

Azure Machine Learning (Data Camp)

Mostafa Elzoghbi
Sr. Technical Evangelist – Microsoft
@MostafaElzoghbi
<http://mostafa.rocks>

Session Objectives & Takeaways

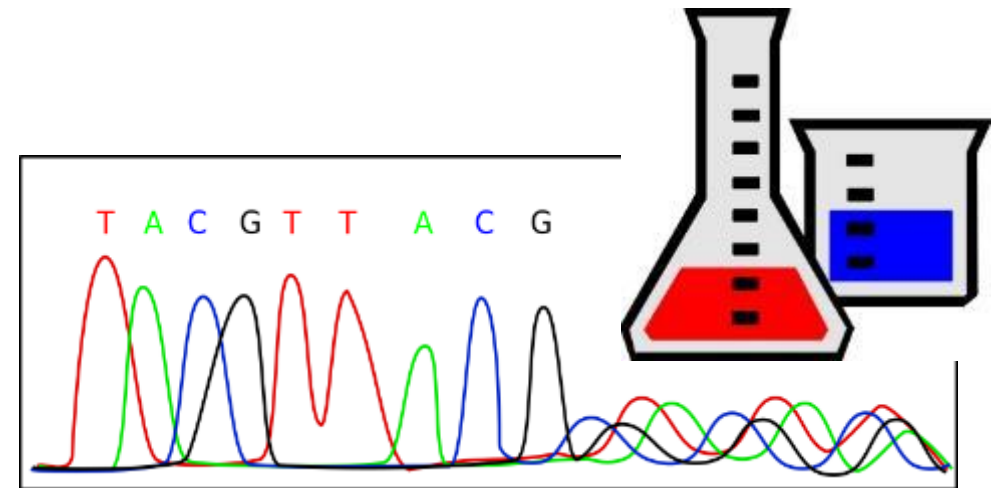
- What is Data Science?
- What is Machine Learning?
- Why do we need Machine Learning?
- Azure ML studio platform capabilities.
- Machine Learning basic concepts & algorithms:
 - How to build a model
 - Supervised vs Unsupervised models
 - Evaluate a model.
 - ML algorithms: Regression, classification, clustering and recommendations.

Data Science Involves

- Data science is about using **data** to **make decision** that **drive actions**.
- Data science process involves:
 - Data selection
 - Preprocessing
 - Transformation
 - Data Mining
 - Delivering value from data : Interpretation and evaluation

Data Science

- Data Science is far too complex
 - Cost of accessing/using efficient ML algorithms is high
 - Comprehensive knowledge required on different tools/platforms to develop a complete ML project
 - Difficult to put the developed solution into a scalable production stage
- Need a simpler/scalable method:
Azure Machine Learning Service



What is Machine Learning ?

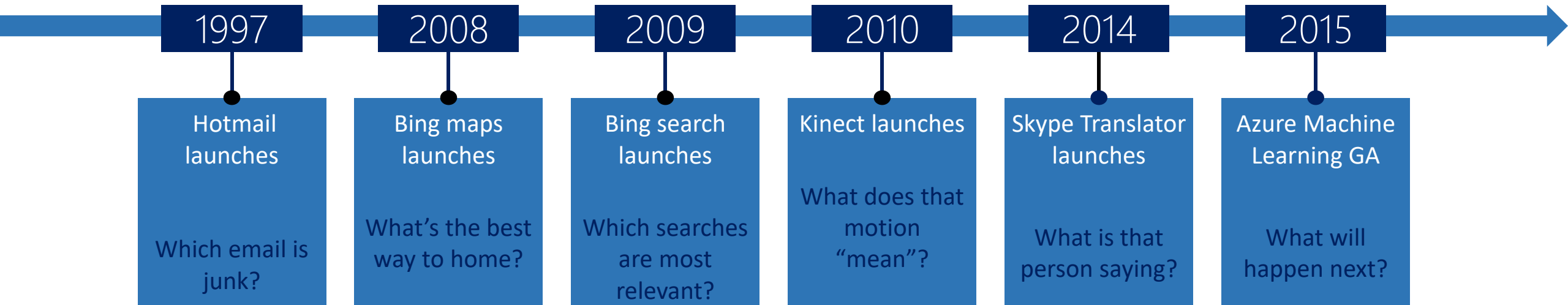
- Using **known data**, develop a **model** to predict **unknown data**.

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

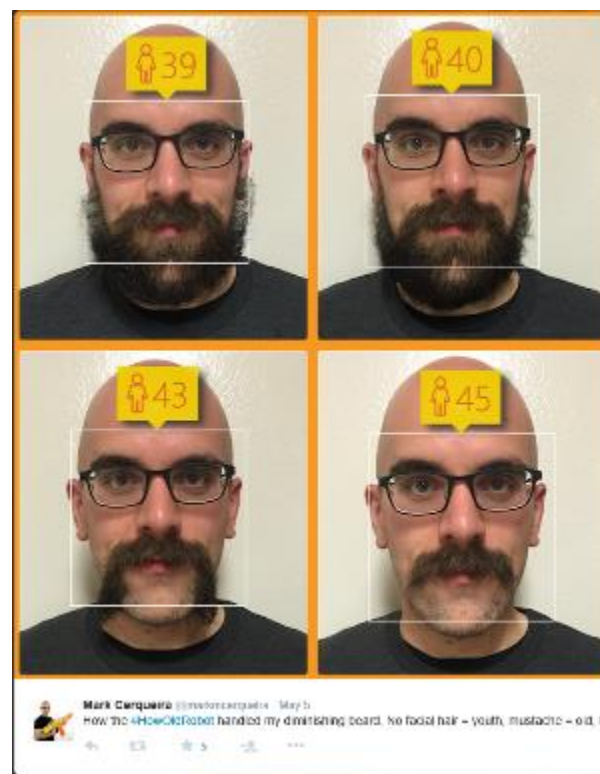
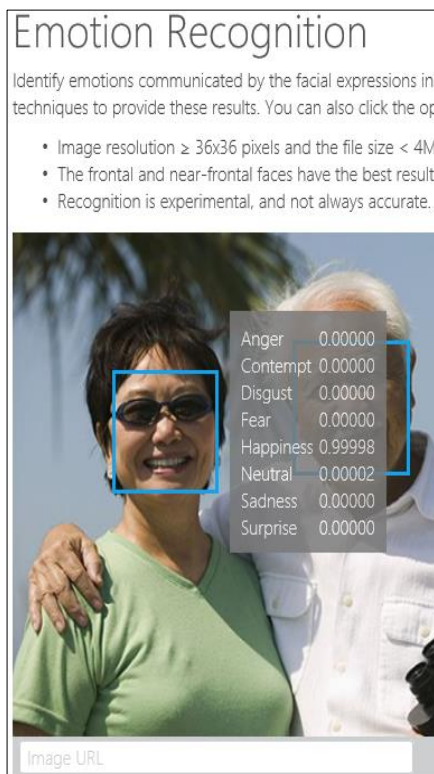
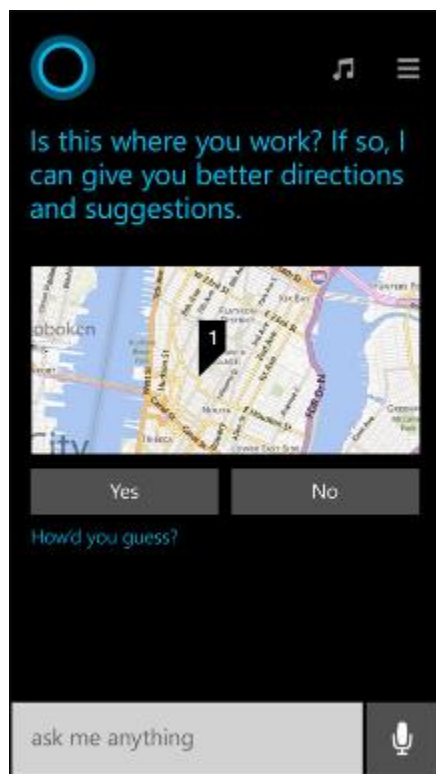
Model: Known data + Algorithms (ML algorithms)

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

Microsoft & Machine Learning



Why Machine Learning?





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.

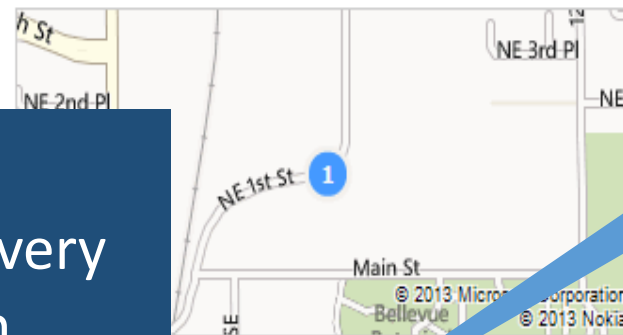
[Seattle's Best Mortgage](#)

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?

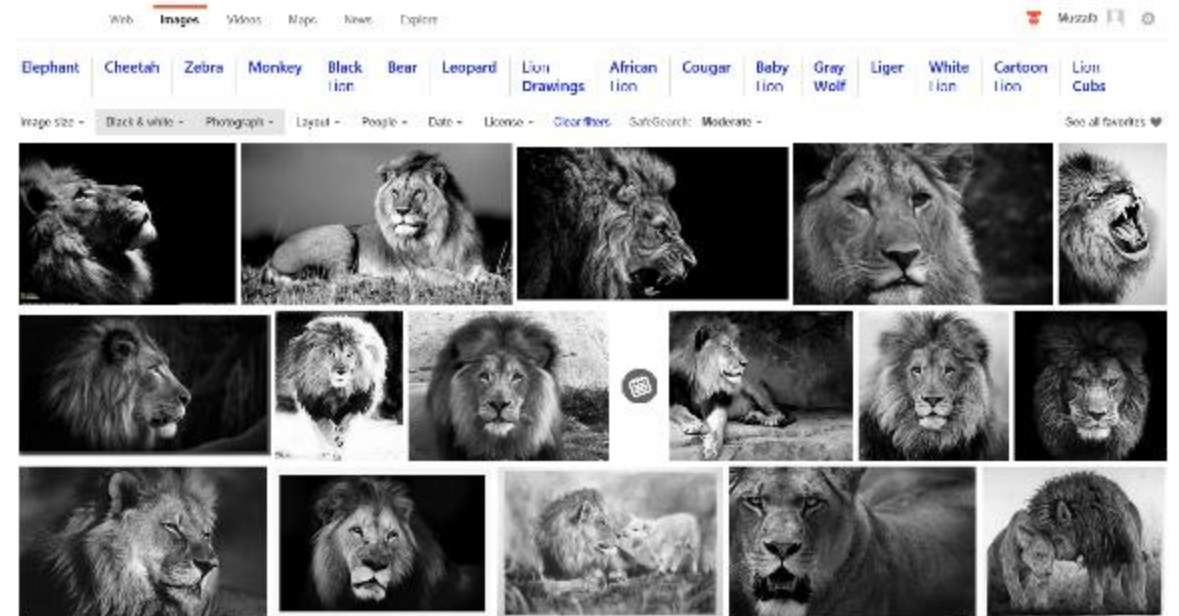
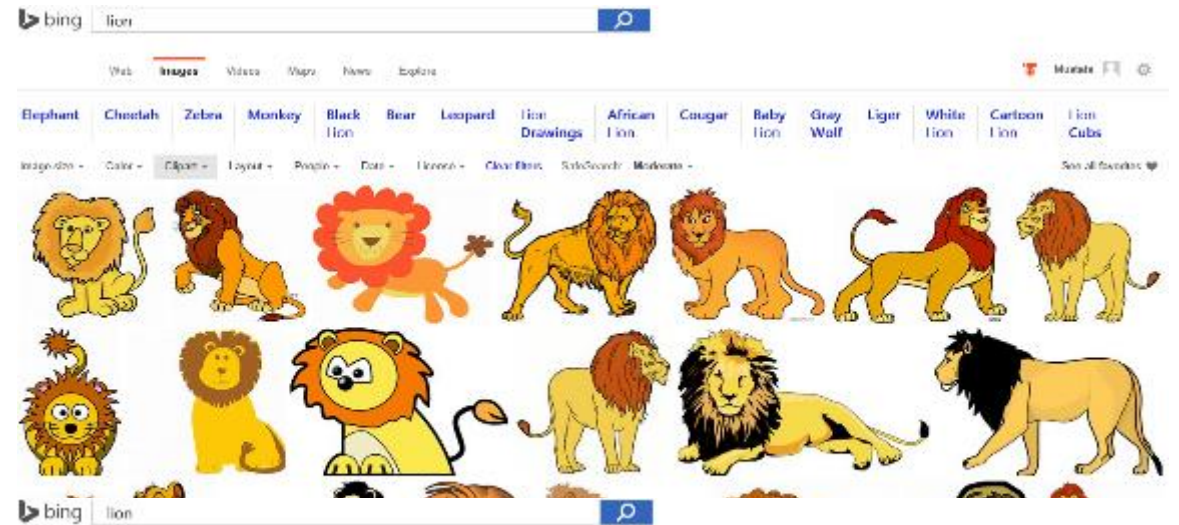
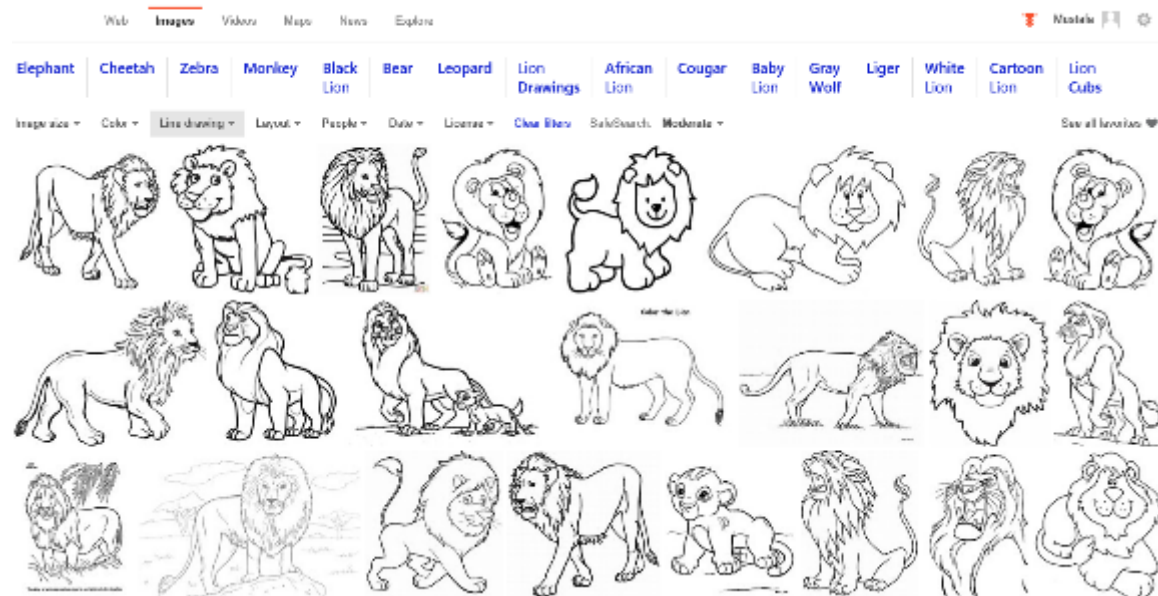
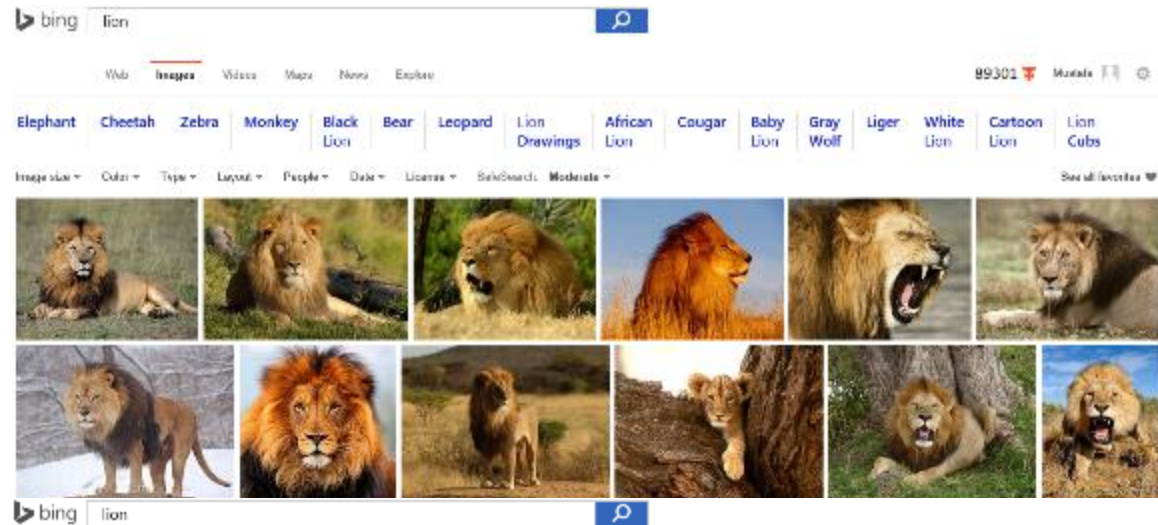
Are any of these pages malicious?

What pages should we index?

What ad pricing will optimize revenue?

Machine learning enables nearly every value proposition of web search.

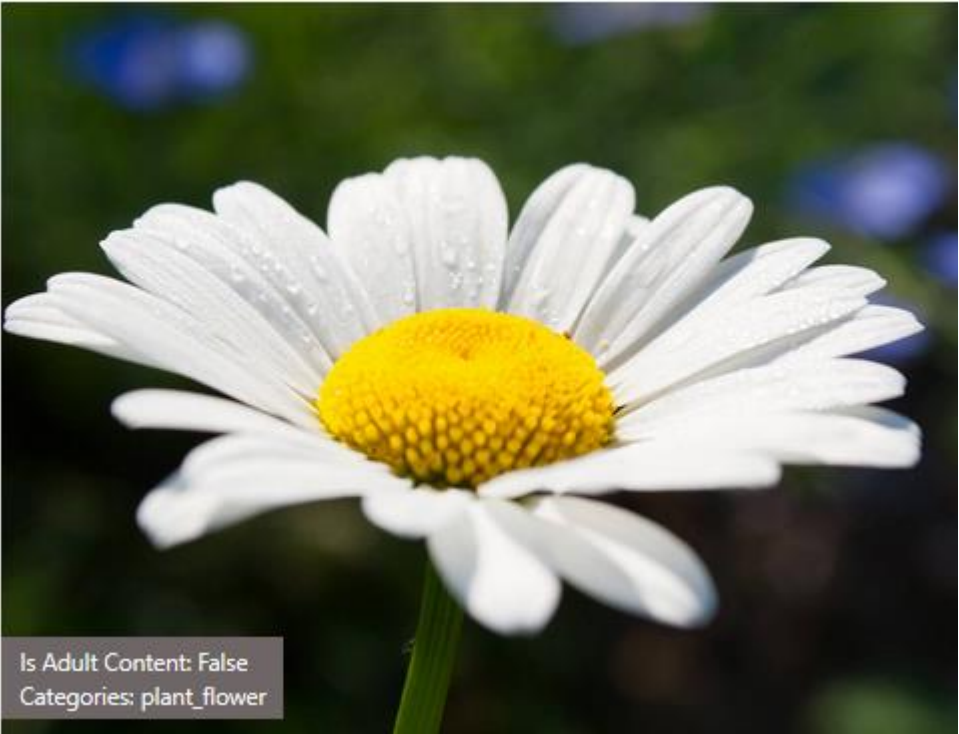
Image Analyze









Accent Color: Which border color is the best?



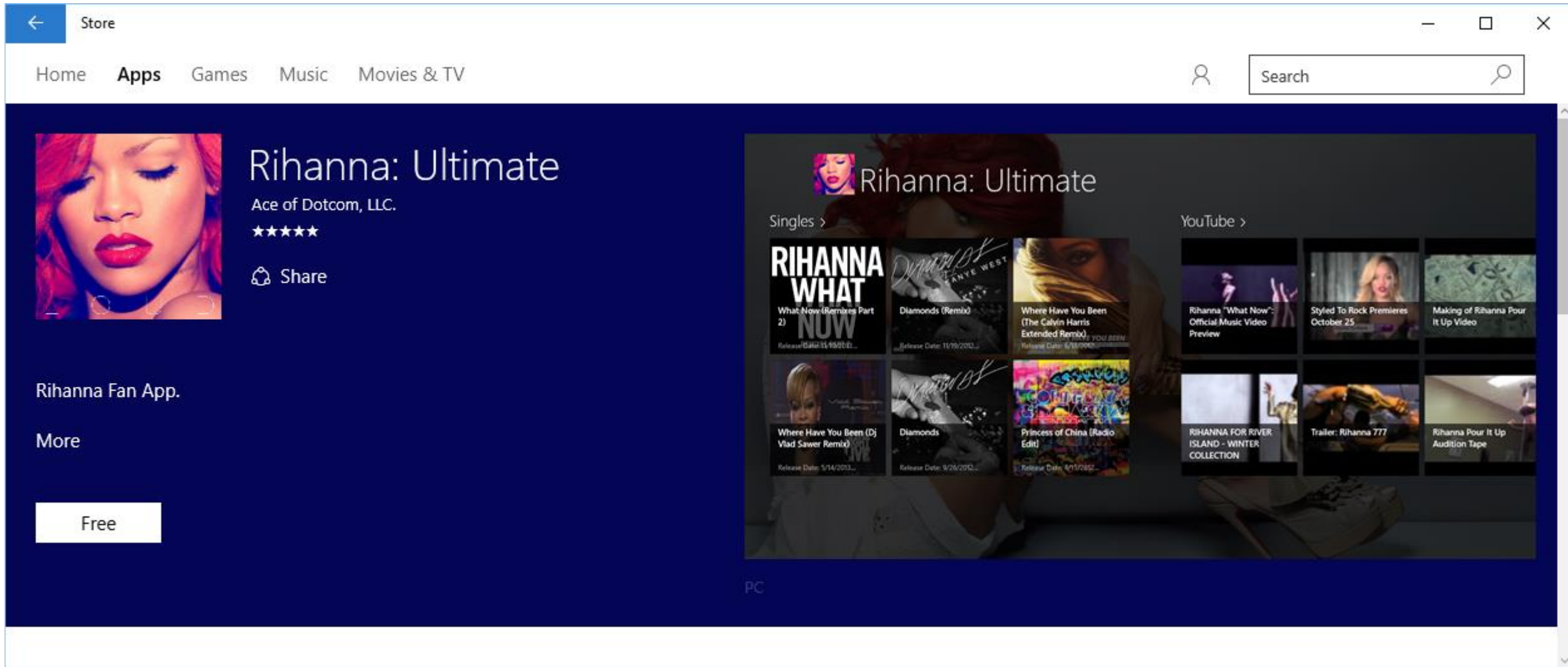
Accent Color: Analyze Image



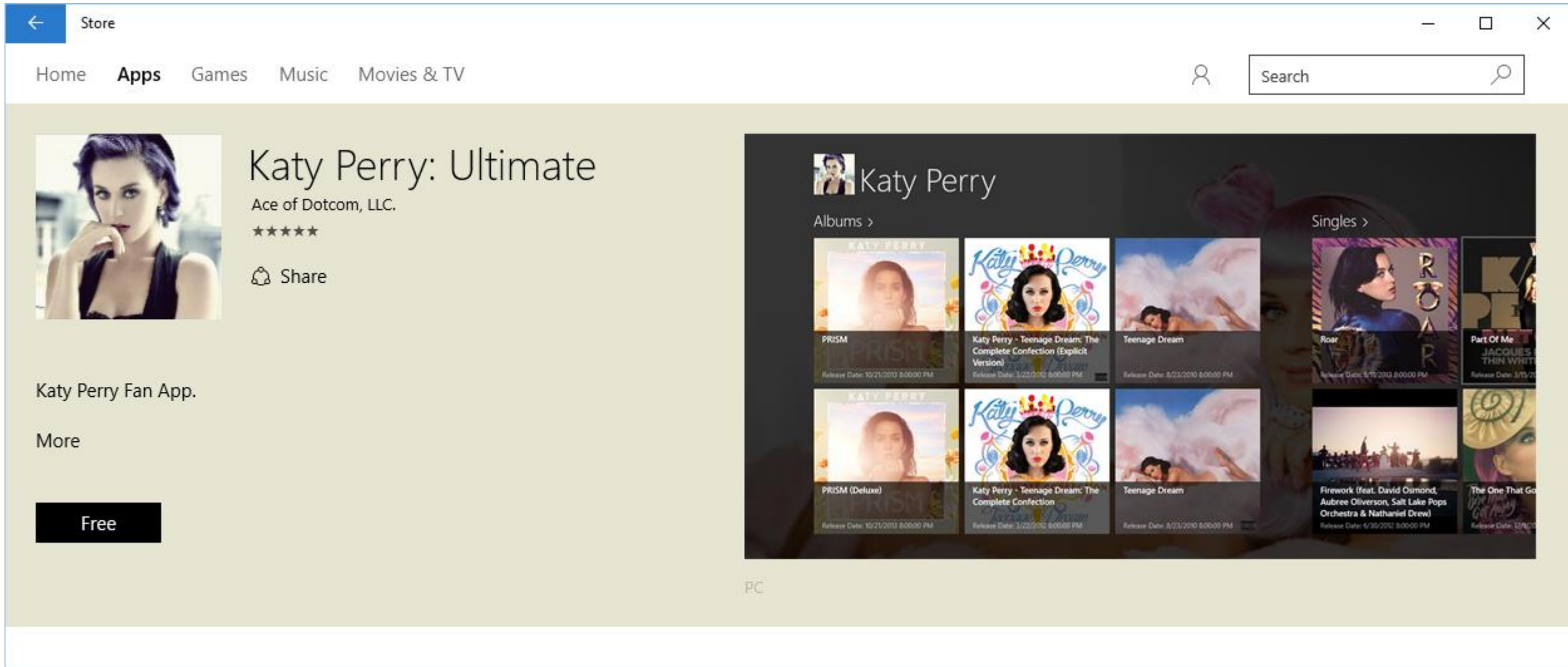
Features:

Feature Name	Value
Image Format	Jpeg
Image Dimensions	1500 x 1155
Clip Art Type	0 Non-clipart
Line Drawing Type	0 Non-LineDrawing
Black & White Image	False
Is Adult Content	False
Adult Score	0.02942577563226223
Is Racy Content	False
Racy Score	0.018232977017760277
Categories	[{ "name": "plant_flower", "score": 0.99609375 }]
Faces	[]
Dominant Color Background	
Dominant Color Foreground	
Dominant Colors	  
Accent Color	 #C8A403

Accent Color: Windows 10 Store



Accent Color: Windows 10 Store



Text Analytics: User reviews

Positive

Text Analytics - Preview by [Azure Machine Learning](#) Demo Sample Code Doc

Try out the Azure ML [Text Analytics service](#) - for free.

To use the service in production, you can get access to its API by [signing up](#) for it (click on "Sign Up" on the right pane on that page). For questions/comments, please use the "Feedback" button in the top right.

I love this presentation. I found it very useful.

Analyze

Sentiment:

98 %

Key phrases highlighted below:

I love this **presentation**. I found it very useful.

Negative

Text Analytics - Preview by [Azure Machine Learning](#) Demo S

Try out the Azure ML [Text Analytics service](#) - for free.

To use the service in production, you can get access to its API by [signing up](#) for it (c the right pane on that page). For questions/comments, please use the "Feedback" b

I hate this presentation. I found it very useless.

Analyze

Sentiment:

7 %

Key phrases highlighted below:

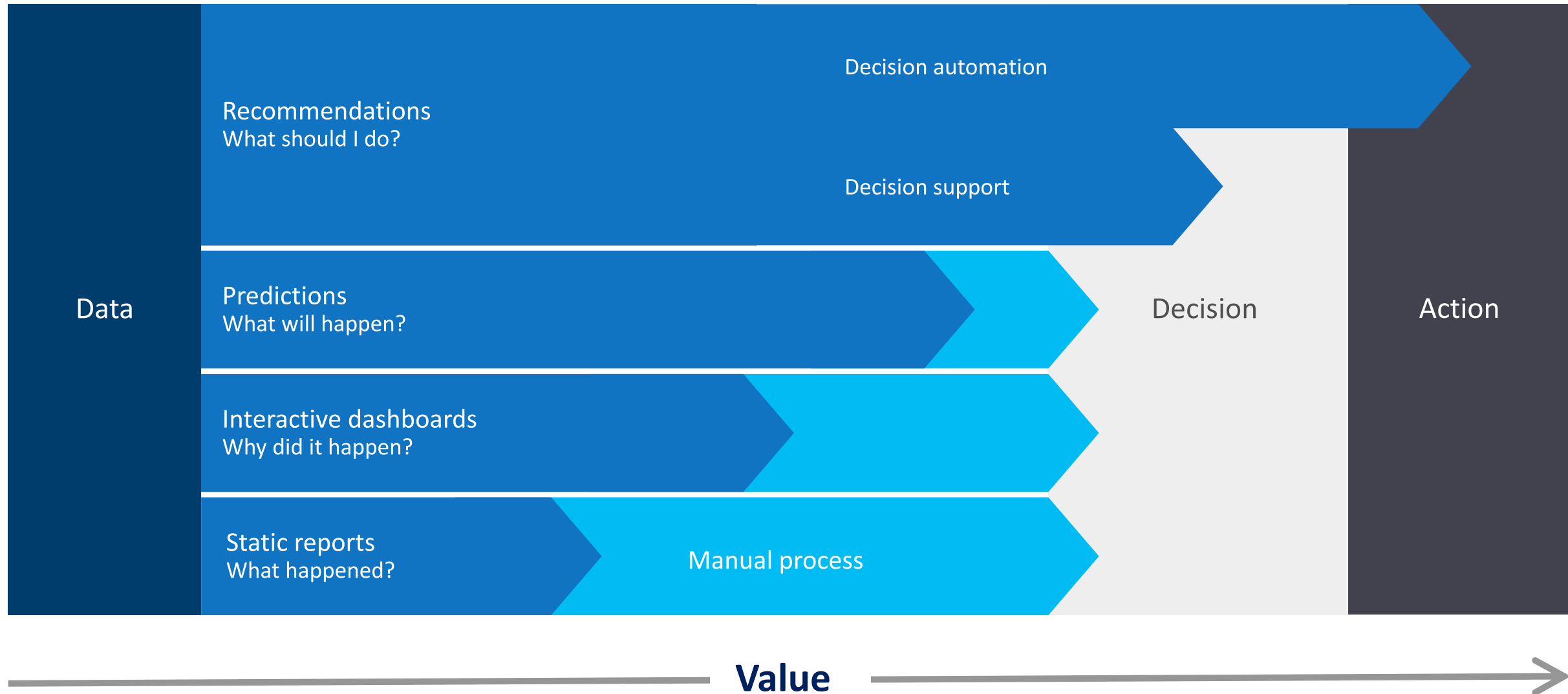
I hate this **presentation**. I found it very useless.

Microsoft Azure Machine Learning

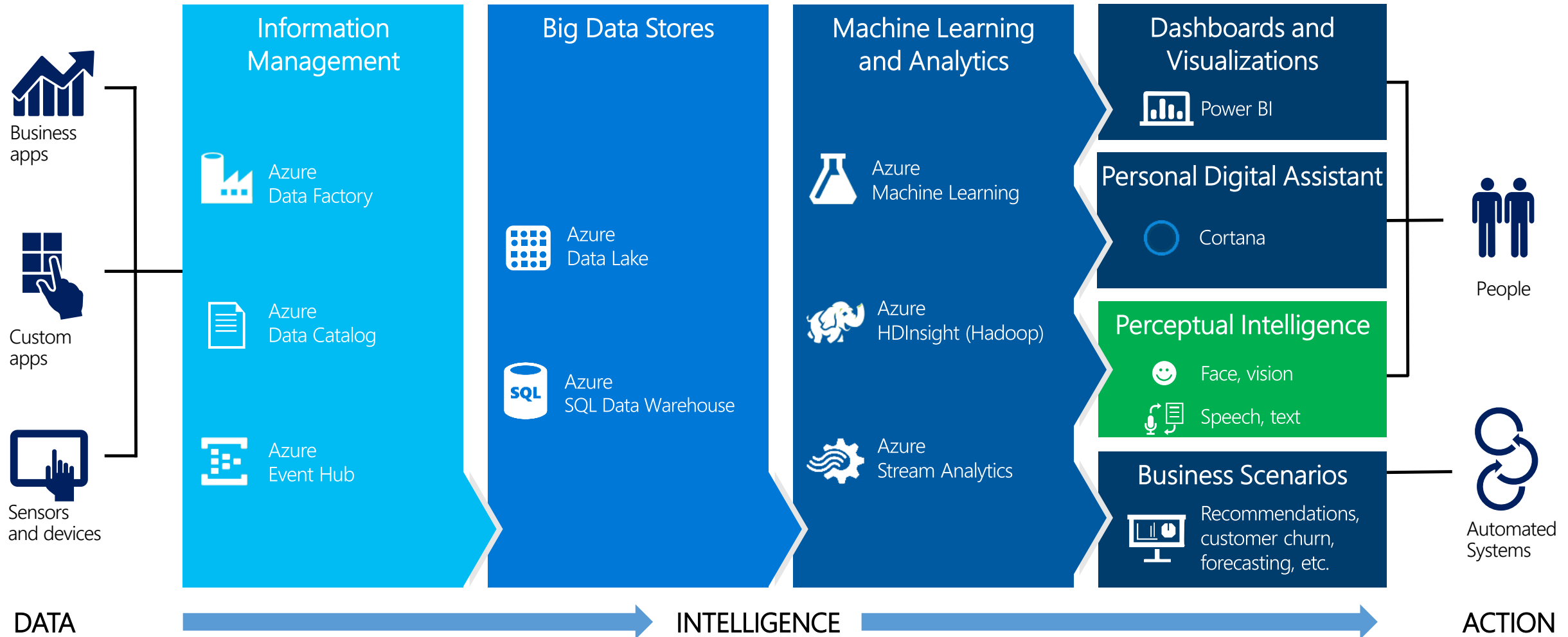
Make machine learning accessible to every enterprise, data scientist, developer, information worker, consumer, and device anywhere in the world.



From data to decisions and actions

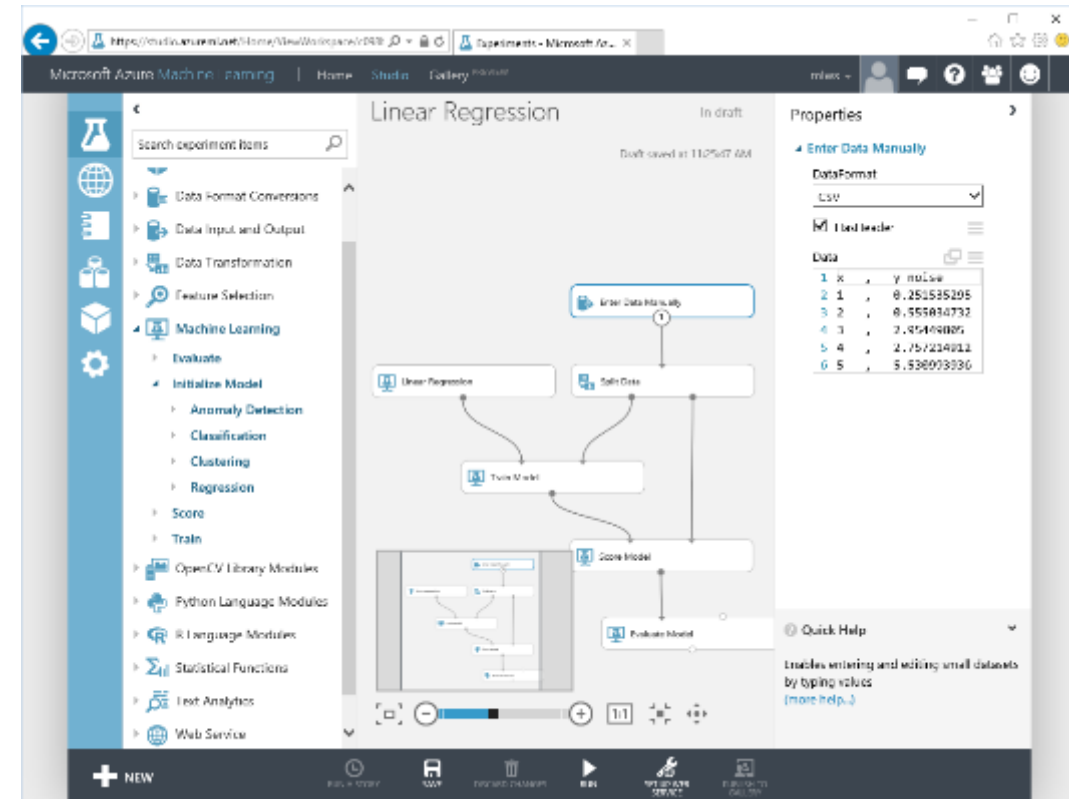


Transform data into intelligent action



Microsoft Azure Machine Learning

- Web based UI accessible from different browsers
- Share|collaborate to any other ML workspace
- Drag & Drop visual design|development
- Wide range of ML Algorithms catalog
- Extend with OSS **R** | **Python** scripts
- Share|Document with IPython|Jupyter
- Deploy|Publish|Scale rapidly (**APIs**)



Microsoft Azure Machine Learning

Built for a cloud-first, mobile-first world

Fully managed

No software to install, no hardware to manage, all you need is an Azure subscription.

Integrated

Drag, drop and connect interface. Data sources with just a drop down; run across any data.

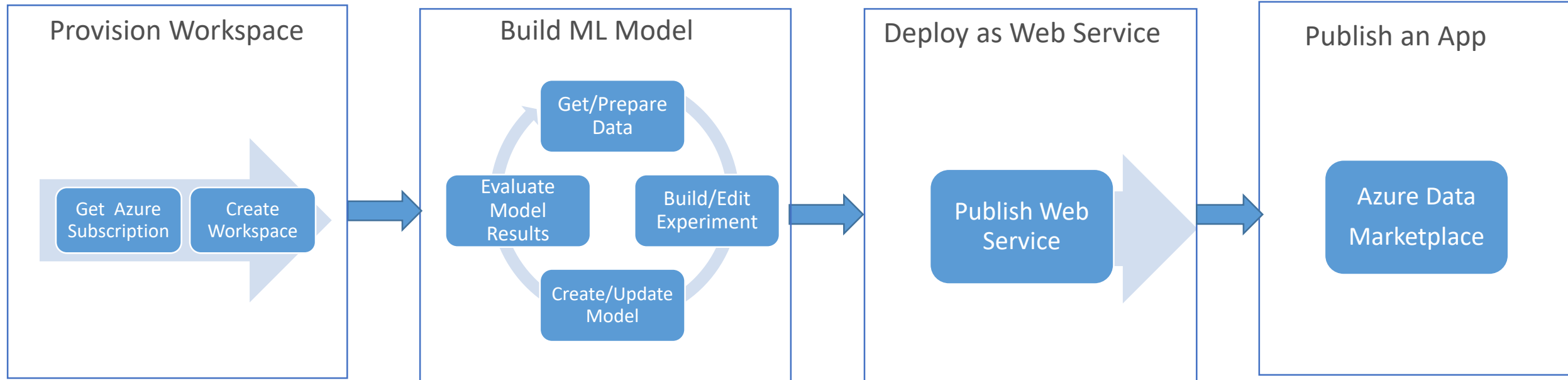
Flexible

Built-in collection of best of breed algorithms with no coding required. Drop in custom R or use popular CRAN packages.

Deploy in minutes

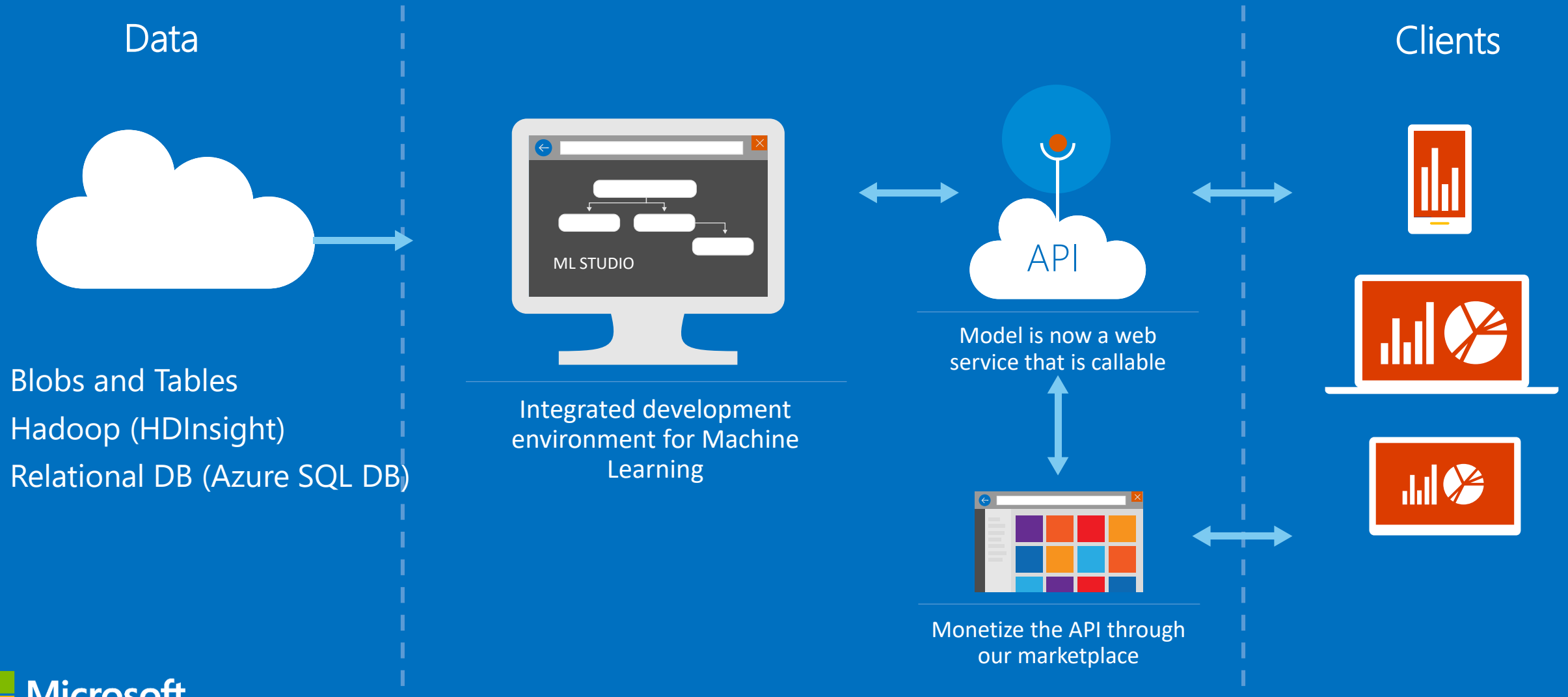
Operationalize models as web services with a single click. Monetize in Machine Learning Marketplace.

Azure Machine Learning Ecosystem



Azure Machine Learning Service

Data -> Predictive model -> Operational web API in minutes

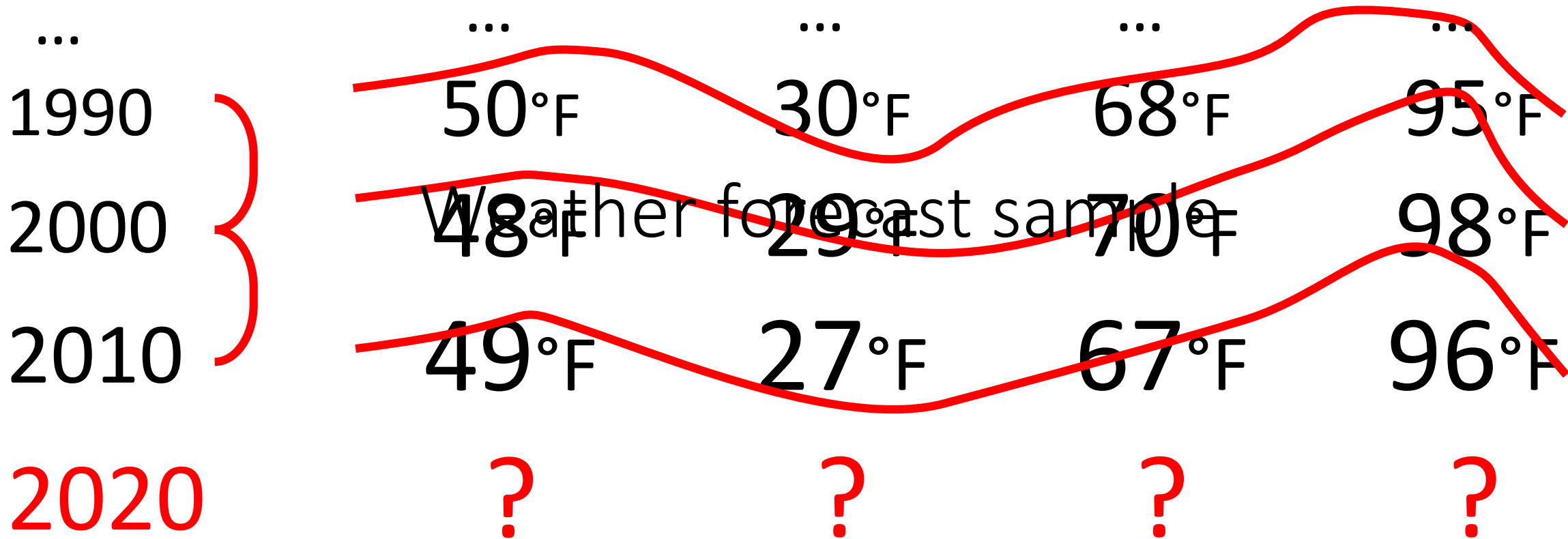
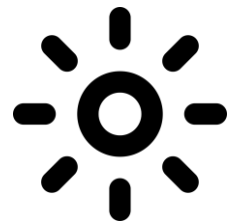


DEMO

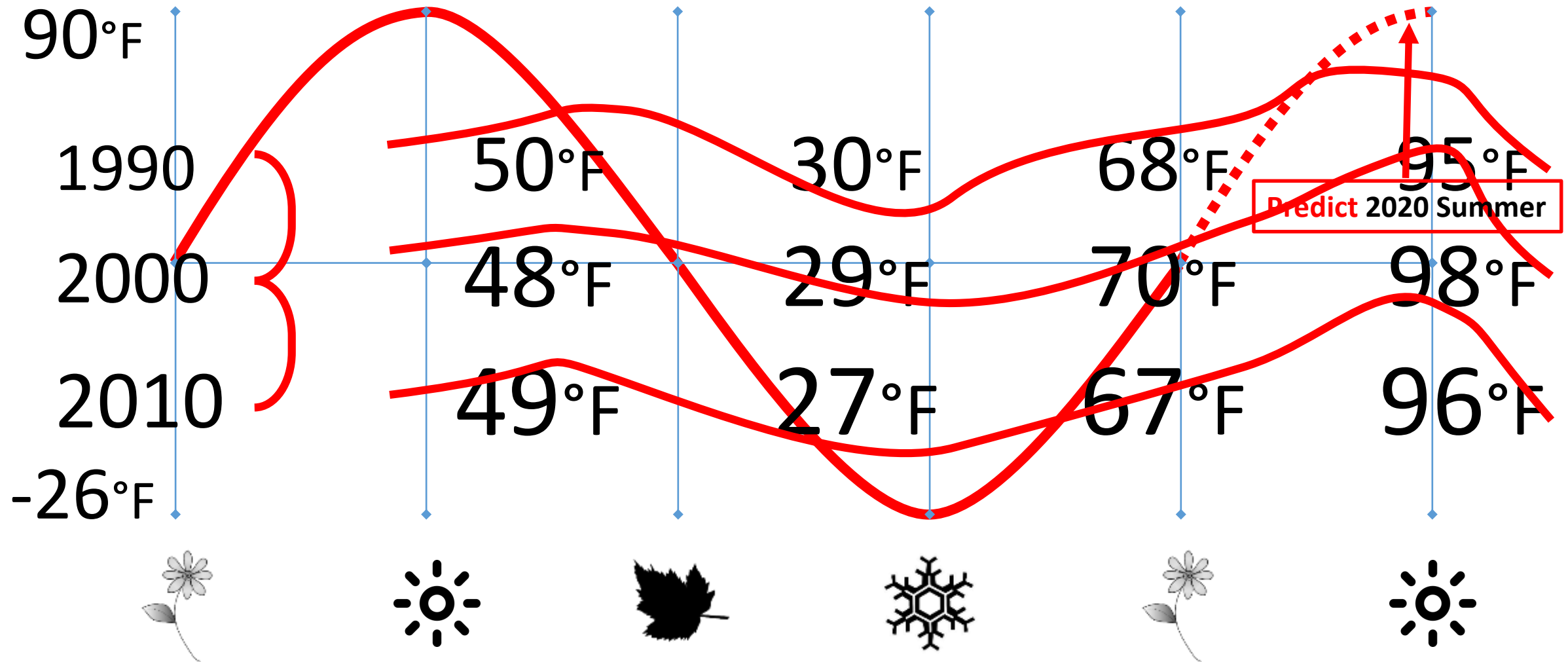
Azure Machine Learning Studio

EXAMPLE

Known data
Model
Unknown data



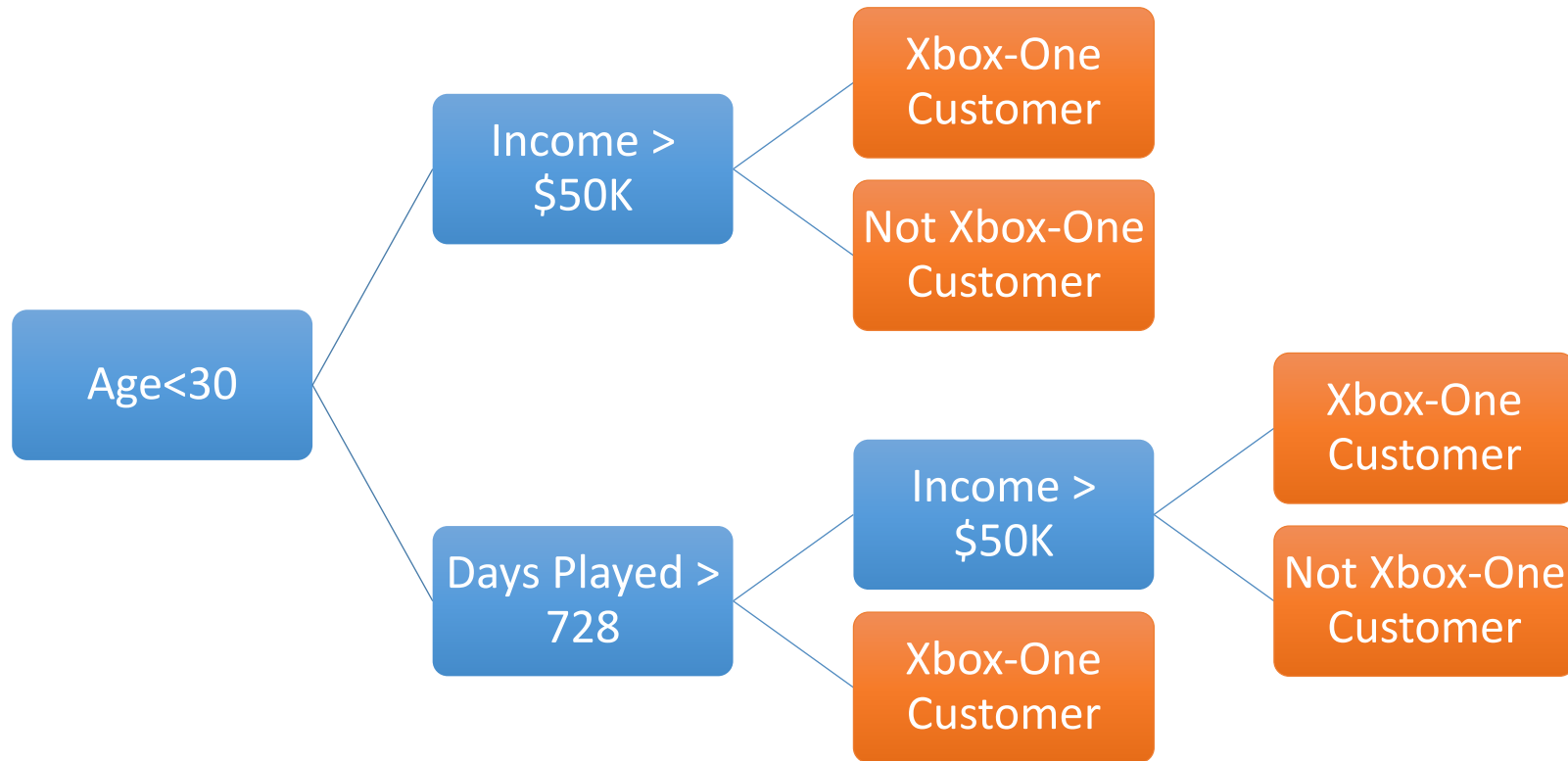
Model (Regression)



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

EXAMPLE

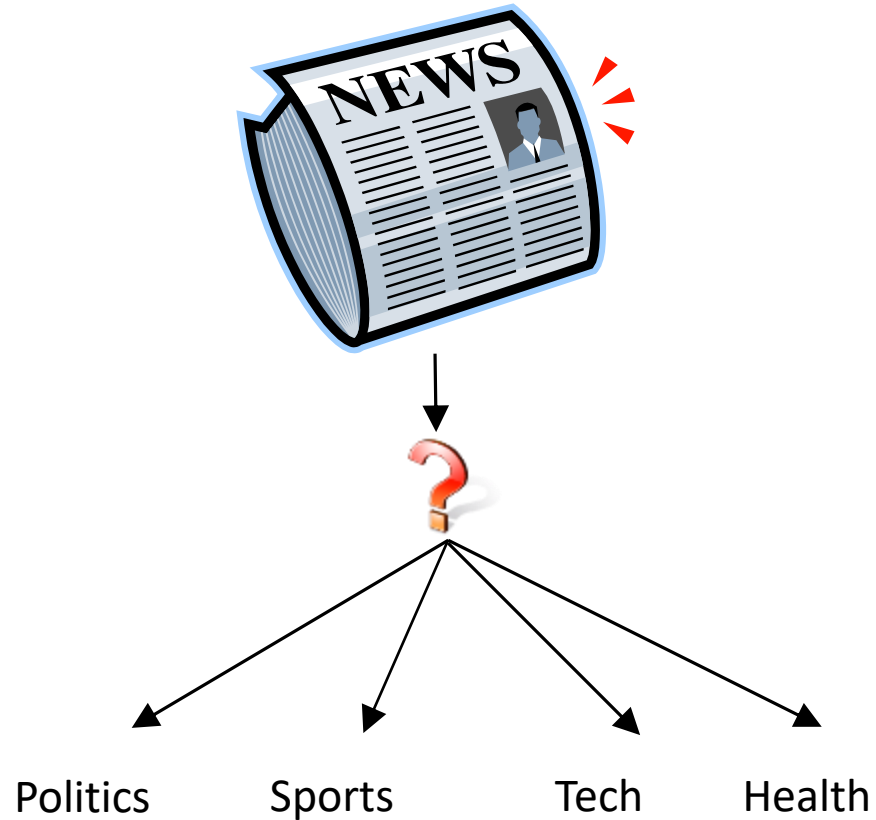
Model (Decision Tree)



EXAMPLE

Model (Classification)

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



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



Tech



Health



Politics



Politics



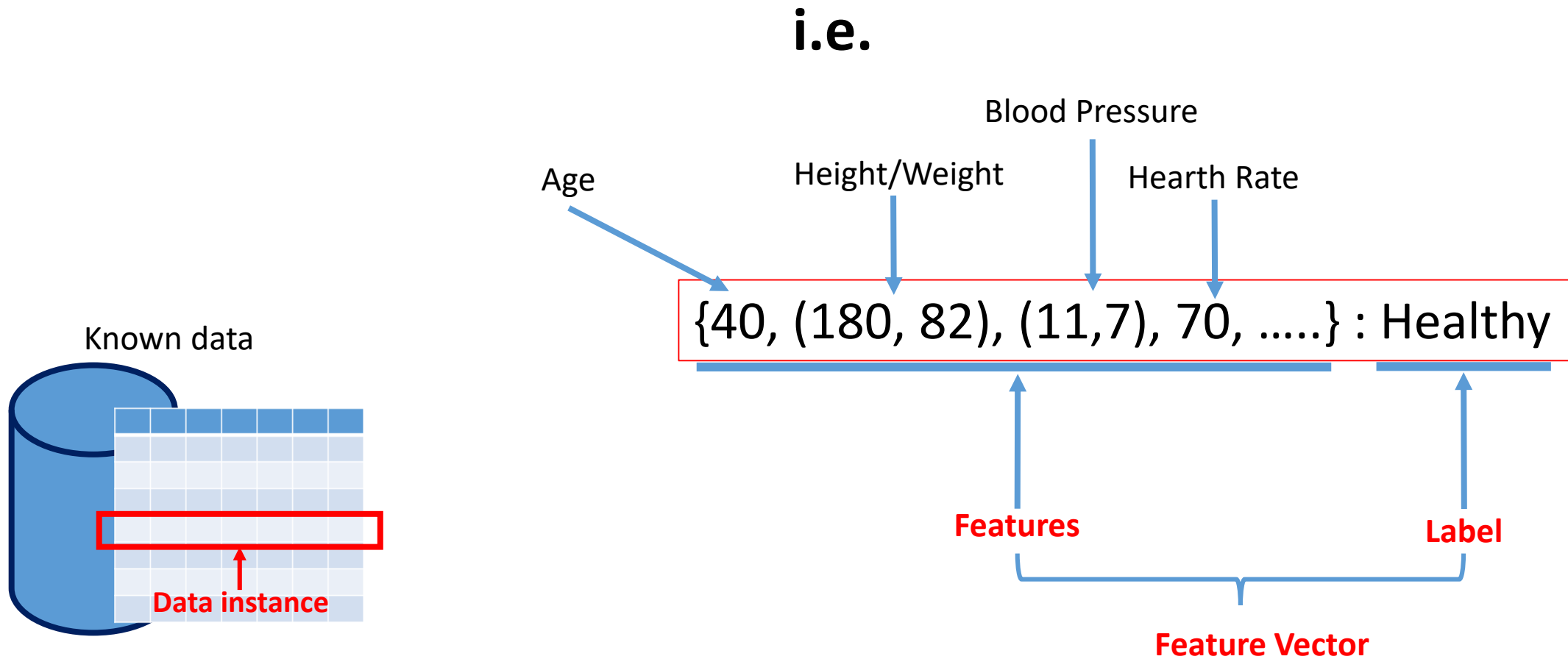
Sports

Feature

[illegible]

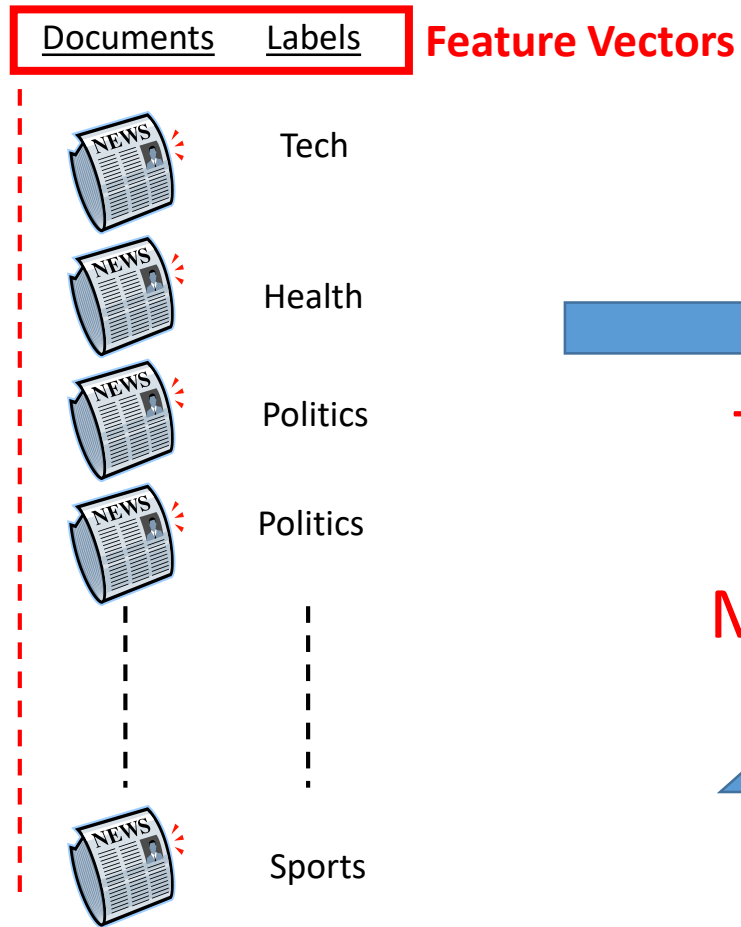
Using **known data**, develop a **model** to predict **unknown data**.

Feature vector



Developing a Model

Training data



Train
the
Model



Base
Model

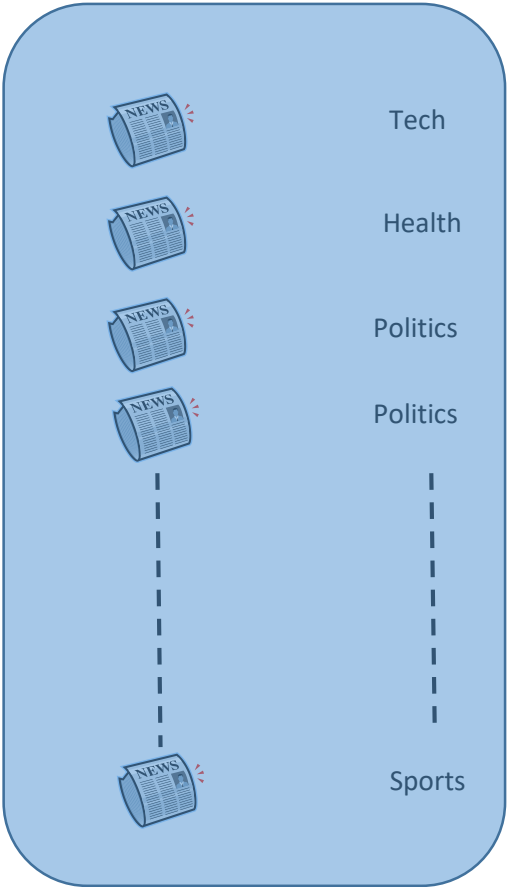


Adjust
Parameters

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

Model's Performance

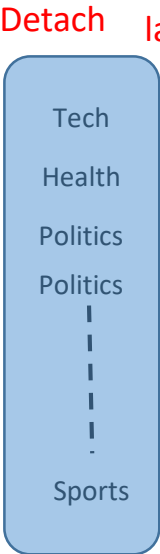
Known data with **true labels**



Split

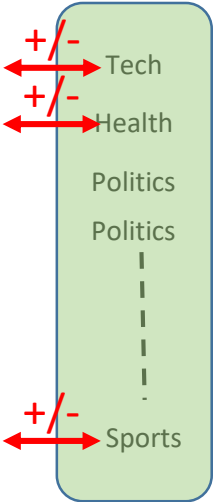
Test data 20%

Training data 80%



True labels

Predicted labels



Test trained model with features

Compare prediction with true labels

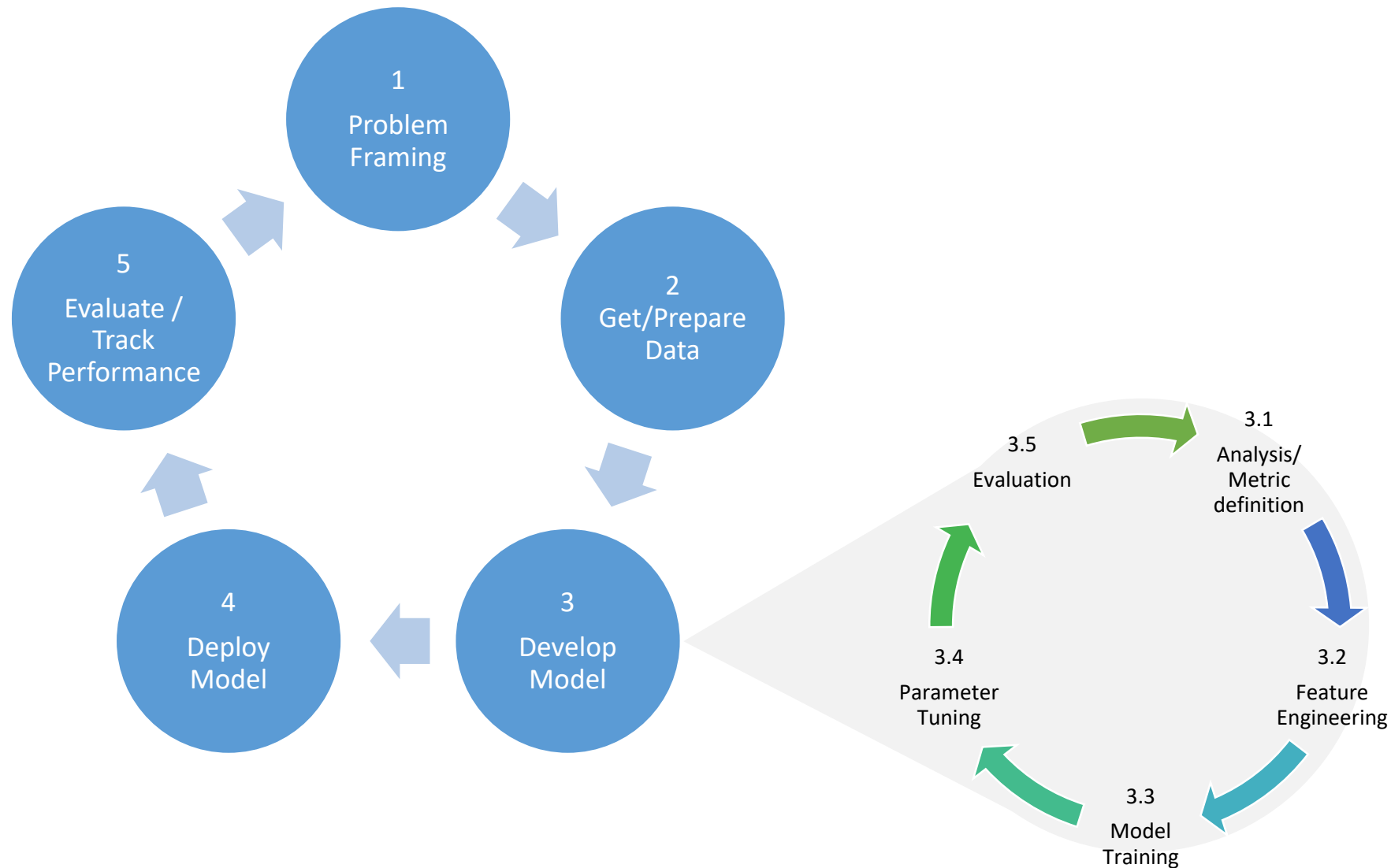
Train the Model



Model's Performance

Difference between "True Labels" and "Predicted Labels"

Steps to Build a Machine Learning Solution



Example use cases

 Finance and risk	 Sales and marketing	 Customer and channel	 Operations and workforce
 Revenue Forecasting	 Sales forecasting	 User segmentation	 Agent allocation
 Portfolio optimization	 Demand forecasting	 Personalized offers	 Warehouse efficiency
 Investment modelling	 Sales lead scoring	 Product recommendation	 Smart buildings
 Fraud detection	 Marketing mix optimization		 Predictive maintenance
 Risk management			 Supply chain optimization

Machine Learning Algorithms

- ML Algorithm defines how your **model** will react
- Which Algorithm to use? Depends on:
 - Data Quality
 - Data Size
 - What you want to predict
 - Time constraint
 - Computation power
 - Memory limits



Machine Learning Algorithms

You can develop solutions by using

- Custom algorithms written in R | Python
- Ready to use ML services from data market
- Existing algorithms

Machine Learning Algorithms

Two major category of algorithms

- Supervised
- Unsupervised

Most commonly used machine learning algorithms are **supervised** (requires **labels**)

- **Supervised** learning examples

- This customer will like *coffee*
- This network traffic indicates a denial of service attack

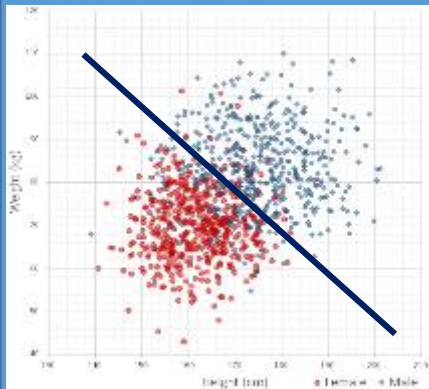
- **Unsupervised** learning examples

- These customers are similar
- This network traffic is unusual

Common Classes of Algorithms

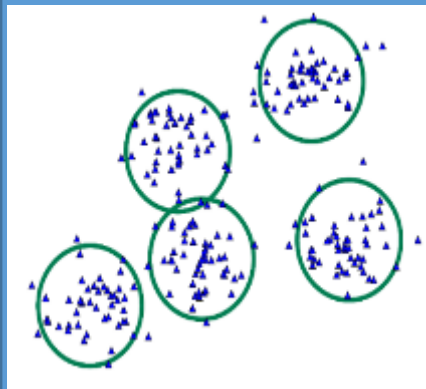
(Supervised|Unsupervised)

Classification



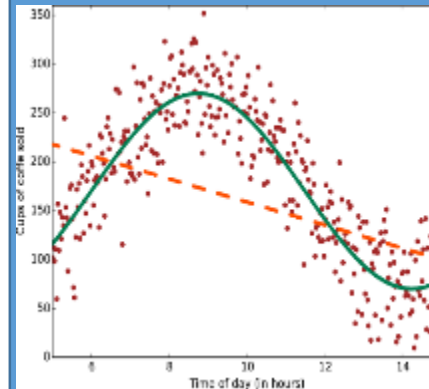
Supervised

Clustering



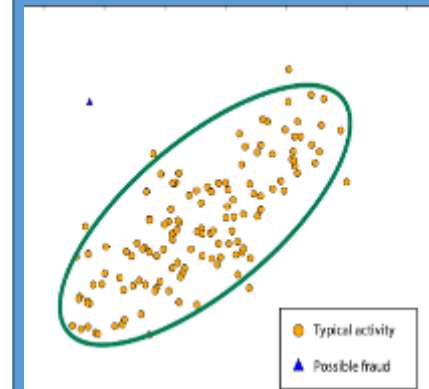
UnSupervised

Regression



Supervised

Anomaly Detection



Supervised

Why you need to know these algorithms?

- If you want to answer a **YES|NO** question, it is **classification**
- If you want to predict a **numerical value**, it is **regression**
- If you want to group data into similar observations, it is **clustering**
- If you want to recommend an item, it is **recommender system**
- If you want to find anomalies in a group, it is **anomaly detection**

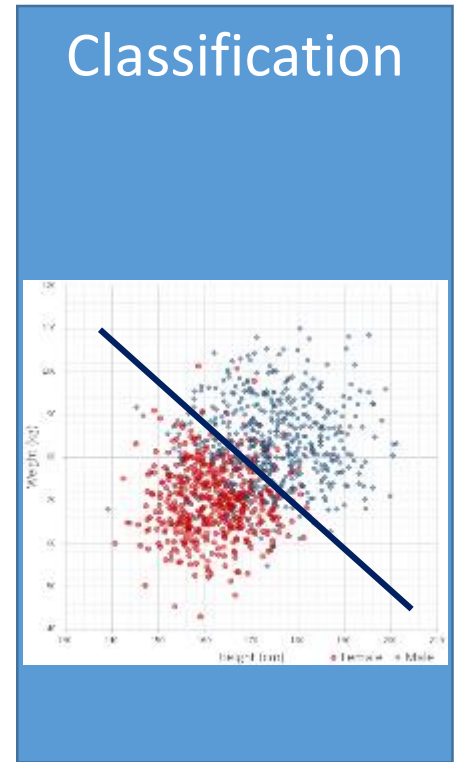
and many other ML algorithms for specific problem

Classification

Scenarios:

- Which customer are more likely to buy, stay, leave (churn analysis)
- Which transactions|actions are fraudulent
- Which quotes are more likely to become orders
- Recognition of patterns: speech, speaker, image, movement, etc.

Algorithms: Boosted Decision Tree, Decision Forest, Decision Jungle, Logistic Regression, SVM, ANN, etc.

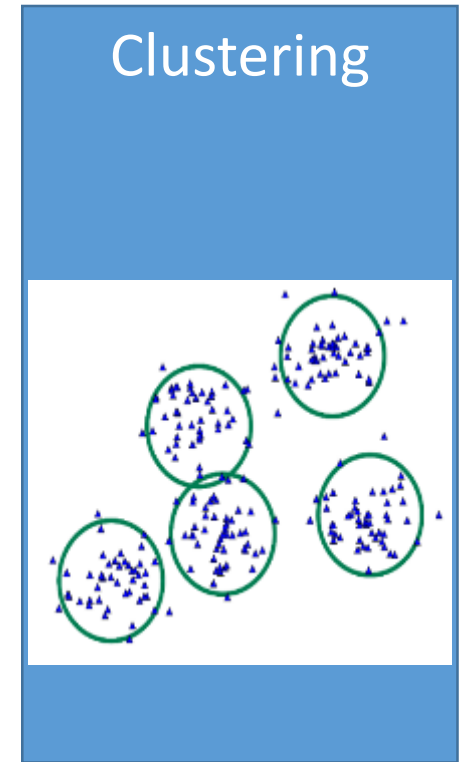


Clustering

Scenarios:

- Customer segmentation: divide a customer base into groups of individuals that are similar in specific ways relevant to marketing, such as age, gender, interests, spending habits, etc.
- Market segmentation
- Quantization of all sorts, such as, data compression, color reduction, etc.
- Pattern recognition

Algorithms: K-means

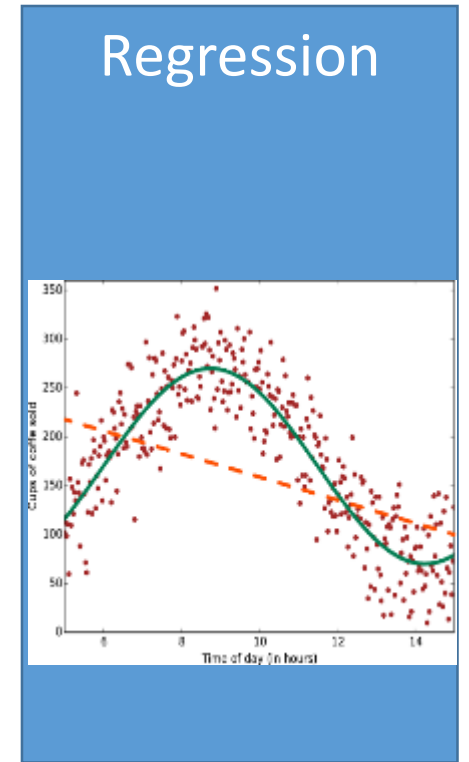


Regression

Scenarios:

- Stock prices prediction
- Sales forecasts
- Premiums on insurance based on different factors
- Quality control: number of complaints over time based on product specs, utilization, etc.
- Workforce prediction
- Workload prediction

Algorithms: Bayesian Linear, Linear Regression, Ordinal Regression, ANN, Boosted Decision Tree, Decision Forest



Regression versus Classification

Does your customer want to predict | estimate a number (regression) or apply a label | categorize (classification)?

- Regression problems

- Estimate household power consumption
- Estimate customer's income



- Classification problems

- Power station will | will not meet demand
- Customer will respond to advertising



Binary versus Multiclass Classification

Does your customer want a yes|no answer?

- Binary examples

- click prediction
- yes|no
- over|under
- win|loss



- Multiclass examples

- kind of tree
- kind of network attack
- type of heart disease



DEMO

Machine Learning Basics Infographic

References

- Free e-book “Azure Machine Learning”
- <https://mva.microsoft.com/ebooks#9780735698178>
- Azure Machine Learning documentation
- <https://azure.microsoft.com/en-us/documentation/services/machine-learning/>
- Data Science and Machine Learning Essentials
- www.edx.org
- Azure ML HOL (GitHub):
- <https://github.com/Azure-Readiness/hol-azure-machine-learning/>

HOL Document

- Access Azure HOL Doc: <https://aka.ms/azuremlhol>

Thank you

- Check out my blog for Azure ML articles: <http://mostafa.rocks>
- Follow me on Twitter: @MostafaElzoghbi