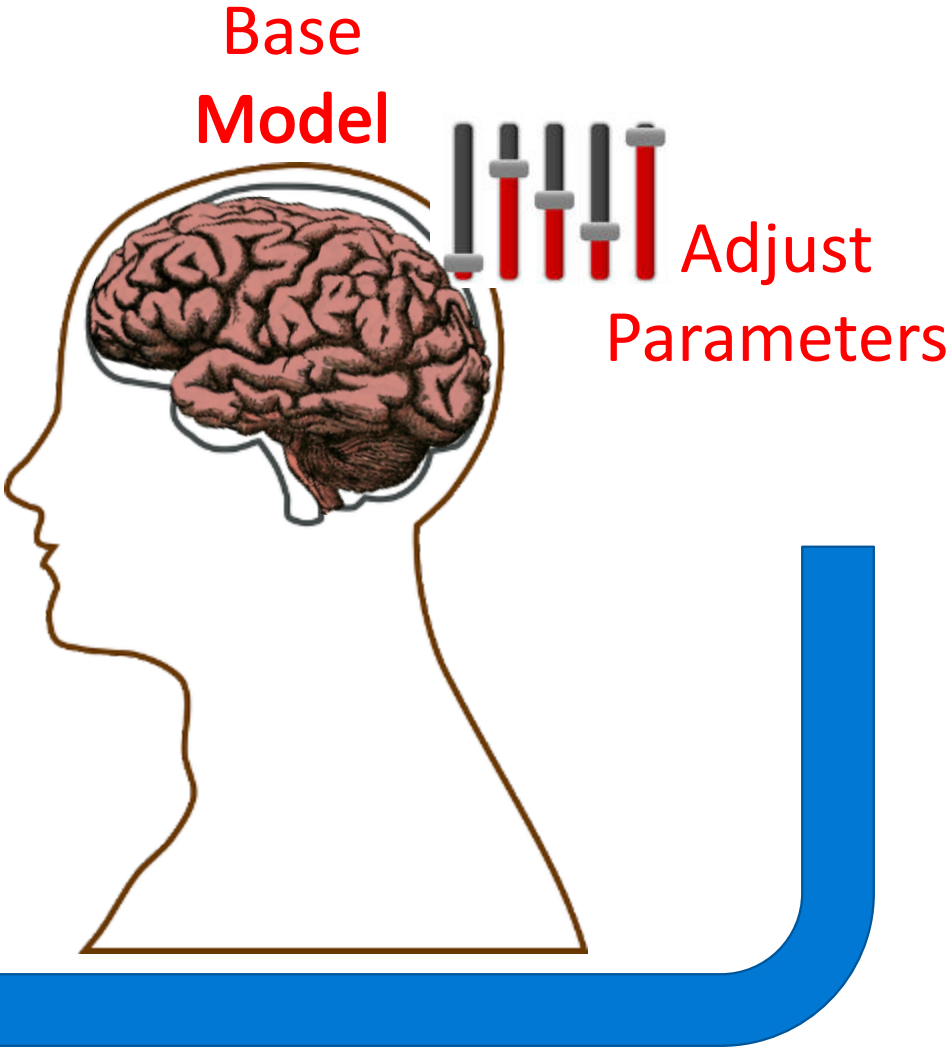
Developing a Model

Training data

<u>Documents</u>	<u>Labels</u>	Feature Vectors
	Tech	
	Health	
	Politics	
	Politics	
...	...	
	Sports	



Train
the
Model



Using known data, develop a model to predict unknown data.

DATA SCIENCE & AI

KEY TRENDS

- ▶ Accelerating adoption of AI by developers (consuming models)
- ▶ Rise of hybrid training and scoring scenarios
- ▶ Push scoring/inference to the event (edge, cloud, on-prem)
- ▶ Moving high-end developers into deep learning as non-traditional path to DS / AI dev
- ▶ Growth of diverse hardware arms race across all form factors (CPU / GPU / FPGA / ASIC / device)
- ▶ Demonstrating success of transfer learning techniques while reducing dev complexity

CHALLENGES



Data prep



Model deployment & management



Model lineage & auditing



Explain-ability

AI Platform

AI SERVICES

CONVERSATIONAL AI

Azure Bot Service

TRAINED SERVICES

Cognitive Services

CUSTOM SERVICES

Azure Machine Learning

AI INFRASTRUCTURE

AI ON DATA

Cosmos
DB

SQL
DB

SQL
DW

Data
Lake

Spark

DSV
M

Batch
AI

ACS

Edge

AI COMPUTE

CPU, GPU, FPGA

AI TOOLS

CODING & MANAGEMENT TOOLS

VS/VS Code
Tools for AI

Azure ML
Studio

Azure ML
Workbench

Others (Pycharm, Jupyter notebooks...)

DEEP LEARNING FRAMEWORKS

Cognitive
Toolkit

TensorFlow

Caffe

Others (Scikit-learn, MXNet, Keras, ...)

Open AI Ecosystem

Gluon

ONNX

Machine Learning & AI Portfolio

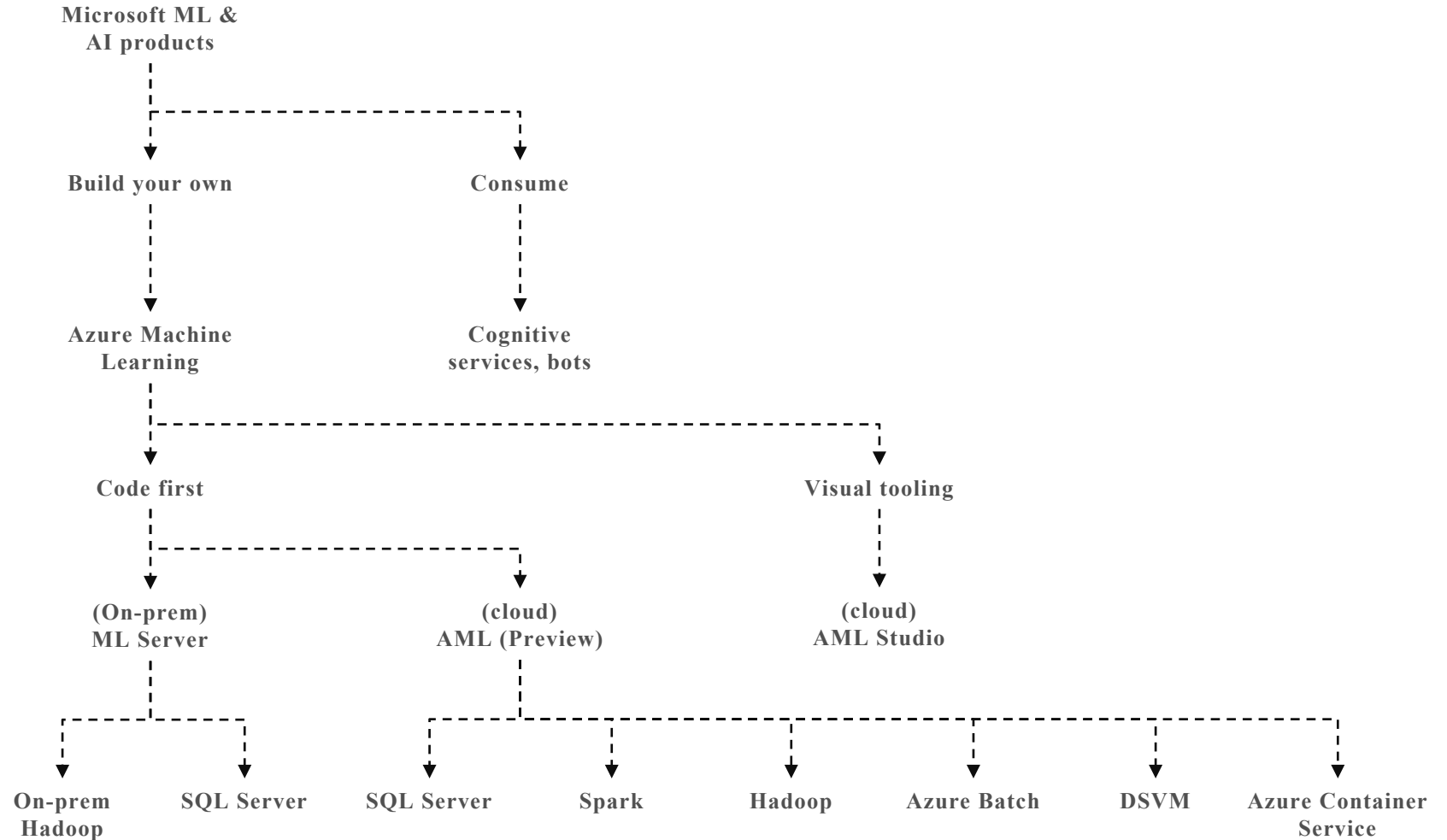
When to use what?

Build your own or consume pre-trained models?

Which experience do you want?

Deployment target

What engine(s) do you want to use?



Azure ML Experimentation

- Manage training jobs locally, scaled-up or scaled-out
- Run distributed ML training jobs (e.g Tensorflow, CNTK, etc.)
- Conduct a hyperparameter search on traditional ML or DNN
- Service side capture of run metrics, output logs and models
- Leaderboards, side by side run comparison and model selection
- Use your favorite IDEs, editors, notebooks, and frameworks

USE ANY FRAMEWORK OR LIBRARY



USE ANY TOOL

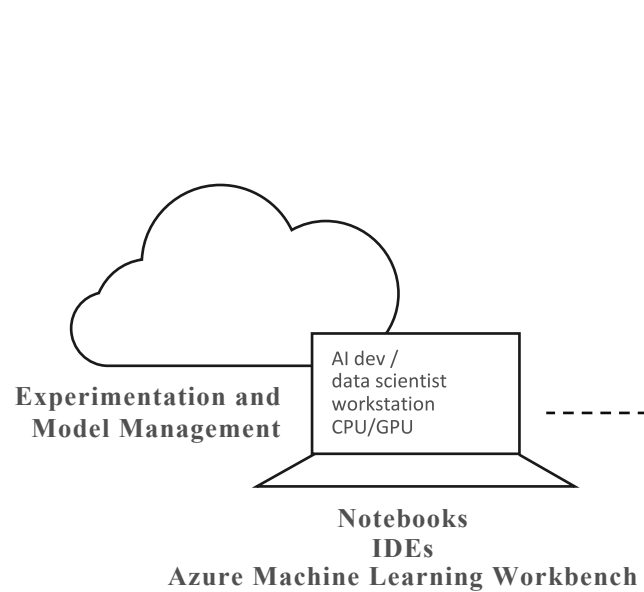


USE THE MOST POPULAR INNOVATIONS



AZURE MACHINE LEARNING

AZURE MACHINE LEARNING SERVICES



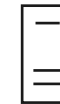
TRAIN & DEPLOY OPTIONS

AZURE



Spark
SQL Server
Virtual machines
GPUs
Container services

ON-PREMISES



SQL Server
Machine Learning Server

EDGE
COMPUTING



Azure IoT Edge

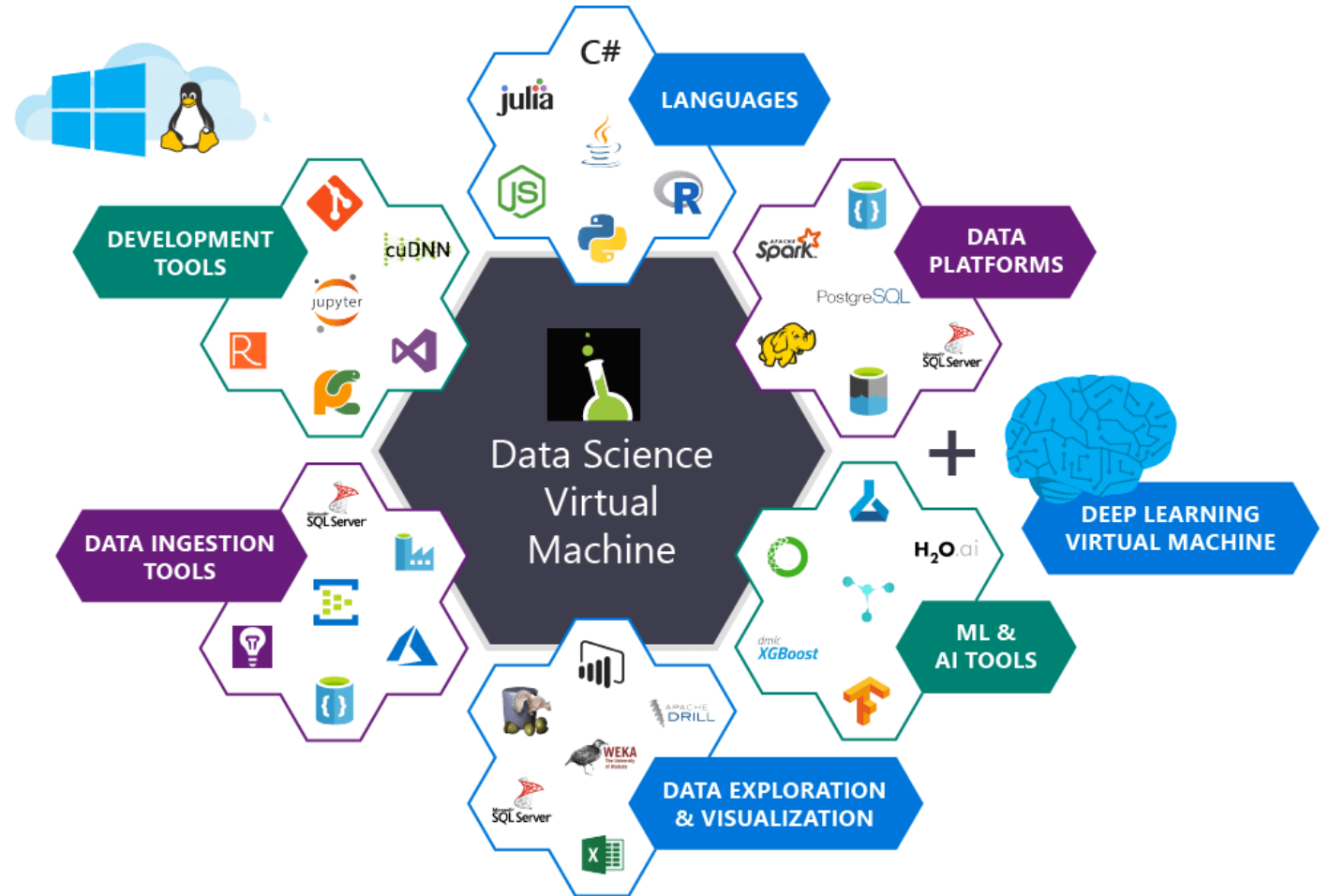
Many ways to operationalize model

- Easiest : ML Studio
 - Subject to ML Studio limitations
- Second easiest: ML model management
- Custom API with Flask / Falcon
 - Example link: <https://github.com/ms-cse/ML-endpoint-flask>



Data Science VM

- <https://aka.ms/dsvm>
- [ARM Templates](#)



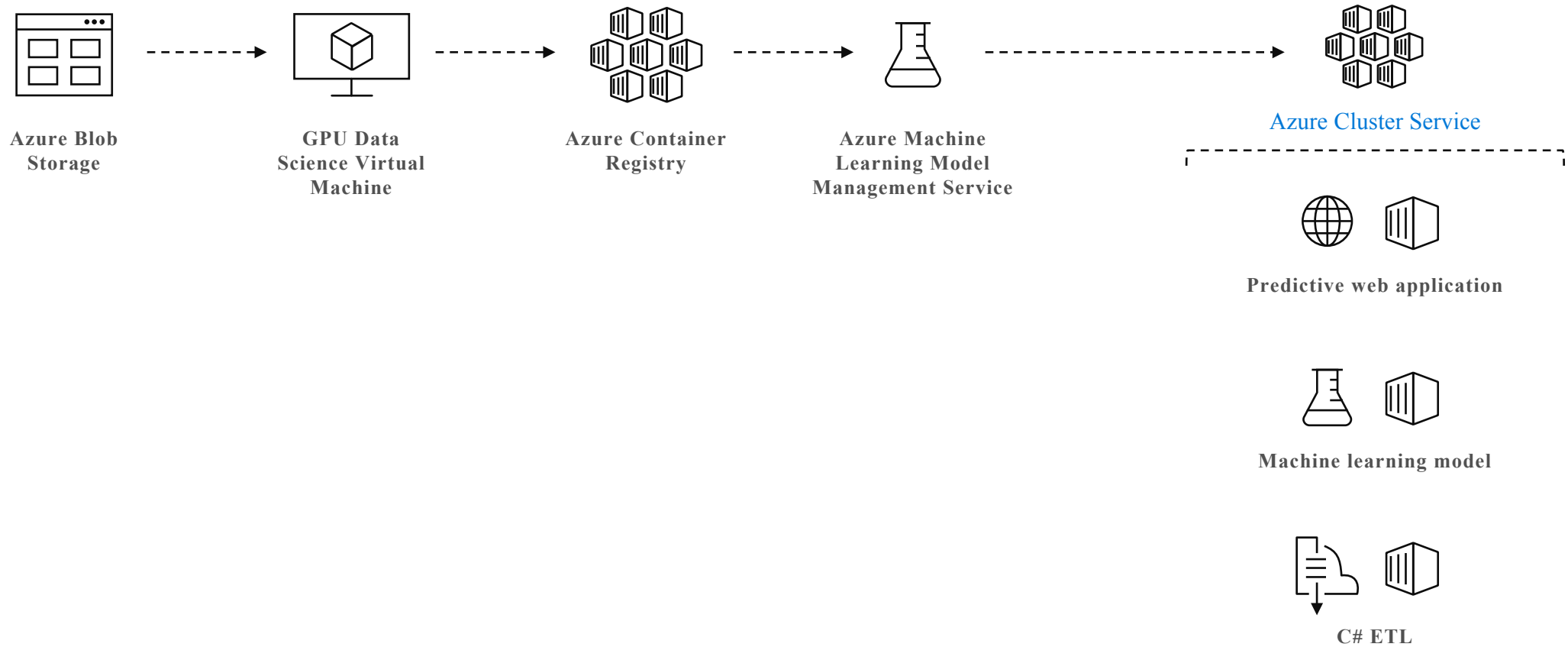
<https://www.pyimagesearch.com/2018/03/21/my-review-of-microsofts-deep-learning-virtual-machine/>

#ntk18

SAMPLE PATTERNS

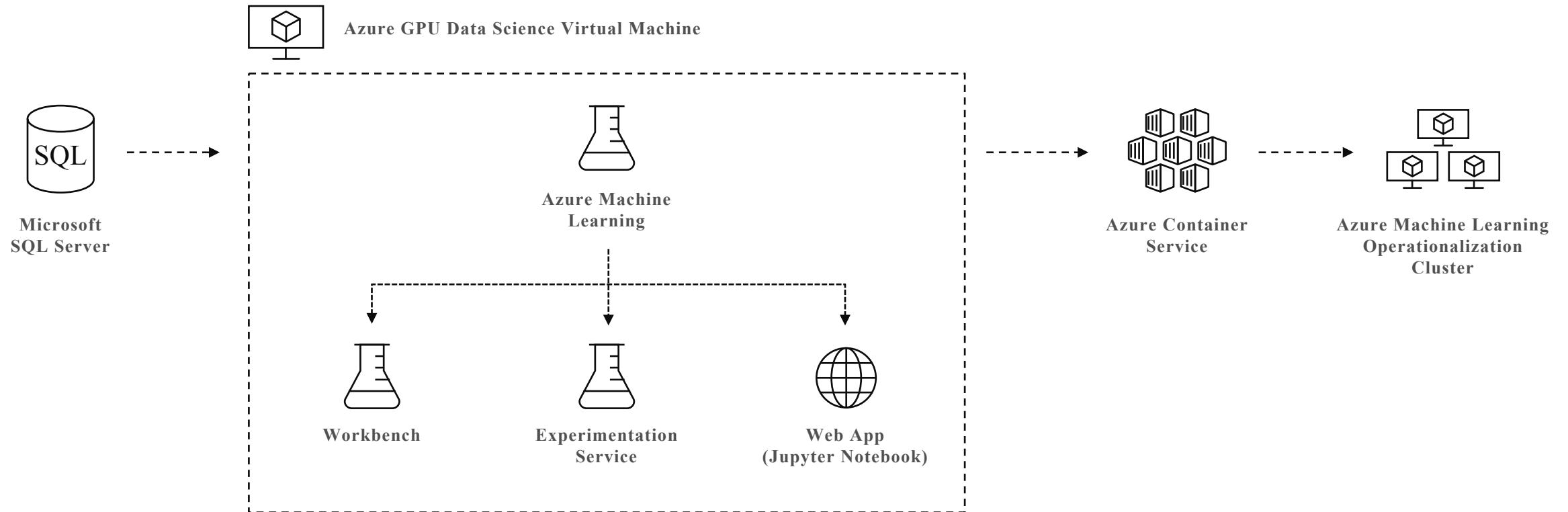
Enhanced image classification

Transfer learning, convolutional neural networks and gradient-boosting decision tree learning algorithms have redefined image classification

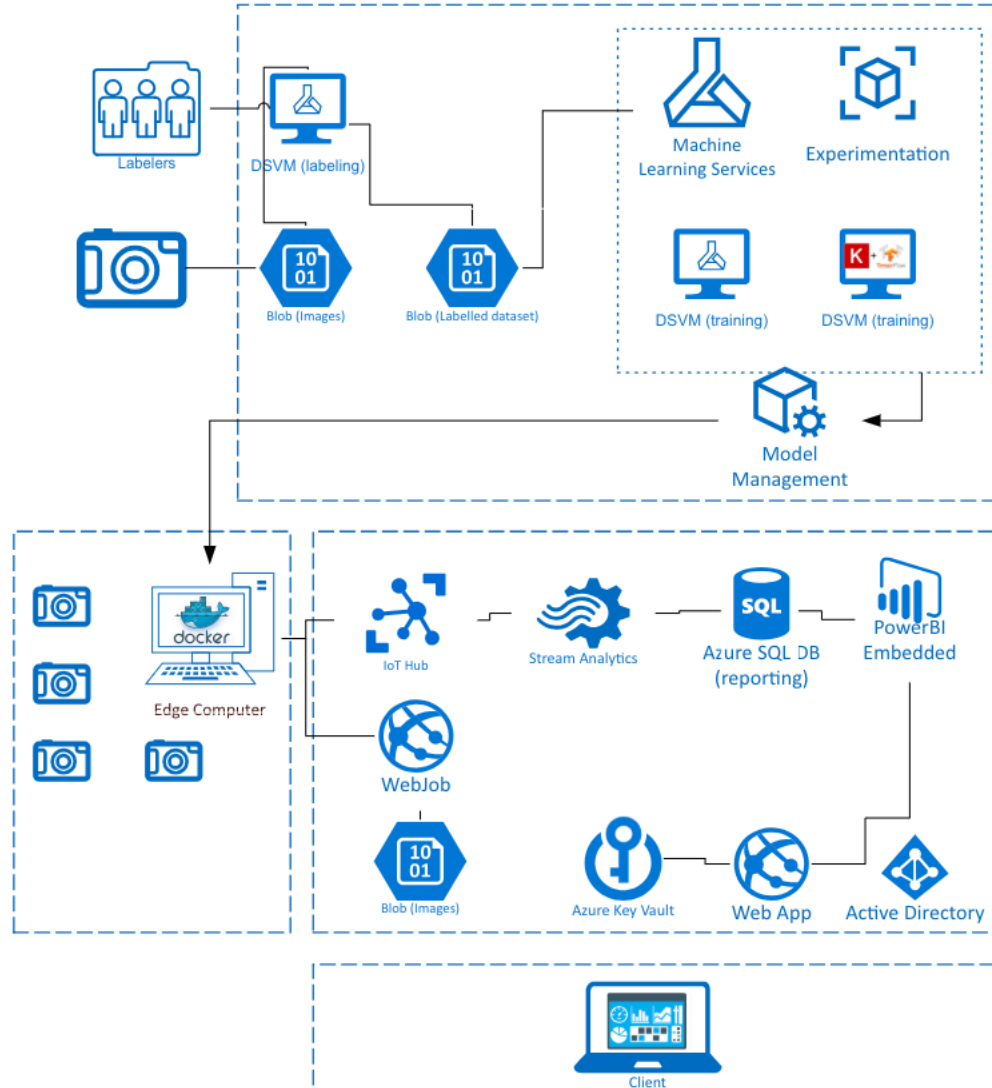


Improved text prediction

Deep learning and natural language processing boosts search efficacy and tagging accuracy



An example from practice: Deep Learning for Predictive maintenance





Drone-based electric grid inspector powered by deep learning

Challenge

- Traditional power line inspection services are costly
- Demand for low cost image scoring and support for multiple concurrent customers
- Needed powerful AI to execute on a drone solution

Solution

- Deep learning to analyze multiple streaming data feeds
- Azure GPUs support Single Shot multibox detectors
- Reliable, consistent, and highly elastic scalability with Azure Batch Shipyards

eSmart architecture

Data Sources

Ingest

Prepare


Analyze

Publish


Consume




Drone collected images
Batch upload of drone images




Azure Blob
Raw storage



Azure Blob



Cosmos DB



Azure Functions



Azure Batch



Docker Image
DNN contained in a Docker image



Cosmos DB
Contain inventory results and state changes

On-prem command center

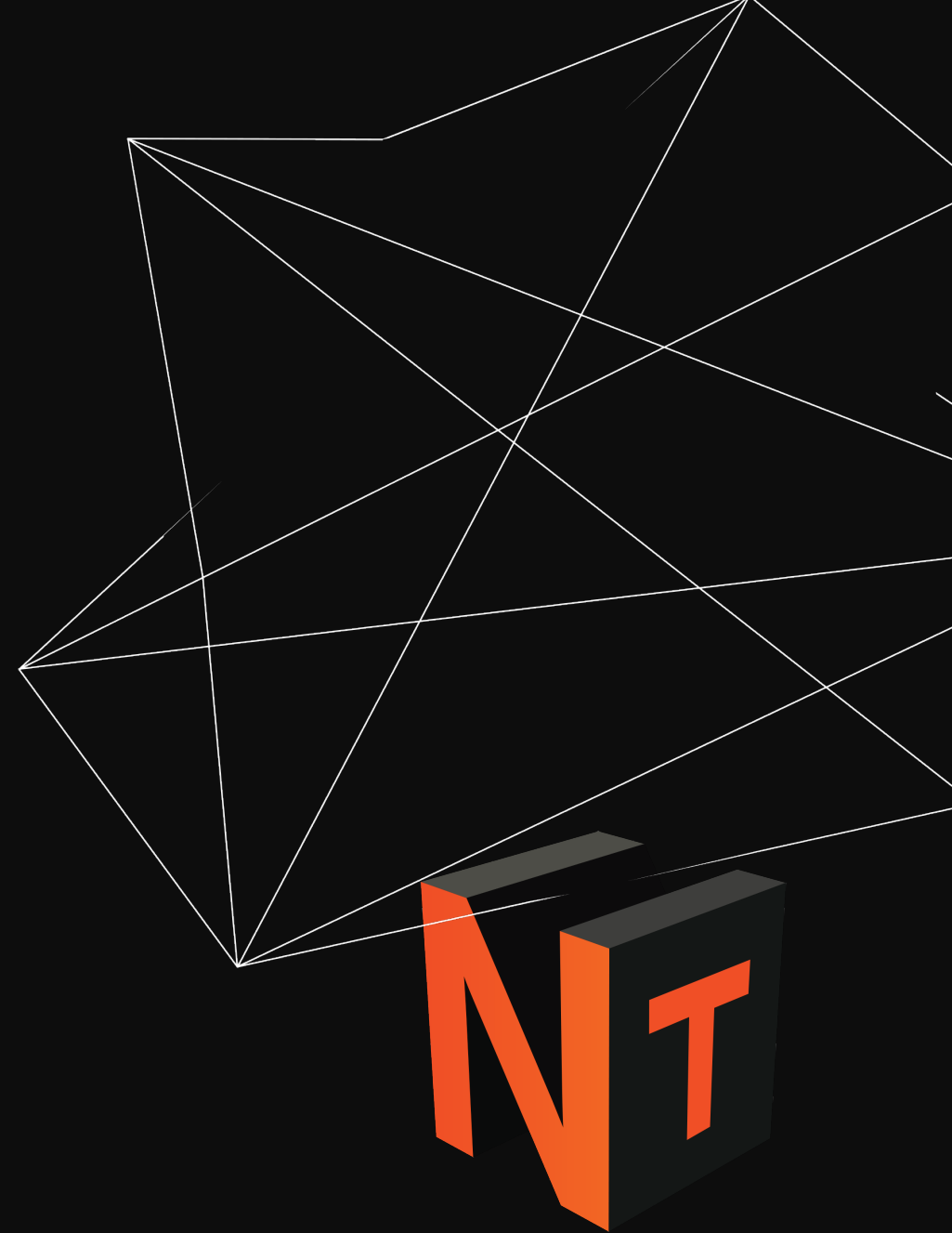
DATA

INTELLIGENCE

ACTION

Iris demo

#ntk18



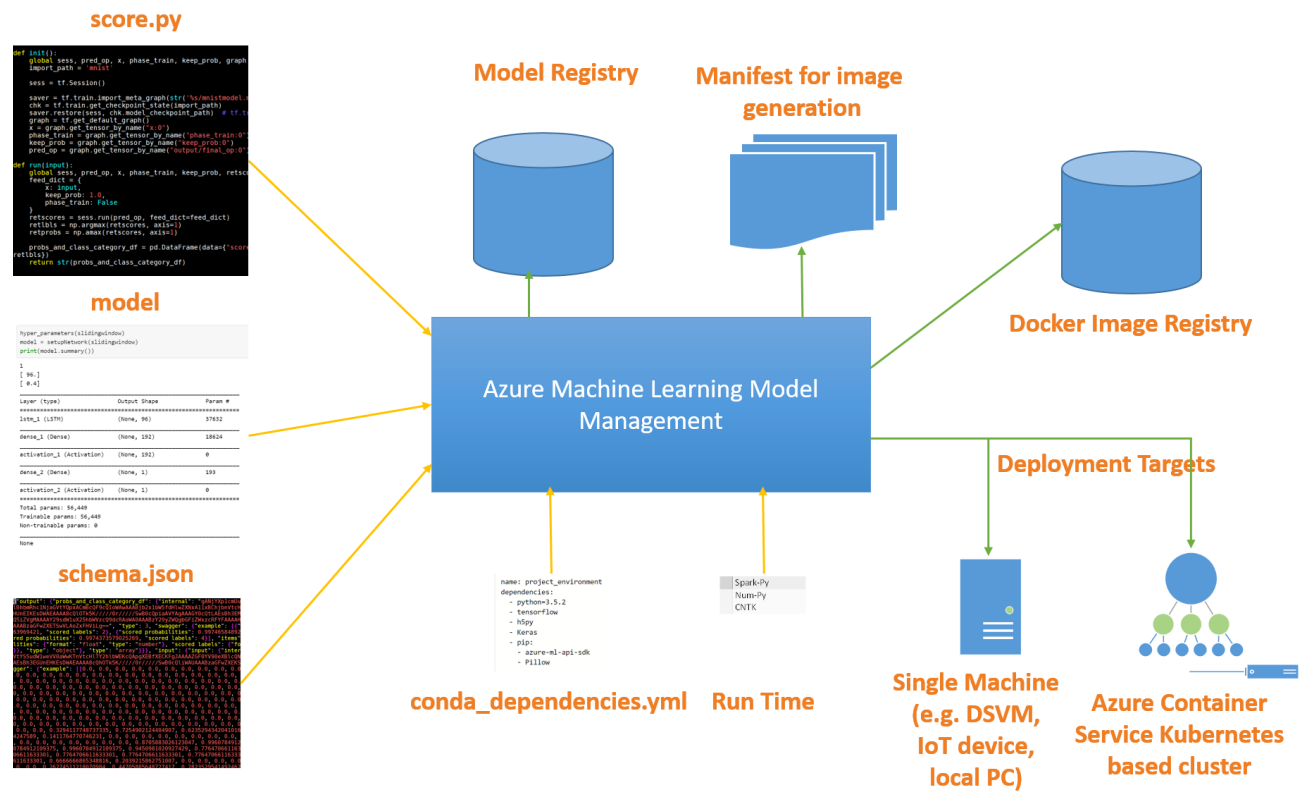
2018
NT Konferenca
Portorož | 22. - 24. maj 2018

Churn prediction

<https://github.com/Azure/MachineLearningSamples-ChurnPrediction>



Operationalizing model

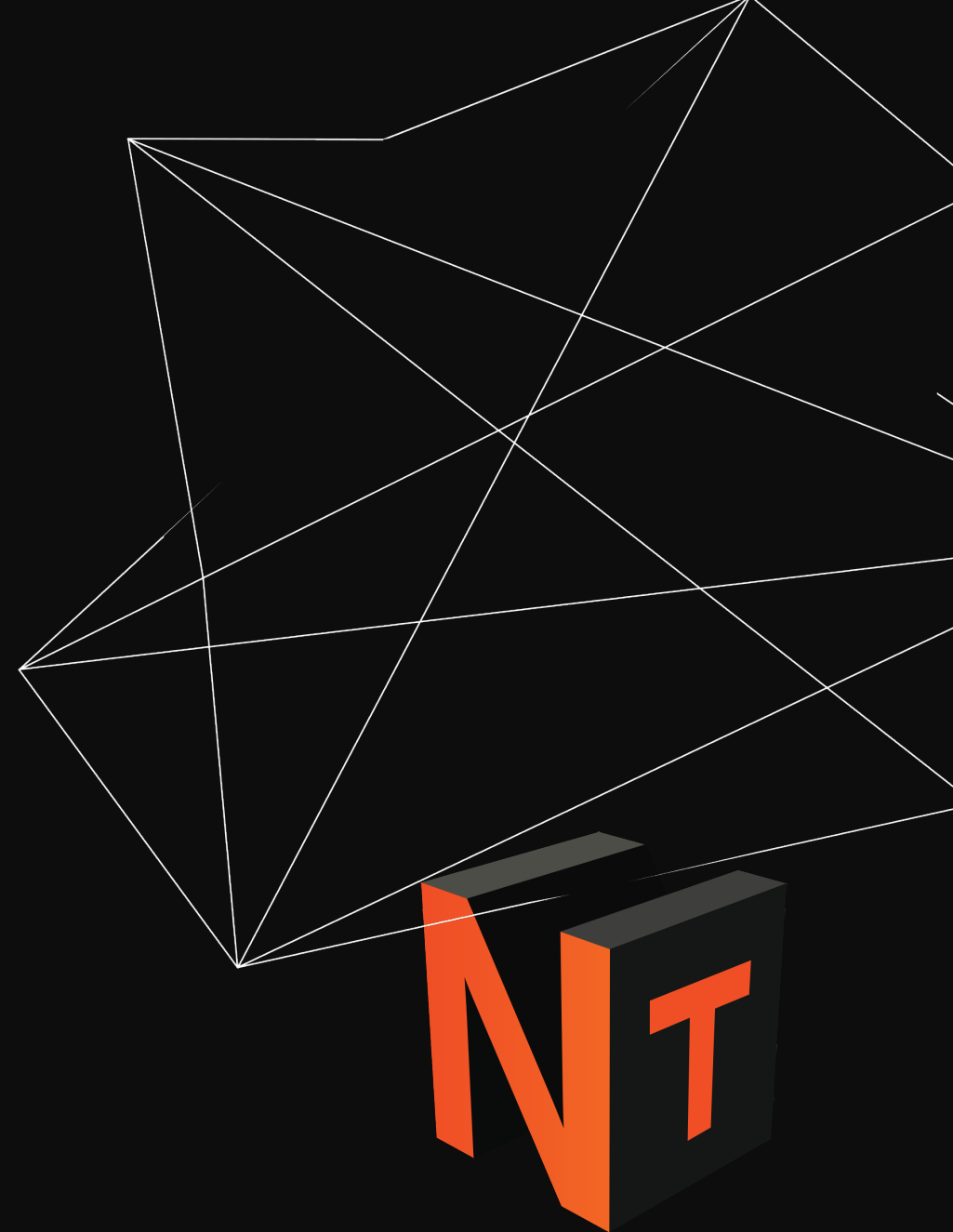


Sentiment Analysis / Text analytics

<https://vykhand.github.io/kaggle-jigsaw>

[How-to: document collection analysis](#)

#ntk18



Provisioning DSVM

- <https://github.com/Azure/DataScienceVM/tree/master/Scripts/CreateDSVM>

```
2
3
4 # Follow instructions of az login to signin to your Azure account. May need to select subscription if you have m
5 az login
6 az group create --name dsvm-win16 --location "northeurppe"
7 az group deployment create --resource-group dsvm-win16 --template-uri https://raw.githubusercontent.com/Azure/D
8
9 az vm show -n dsvm-win16 -g dsvm-win16 -d
10 az vm deallocate -n dsvm-win16 -g dsvm-win16
```



Choosing VMs for your training

Azure VM Comparison

Find and compare Azure Virtual machines specs and pricing on a one page. Check column **Best region price**, it will help you to find in what region that VM is cheaper. Also, you should know that the price in different currencies is different, sometimes the difference is significant, check this [page](#). The data updated daily from Azure API and can be not accurate. This site is not related to Microsoft or Azure. Last update was: 2018-05-20 10:00:28Z GMT

We start accepting requests for Azure Consumption Audit. We are helping reduce Azure costs up to 40%. [Learn more here.](#)

US Dollar (\$) North Europe Cost per month

Search

Cores: RAM:

VM name	# Cores	Memory (GiB)	Max # disks	Linux price	Windows price	Best price region / Diff %
Standard_A4m_v2	4	32	8	212.43	321.93	West US 2 / -28.5%
Standard_A6	4	28	8	362.08	438.00	Korea Central / -30.8%
Standard_A8_v2	8	16	16	279.59	419.75	West US 2 / -13.1%
Standard_B4ms	4	16	8	132.86	159.87	East US / -7%
Standard_B8ms	8	32	16	265.72	305.87	East US / -7%
Standard_D4	8	28	32	426.32	753.36	East US 2 / -8.2%
Standard_D4_v2	8	28	32	385.44	712.48	East US 2 / -13.3%
Standard_D4_v2_Promo	8	28	32	324.12	592.76	East US / -9.9%

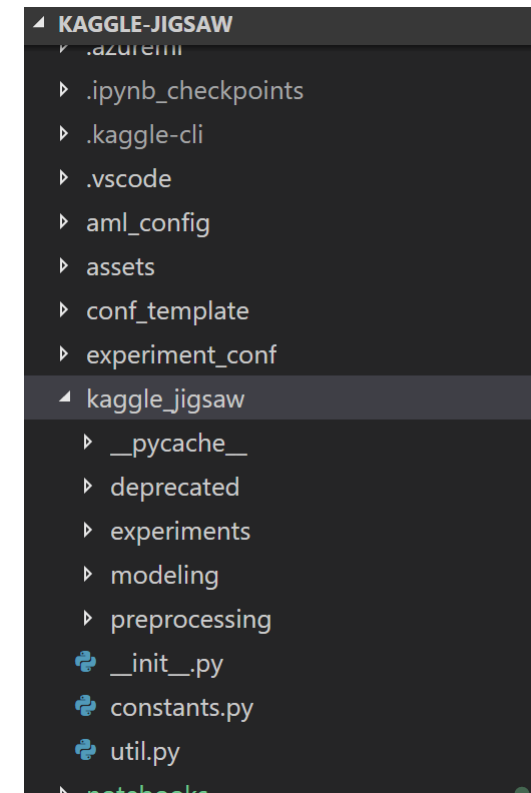
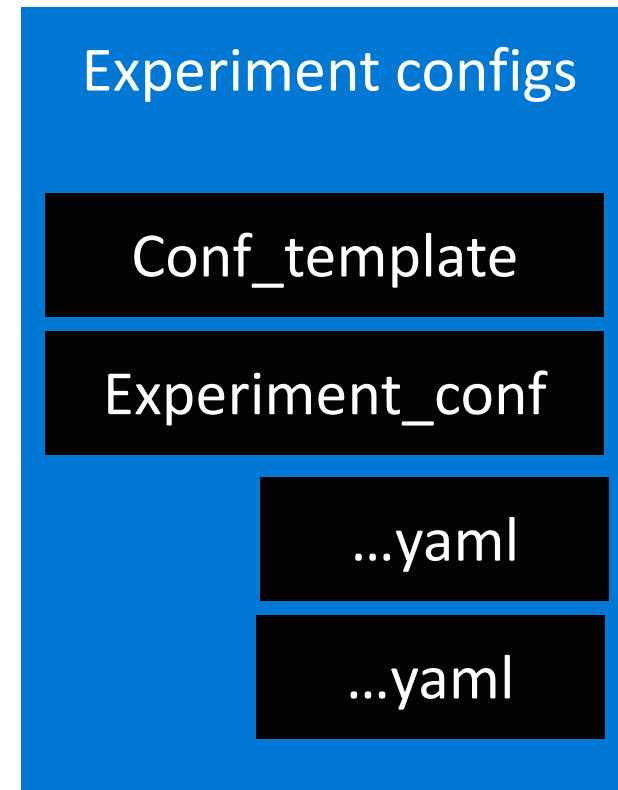
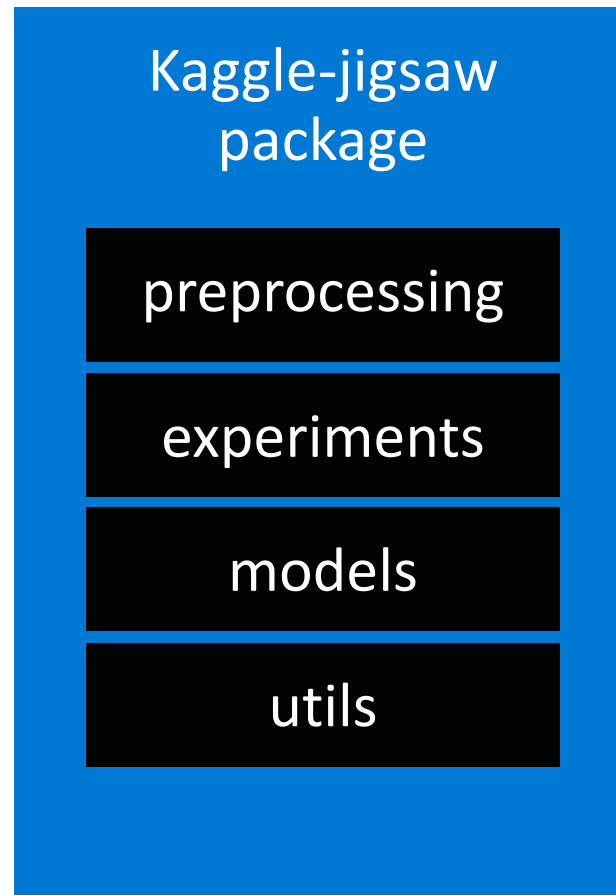
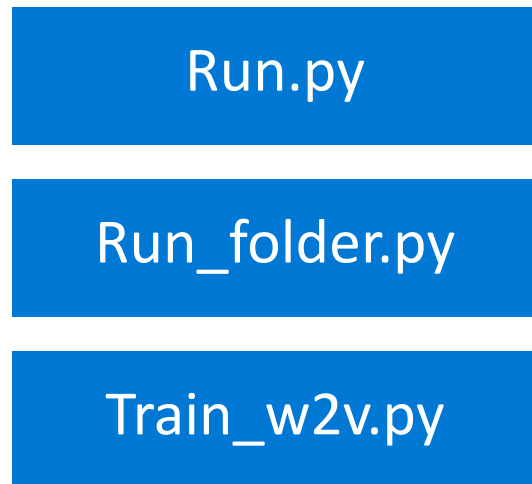
<https://azureprice.net/>



#ntk18

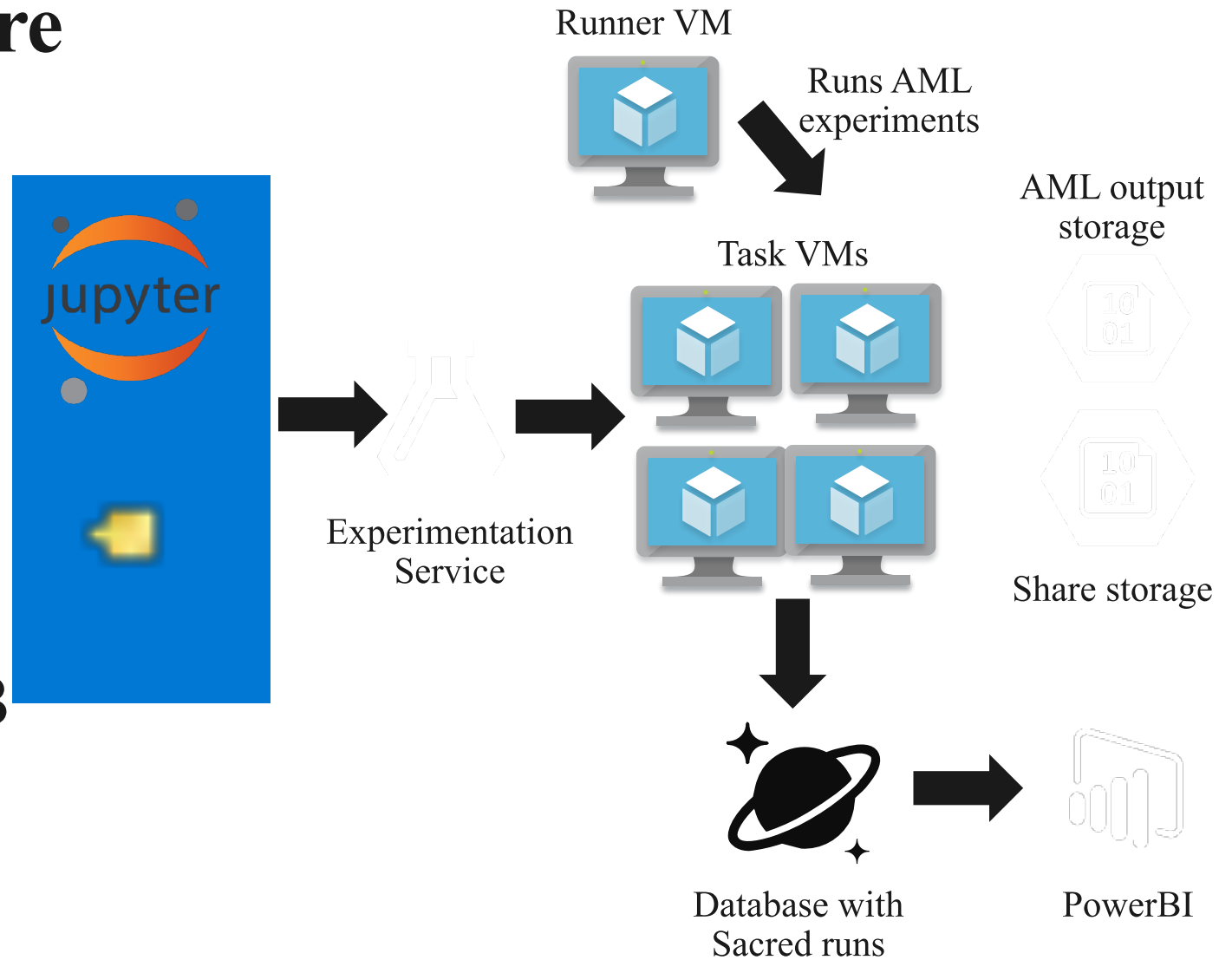
Pipeline: code structure

Everything is a hyperparameter!



Overall infra architecture

- About 6 VMs
 - 2-3 CPU, 2 GPU and a “runner”
- Local win laptop with AMLWB
- MacBook pro with AMLWB
- 2 Azure Storage accounts
- CosmosDB
- PowerBI (used at later stage)



Advanced monitoring - Sacred

- <http://sacred.readthedocs.io>
- This is [where](#) I learned it from
- Very easy to use
- Many advanced features
 - Mongo, SQL backends
 - Slack/Telegram bot integration
 - Seed setting
 - File saving

```
import logging
log = logging.getLogger("jigsaw")

from sacred import Experiment
#, SqlObserver

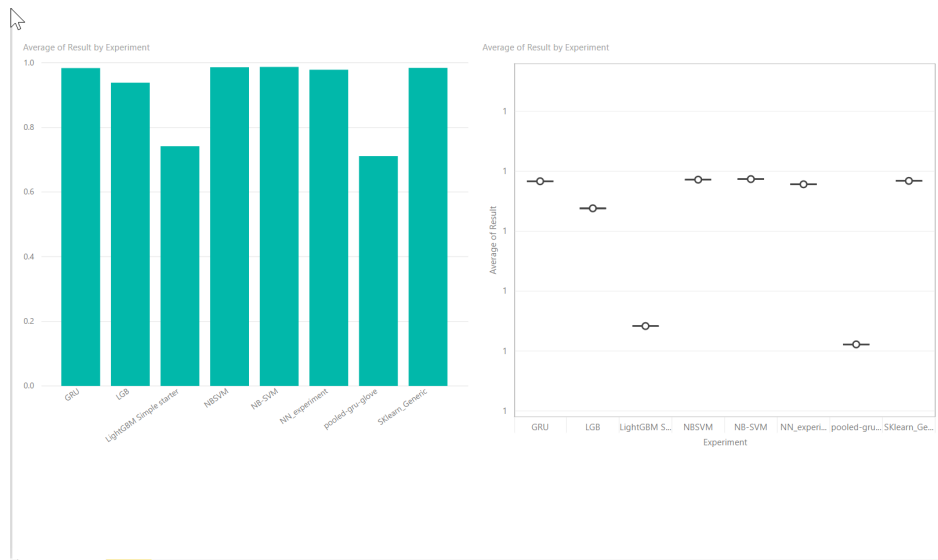
#setting up sacred

NN_experiment = Experiment("NN_experiment")
NN_experiment.observers.append(u.get_mongo_observer())
NN_experiment.observers.append(u.get_telegram_observer())
#sacred_experiment.observers.append(SqlObserver.create(os.environ["AZURE_PGSQL"]))
NN_experiment.logger = log

@NN_experiment.automain
def run_NN_experiment(exec_params, feature_params, model_params, fit_params,
    _seed, _run, _log ):
    """
    flow is defined by exec_params.model_type
    """
    model_type = exec_params["model_type"]
```

Adding PowerBi dashboards

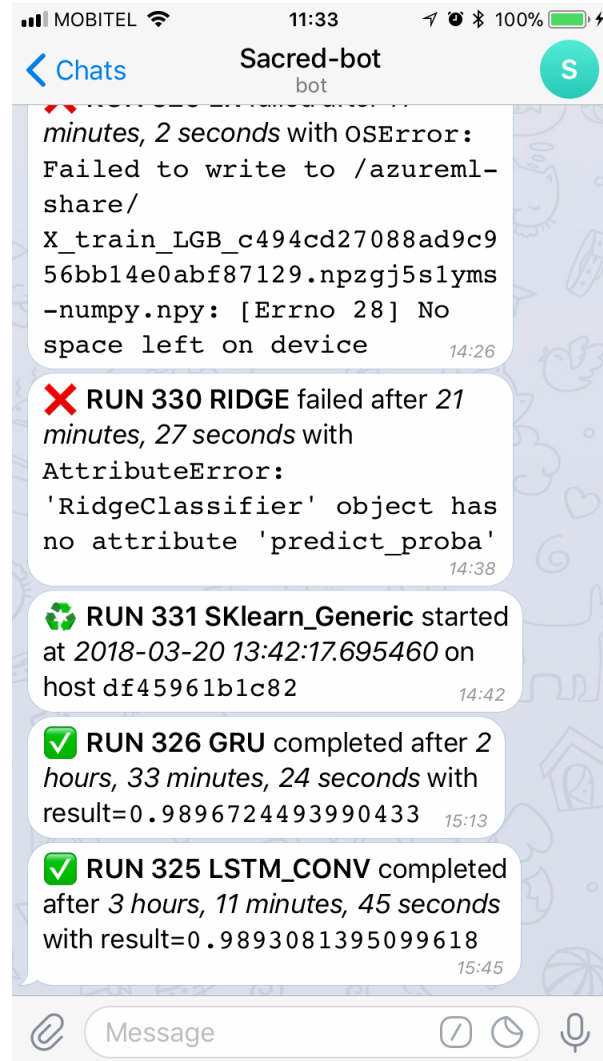
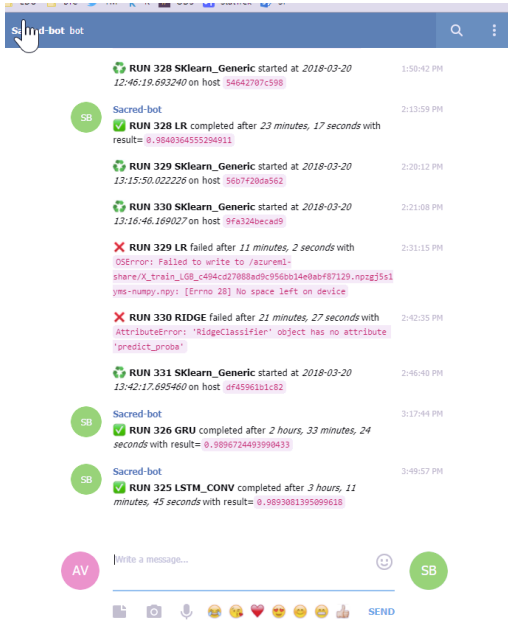
Haven't invested a lot of time here



The screenshot shows the 'cosmos-jigsaw - Query Editor' interface. The main table displays the results of 21 experiments. The columns are: num_folds_new, 1.2 Result, Experiment, Status, StartTime, and StopTime.

num_folds_new	1.2 Result	Experiment	Status	StartTime	StopTime
null	0.741734536	LightGBM Simple starter	COMPLETED	2/23/2018 10:26:54 AM	2/23/2018 10:27:19 AM
null	0.741734536	LightGBM Simple starter	COMPLETED	2/23/2018 10:27:19 AM	2/23/2018 10:27:47 AM
null	0.741734536	LightGBM Simple starter	COMPLETED	2/23/2018 6:17:11 PM	2/23/2018 6:17:42 PM
null	0.741734536	LightGBM Simple starter	COMPLETED	2/23/2018 6:17:42 PM	2/23/2018 6:18:12 PM
null	0.994690255	pooled-gru-glove	COMPLETED	2/26/2018 10:45:35 PM	2/27/2018 2:28:33 AM
null	0.994875522	pooled-gru-glove	COMPLETED	2/27/2018 6:37:15 AM	2/27/2018 10:11:54 AM
null	0	pooled-gru-glove	COMPLETED	2/27/2018 9:37:41 PM	2/27/2018 10:22:10 PM
null	0.997117302	pooled-gru-glove	COMPLETED	2/27/2018 11:06:24 PM	2/28/2018 1:55:41 AM
null	0.996795287	pooled-gru-glove	COMPLETED	2/28/2018 8:12:06 AM	2/28/2018 11:09:29 AM
null	0.994458982	pooled-gru-glove	COMPLETED	2/28/2018 7:06:30 PM	2/28/2018 10:02:09 PM
null	0	pooled-gru-glove	COMPLETED	2/28/2018 10:47:19 PM	2/28/2018 11:19:32 PM
FALSE	5	0.986890755 NB-SVM	COMPLETED	3/5/2018 10:30:59 AM	3/5/2018 11:54:23 AM
FALSE	5	0.986890755 NB-SVM	COMPLETED	3/5/2018 12:23:57 PM	3/5/2018 2:02:53 PM
FALSE	10	0.987042738 NB-SVM	COMPLETED	3/5/2018 2:49:16 PM	3/5/2018 6:09:31 PM
FALSE	5	0.987220949 NB-SVM	COMPLETED	3/5/2018 9:35:16 PM	3/5/2018 11:22:10 PM
FALSE	5	0.526764159 LGB	COMPLETED	3/9/2018 12:35:18 PM	3/9/2018 1:59:57 PM
FALSE	5	0.983630905 LGB	COMPLETED	3/9/2018 5:02:32 PM	3/9/2018 6:28:38 PM
FALSE	5	0.985213868 LGB	COMPLETED	3/9/2018 6:34:53 PM	3/9/2018 10:34:24 PM
FALSE	5	null LGB	COMPLETED	3/10/2018 12:27:39 AM	3/10/2018 3:49:42 AM
FALSE	5	null LGB	COMPLETED	3/10/2018 9:30:22 AM	3/10/2018 11:07:05 AM
FALSE	5	null LGB	COMPLETED	3/10/2018 4:51:40 PM	3/10/2018 8:49:21 PM

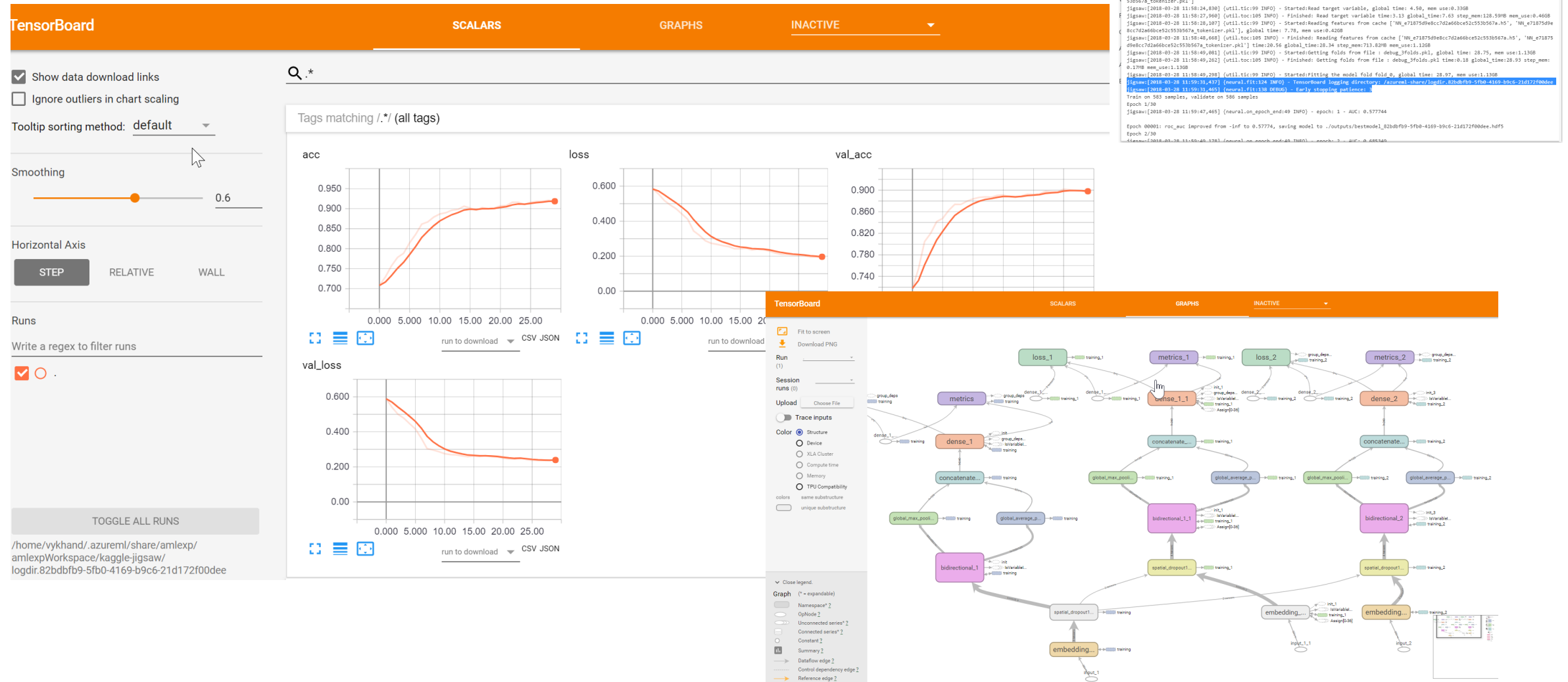
Telegram integrations



```
def get_telegram_observer():
    base_path = os.path.dirname(__file__)
    conf_file = os.path.join(base_path, "..", "aml_config", "telegram_conf.json")
    obs = TelegramObserver.from_config(conf_file)
    # copied that from the source file and added RUN id
    obs.started_text = "🔄 *RUN {_id} {experiment[name]}* " \
        "started at {_start_time}_ " \
        "on host `{host_info[hostname]}"`"
    obs.completed_text = "✅ *RUN {_id} {experiment[name]}* " \
        "completed after {_elapsed_time}_ " \
        "with result=`{result}`"
    obs.interrupted_text = "⚠️ *RUN {_id} {experiment[name]}* " \
        "interrupted after {_elapsed_time}_"
    obs.failed_text = "❌ *RUN {_id} {experiment[name]}* failed after " \
        "{_elapsed_time}_ with `{error}`\n\n" \
        "#Backtrace:\n``{backtrace}```"
    return obs
```

TODO: had practice refactor

Integrating with tensorboard



Predictive maintenance

<https://github.com/Azure/MachineLearningSamples-PredictiveMaintenance>

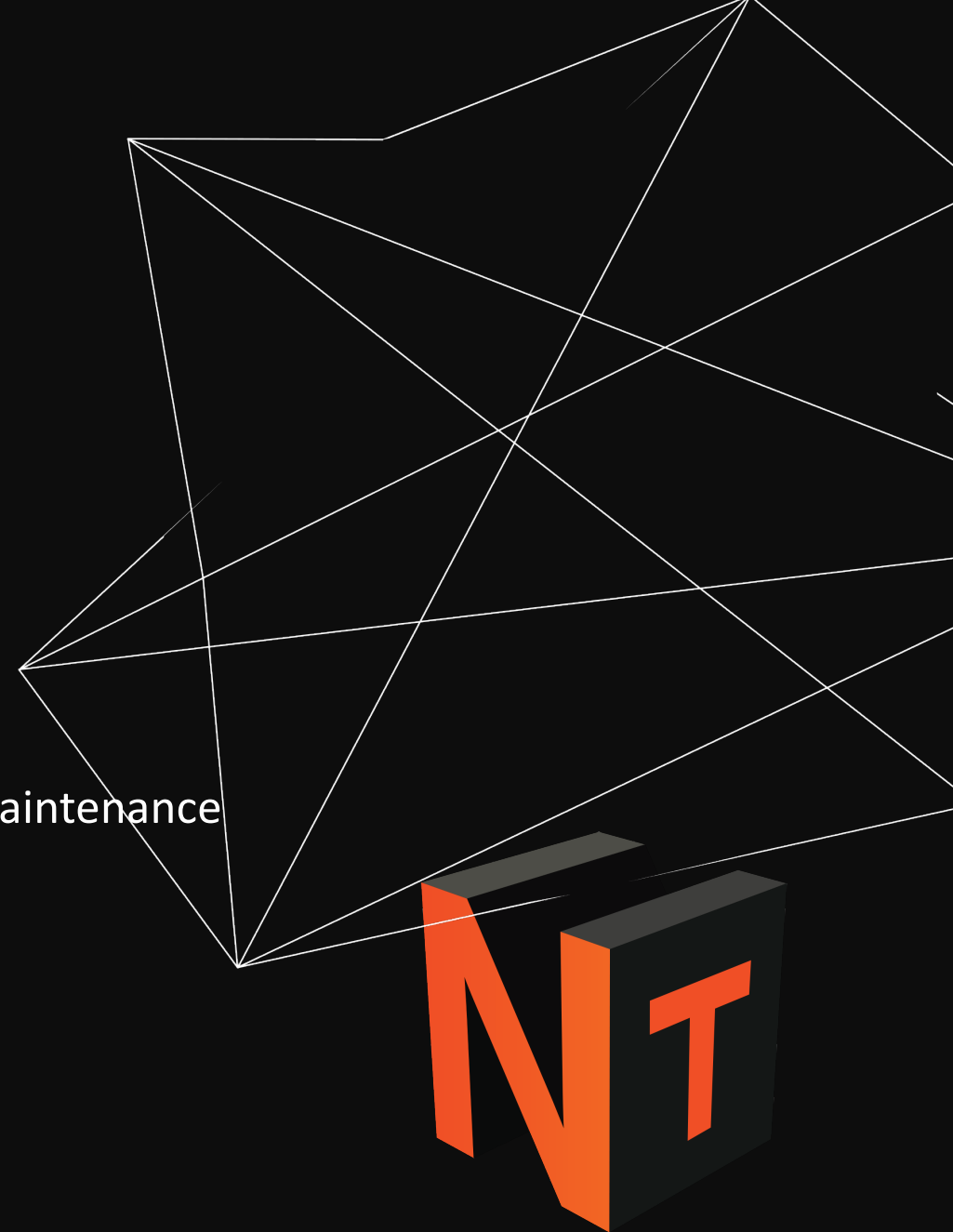


Image classification / object detection



Labelling images



Reusing the object detection example



Microsoft Build 2018 announcements



#ntk18

2018
NT Konferenca
Portorož | 22. - 24. maj 2018

Microsoft Build 2018 announcements

- Python SDK for Azure ML (preview)
- [Python packages for ML](#)
 - Computer Vision
 - Forecasting
 - Text Analytics
- ML.Net – machine learning framework for .Net developers
- Azure ML with project Brainwave (private review)
 - FPGA acceleration

<https://azure.microsoft.com/en-us/blog/azure-ai-platform-announcements-new-innovations-for-developers/>



Azure ML SDK for Python

A Control Plane API for the Data Scientist interacting with the Azure ML Cloud Services

- Execution of Training on different compute targets
- Manage Data Connections
- Provision Compute
- Manage Compute Targets
- Query and manipulate history of training runs
- Download artefacts of a training run
- Deploy a model to production

