



1

Microsoft  
tech·days  
Kistamässan Stockholm  
22-24 oktober 2019

# 4 machine learning questions Azure can answer

Dr. Nico Jacobs  
@SQLWaldorf

2

*automatic*

*"Making predictions is hard,  
especially about the future"*

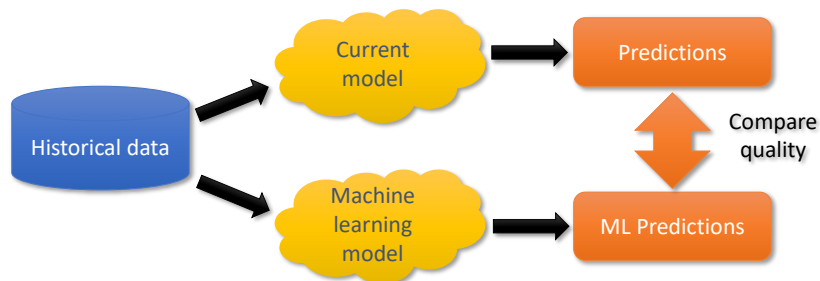
3



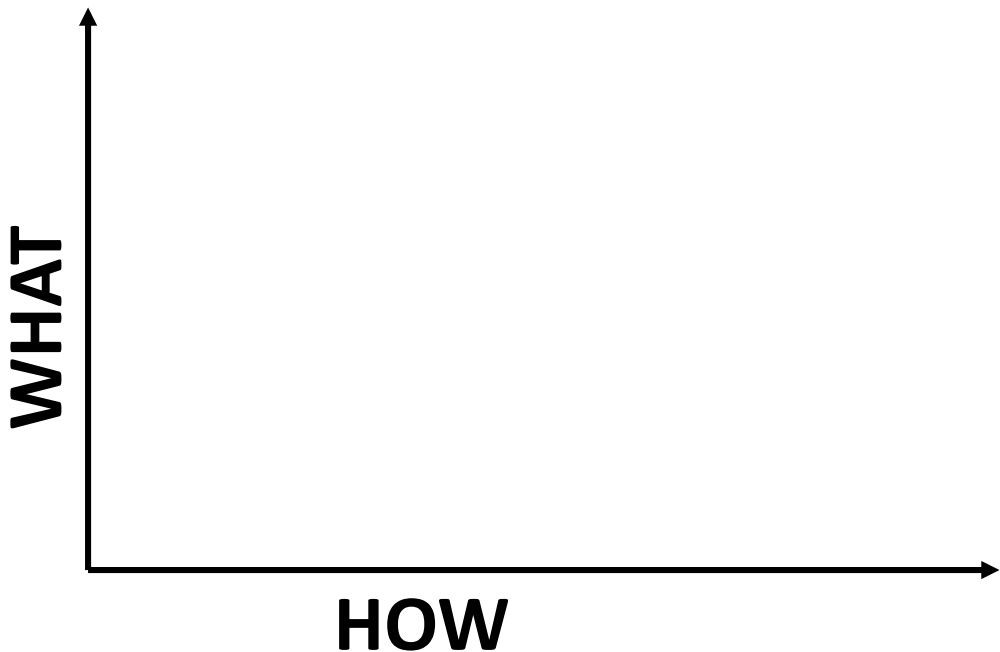
4

# Machine learning models

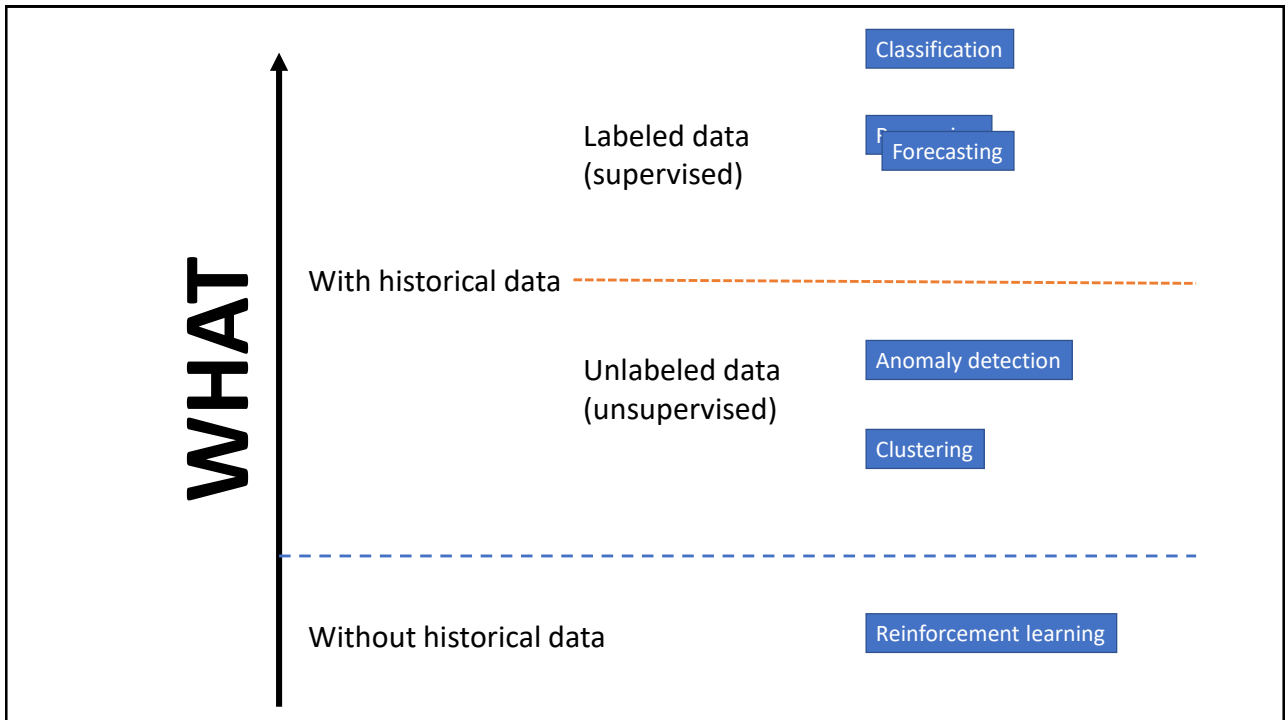
- Machine learning algorithms build models from historical data
- We compare the performance of a machine learning model against the currently used model



5



6



7

## Learning functions

- In machine learning a function is generated which maps  $n$  input parameters on a result:

$$f(i_1, i_2, \dots, i_n) \rightarrow \text{result}$$

- $f$  is the model to be learned
- The inputs  $I$  are called features
- The result is called the label
- If  $f$  is created based on labeled data it is called supervised learning, else it is called unsupervised learning

8

## Learning functions

- In machine learning a function is generated which maps n input parameters on a result:

$$f(i_1, i_2, \dots, i_n) \rightarrow \text{result}$$

- In supervised learning if the result is a class this is called classification
- In supervised learning if the result is a continuous value this is called regression
  - Unless the inputs form a time series, then it's called forecasting

9

## Classification: Which group, which class?

- Will a customer churn or not?
  - $f(\text{age, gender, totalSpending}) \rightarrow \text{willChurn}$
- Will this machine fail in the next week or not?
- Should we advertise computers, bikes or clothing to this customer?

10

## Classification: Which group, which class?

- Classification predicts a discrete value
  - A prediction is right or wrong
- Most techniques distinguish between two classes
- Some techniques will combine multiple two-class classifiers to obtain a multi-class classifier
- Regression problems can be discretized into a classification problem

11

## Regression: How many, how much

- What is the expected cost of a specific house X
  - $f(\text{distanceShop}, \text{distanceHighway}, \text{nrBedrooms}, \text{nrBathrooms}) \rightarrow \text{price}$
- What is the predicted cost of car accidents caused by car insurance customer X?
- How much should we pay employee X?

12

## Regression

- Type of supervised learning
- Some classification problems can be converted into regression problems:
  - High versus low income: classification
  - Predict income: regression
- More difficult to measure quality
  - Answers will always be wrong
  - But we can measure how near they are to the correct answer
  - Different quality criteria
    - No accuracy or precision, but mean absolute error etc.

13

## Forecasting

- Forecasting is a special case of regression
  - $f(\text{result3DaysAgo}, \text{result2DaysAgo}, \text{result1DayAgo}) \rightarrow \text{resultToday}$
- The training data is a time series
  - A value measured at different points in time
- Based upon a base level, a trend and seasonality one predicts the continuation of the time series
- Typical examples:
  - Predicting future stock market value
  - Predicting future demand for specific product

14

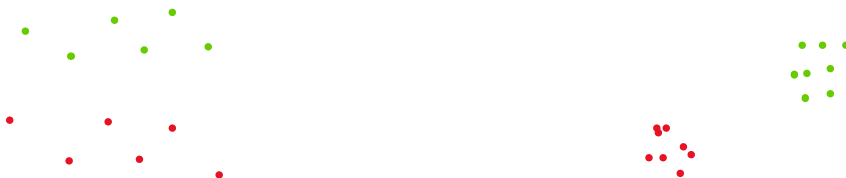
## Clustering: which items belong together?

- Find groups of similar customers
  - $f(\text{age, gender, totalSpending}) \rightarrow \text{clusterN}$
- Find groups of similar messages on social media

15

## Clustering

- Unsupervised learning: training data doesn't specify the clusters
- Since the training data is not labeled there is no way to know if the result is right or wrong
- But we can measure e.g. how compact the clusters are



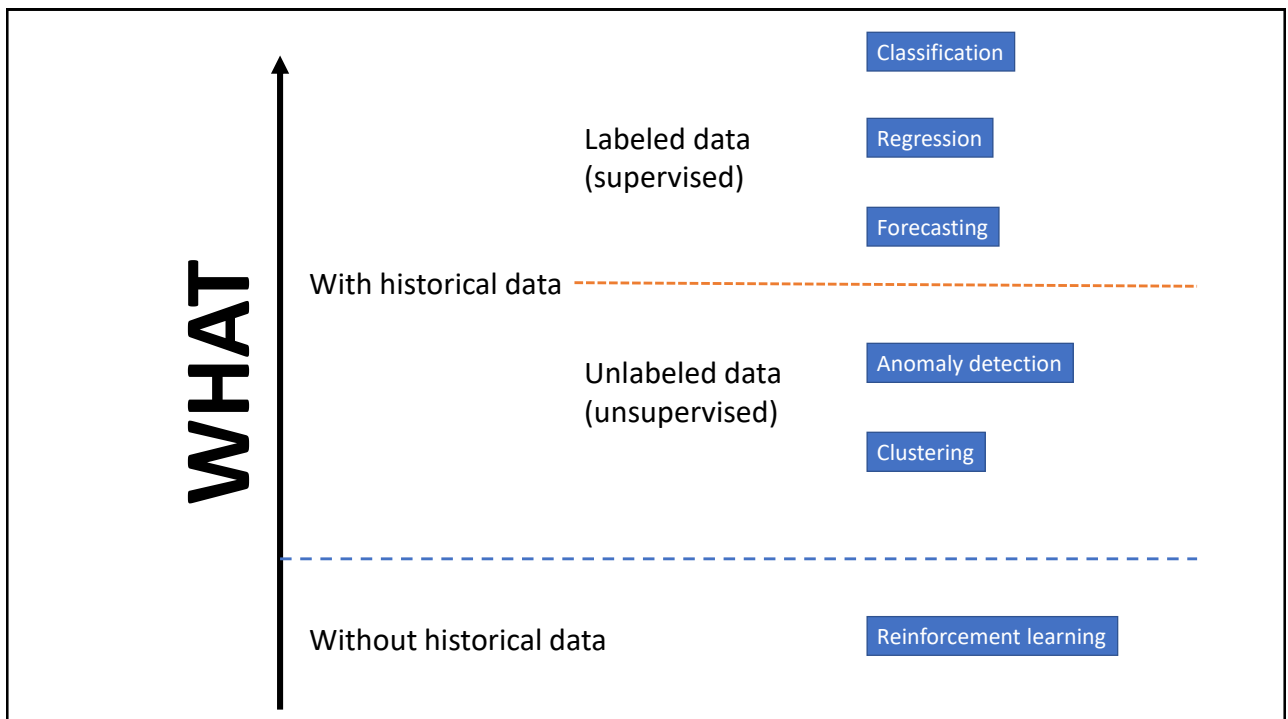
16



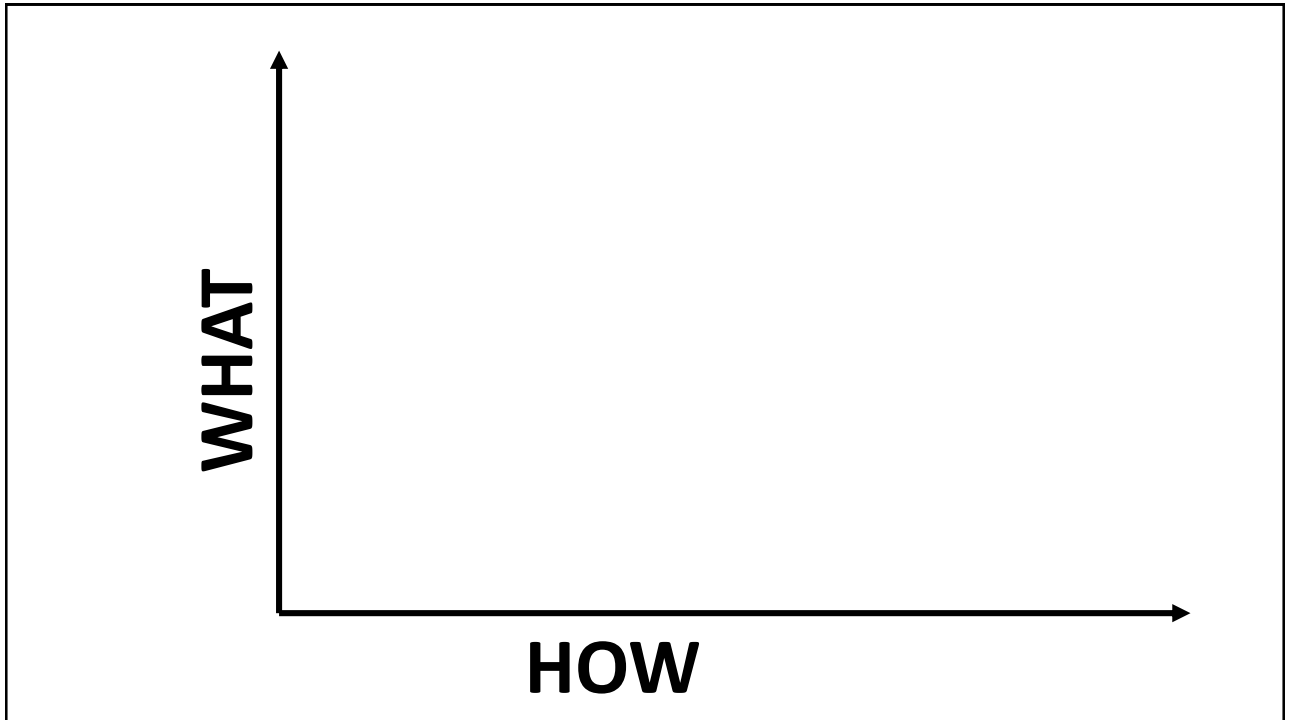
## Anomaly detection: what doesn't fit?

- Which credit card transactions are fraudulent?
- Occurs in supervised as well as unsupervised settings
- Can often be seen as special case of classification or regression

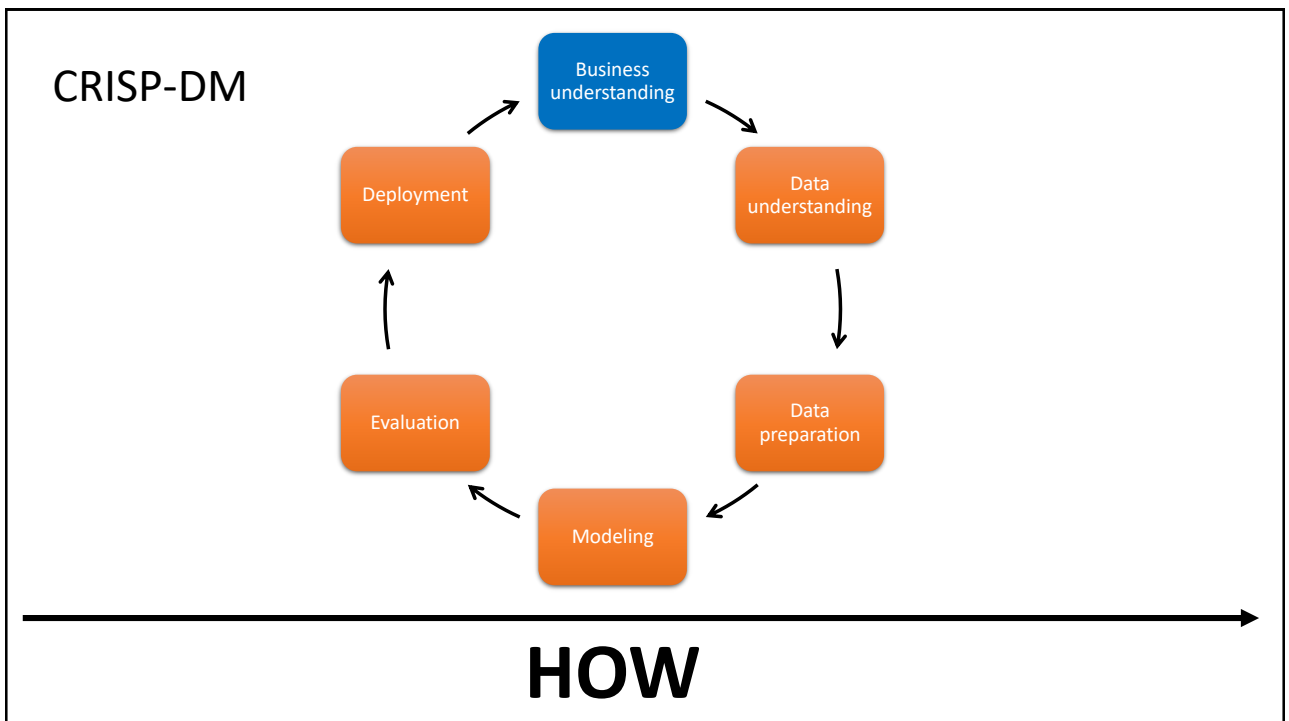
17



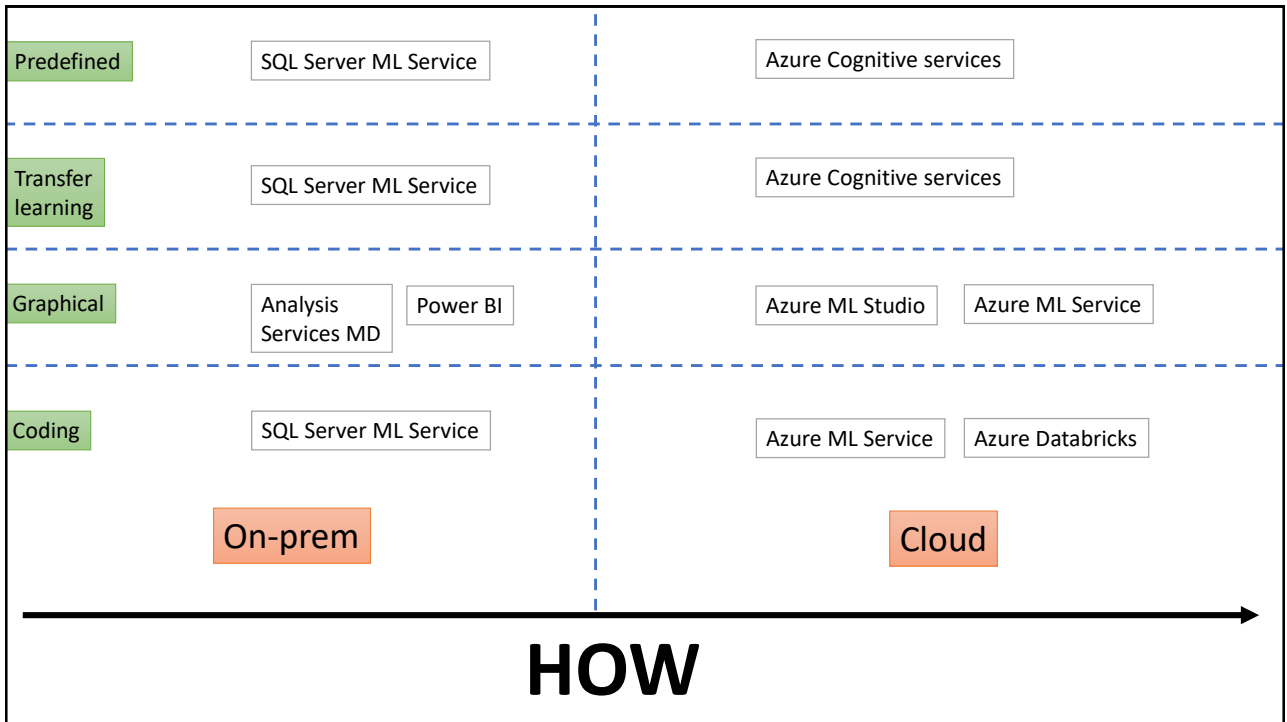
18



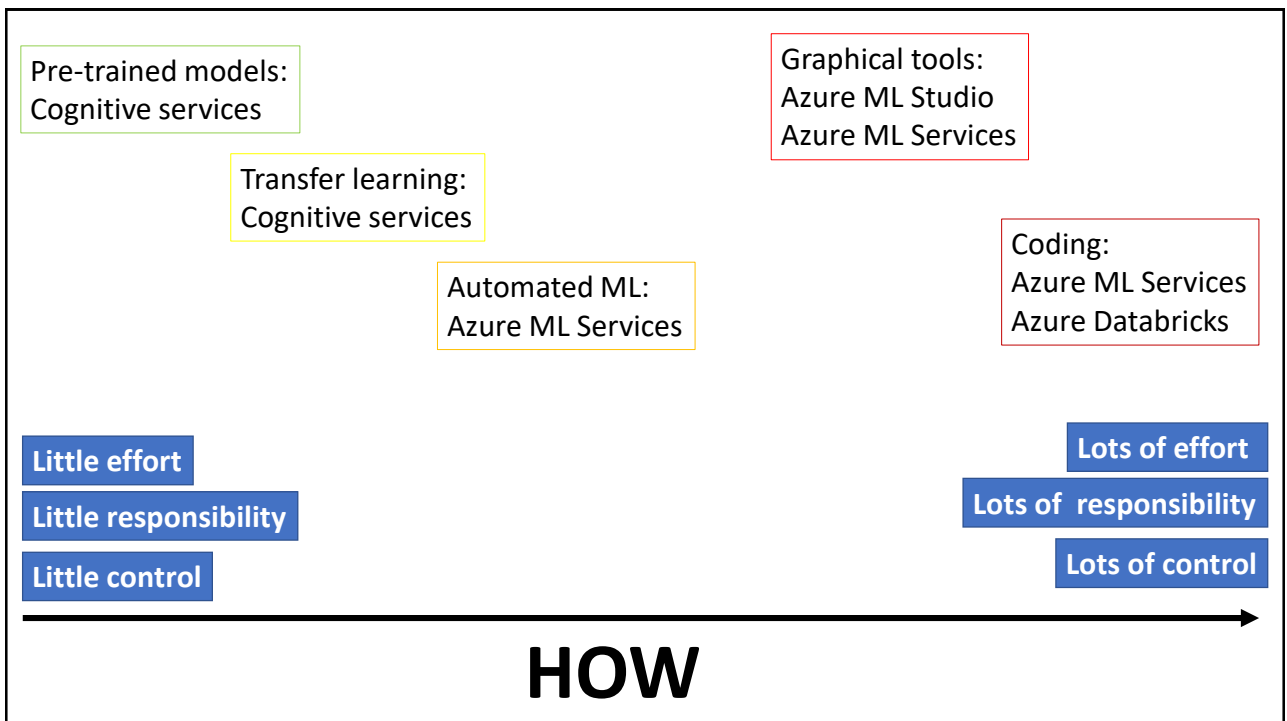
19



20



21



22

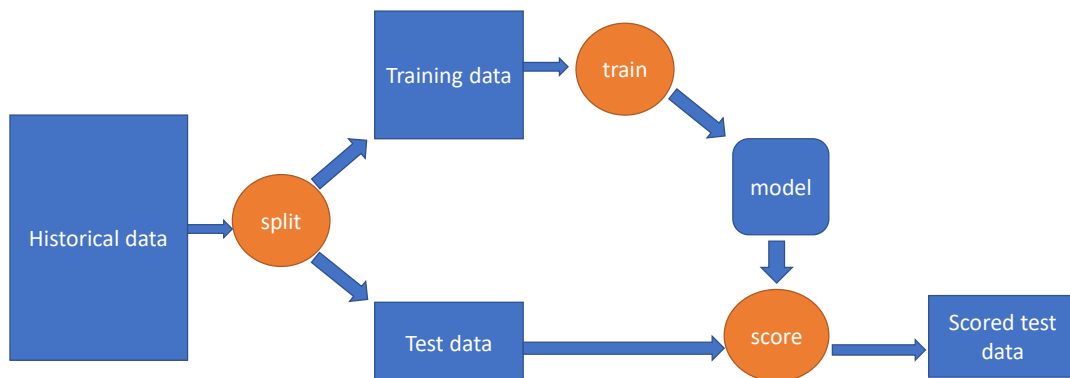
```
EXECUTE sp_execute_external_script
@language = N'R'
, @script = N'
  iris.sub <- c(sample(1:50, 25), sample(51:100, 25), sample(101:150, 25))
  iris.dtree <- rxDTree(Species ~ Sepal.Length + Sepal.Width + Petal.Length + Petal.Width, data=iris[iris.sub,])
  model <- rxSerializeModel(iris.dtree, realtimeScoringOnly = TRUE)
, @params = N'@model varbinary(max) OUTPUT'
, @model = @model OUTPUT
INSERT [dbo].[ml_models]([model_name], [model_version], [native_model_object])
VALUES('iris.dtree', 'v1', @model) ;
```

```
DECLARE @model varbinary(max) = (
  SELECT native_model_object
  FROM ml_models
  WHERE model_name = 'iris.dtree'
  AND model_version = 'v1');
SELECT flower.*, prediction.*
FROM PREDICT(MODEL = @model, DATA = dbo.iris_rx_data as flower)
WITH(setosa_Pred float, versicolor_Pred float, virginica_Pred float) as prediction
ORDER BY [Sepal.Width]
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species	setosa_Pred	versicolor_Pred	virginica_Pred
5	2	3.5	1	versicolor	0	1	0
6	2.2	4	1	versicolor	0	1	0
6.2	2.2	4.5	1.5	versicolor	0	1	0
6	2.2	5	1.5	virginica	0	0.0384615384615385	0.961538461538462

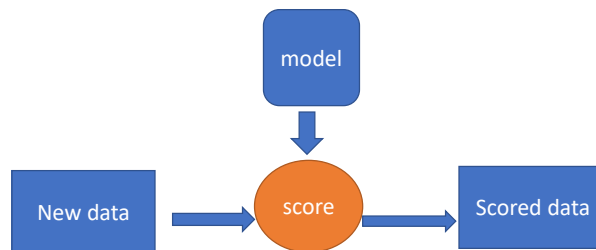
23

## Training a model



24

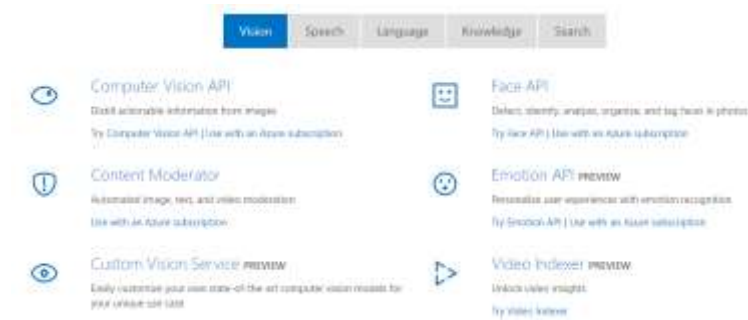
## Using a model



25

## Cognitive Services

- Cognitive services provides web services hosted in Microsoft Azure to convert text, sound, photos, ...into an easier to analyze format
  - Usually json documents are returned



26

## Two types of services

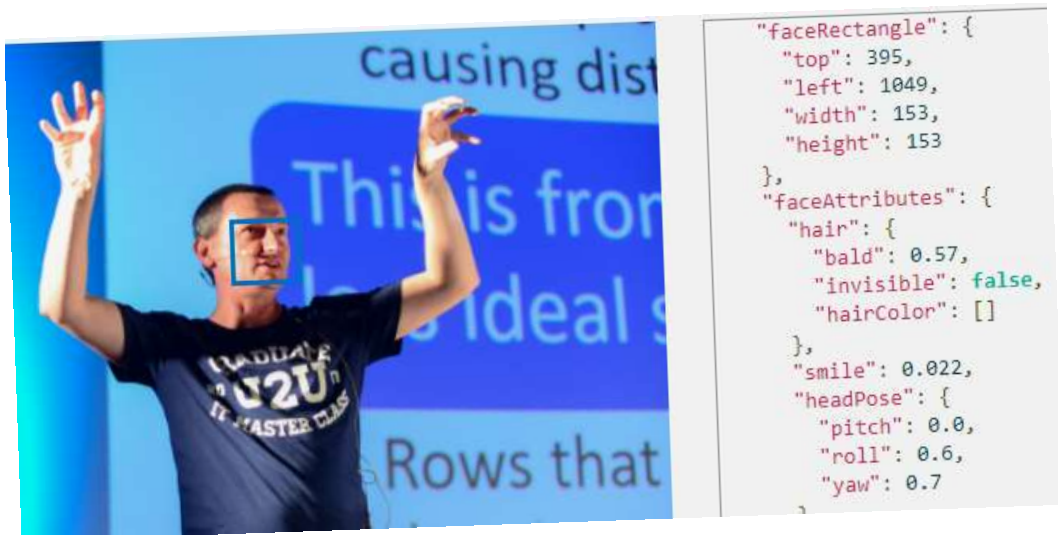
### Pre-trained

- We just provide the input and get back a result

### Transfer learning

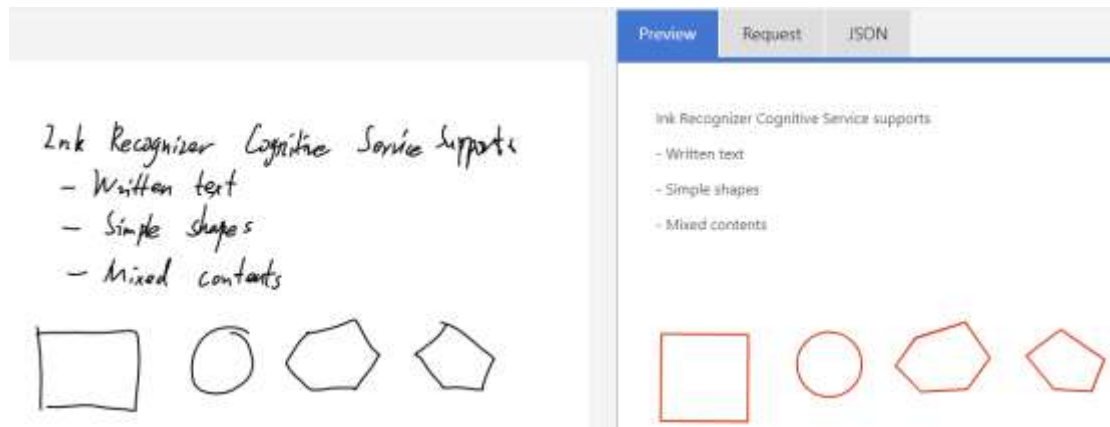
- Microsoft provides a framework within which we can further train or customize the service

27



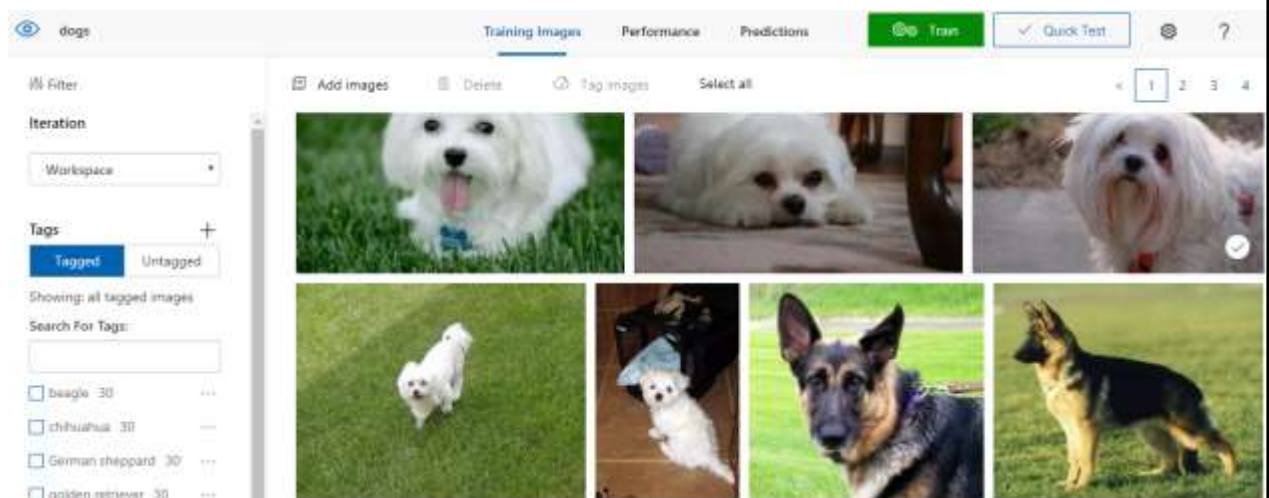
28

## Ink recognizer



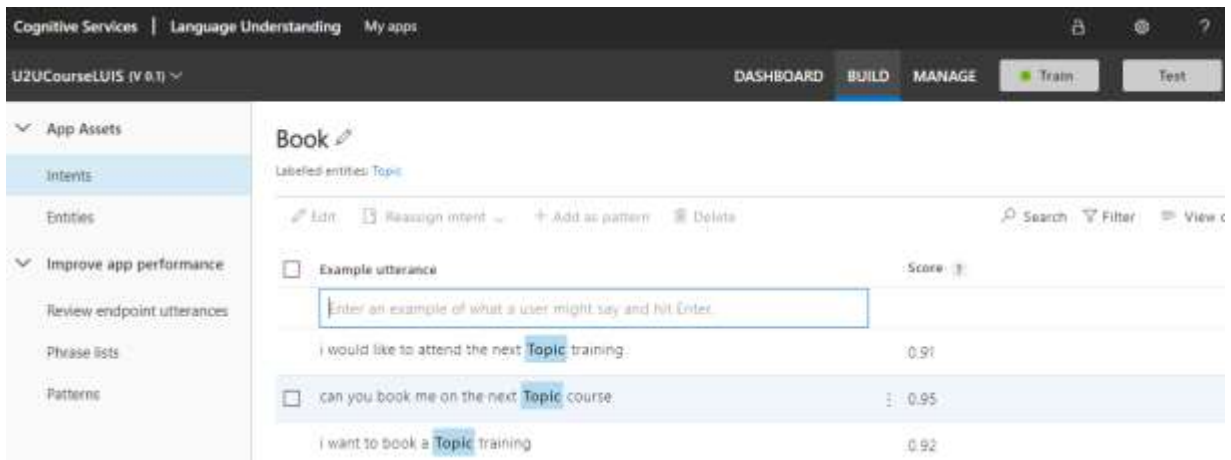
29

## Transfer learning: custom vision



30

# Transfer learning: LUIS



31

# Azure Machine Learning Studio

- Azure ML Studio starts from <https://studio.azureml.net>
- It comes with a graphical interface in which users develop a data mining workflow
- The models resulting from these flows can then be converted into an Azure web service for scoring new data



32



## Demo: ML Studio

Titanic data set



33

## 3 ways to use Azure ML Services

### Do It Yourself

- Ideal if you want to write or reuse Python code
- Scikit-learn, Tensorflow, PyTorch, Keras,...

### Graphical approach

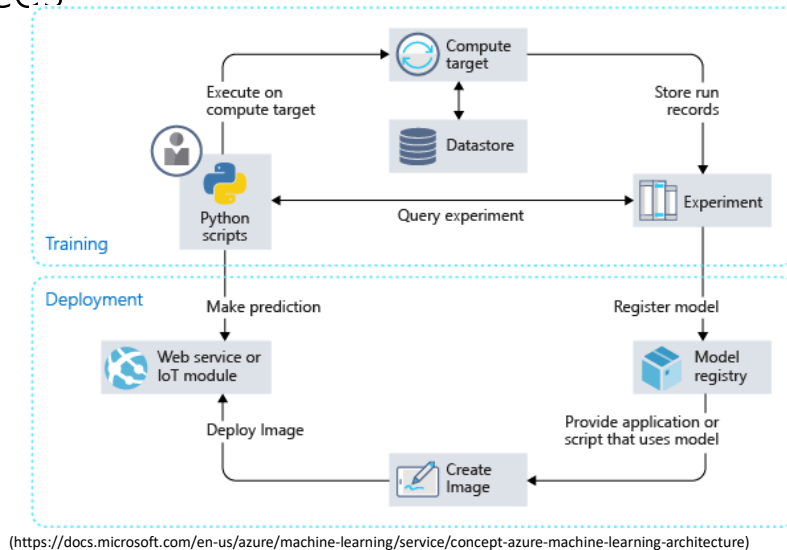
- Like Azure Machine Learning Studio, where a model is created with a drag-and-drop interface

### Automated ML

- Automated ML uses a trial-and-error approach to build a reasonable model

34

# Architecture Azure Machine Learning Services

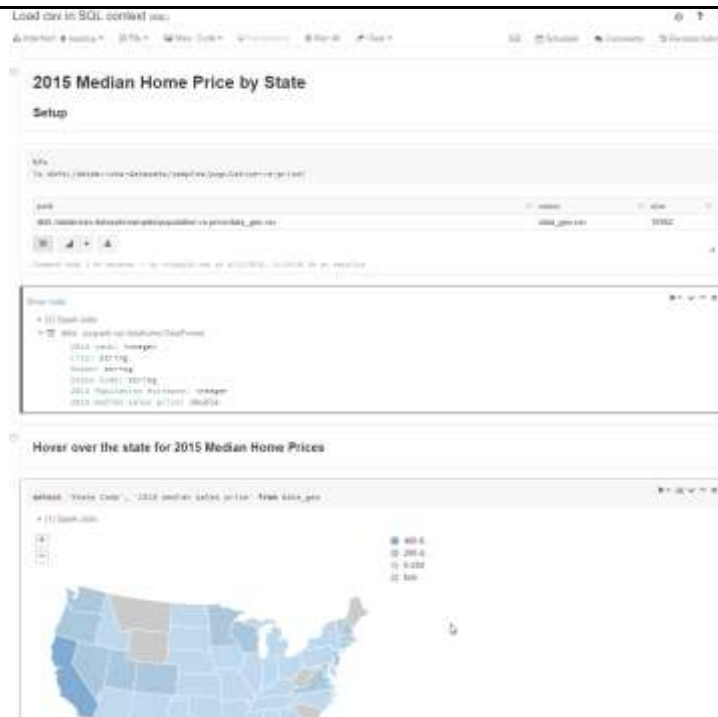


35

## Demo: Azure ML Services

36

- Designed by Spark developers
- Collaborative platform
- 10 times faster than vanilla spark
- Easy setup and administration
- Native integration with Azure services
  - Power BI
  - SQL Data Warehouse
  - Cosmos Db
  - Blob Storage
- Azure Active Directory integration
- Enterprise grade SLA



Microsoft  
tech·days  
Kistamässan Stockholm  
22-24 oktober 2019

Nico Jacobs  
@SQLWaldorf  
Nico@u2u.be

# Thank you!

Please take the opportunity to evaluate this session in the app

Download slidedeck from