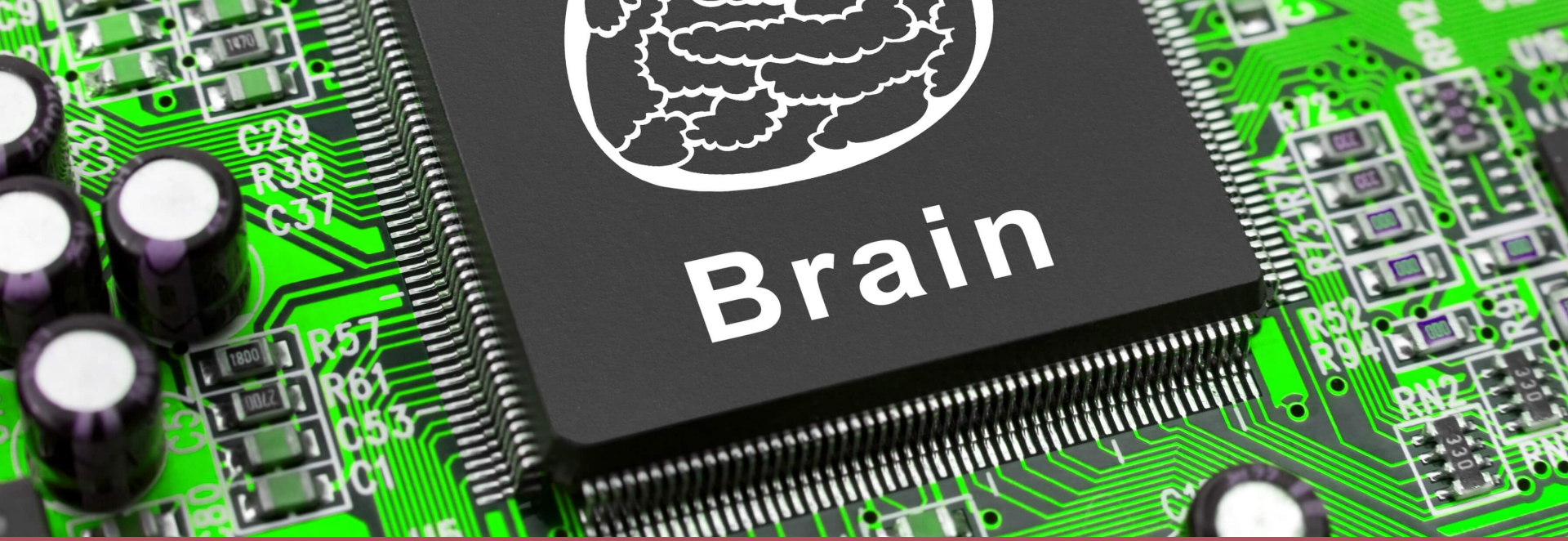


Machine Learning

Your first steps

June 14th, 2017

Quintor

David Stibbe

- Developer
- dstibbe@quintor.nl
- Twitter: @dstibbe

Program

- Intelligence, Artificial Intelligence and ML
- Methods
- Machine learning applied in everyday life
- How to get started

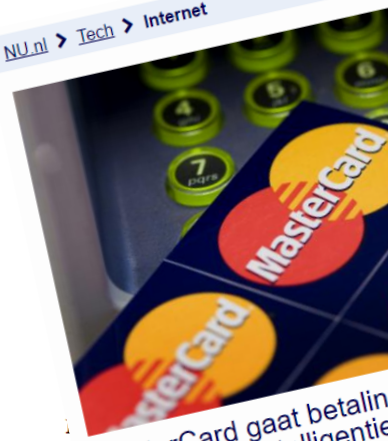
Intelligence, Artificial Intelligence and Machine Learning

DEC 29, 2016 @ 01:01 AM 26,457 VIEWS

Free Webcas

4 Mind-Bl… Artificial Intelligence

NU.nl › Tech › Internet



MasterCard gaat betalin… kunstmatige intelligentie…

Gepubliceerd: 30 november 2016 11:33
Laatste update: 30 november 2016 18:44

Creditcardbedrijf MasterCard ga… kunstmatige intelligentie verifië…

Met de kunstmatige intelligentie … weggefilterd, terwijl legitiem beta… systeem zal wereldwijd door Ma…

De kunstmatige intelligentie ki… besteedt. Op deze manier wil… wanneer er verdachte betaling… worden geblokkeerd als er abnormaal winkel…

Op het moment kijkt een computer al of een betaling wel klopt, maar …

EYEWITNESS NEWS

Independent.
Insightful.
International

LATEST SPORT LIFESTYLE POLITICS OPINION

HOW DO YOU TEACH A DRIVERLESS CAR TO DRIVE?

How does a driverless car learn how to drive, especially … something goes wrong?

FILE: Pilot models of the Uber self-driving car is disp… Technologies Center in Pittsburgh, Pennsylvania. Pi…

DIGITAL TRENDS

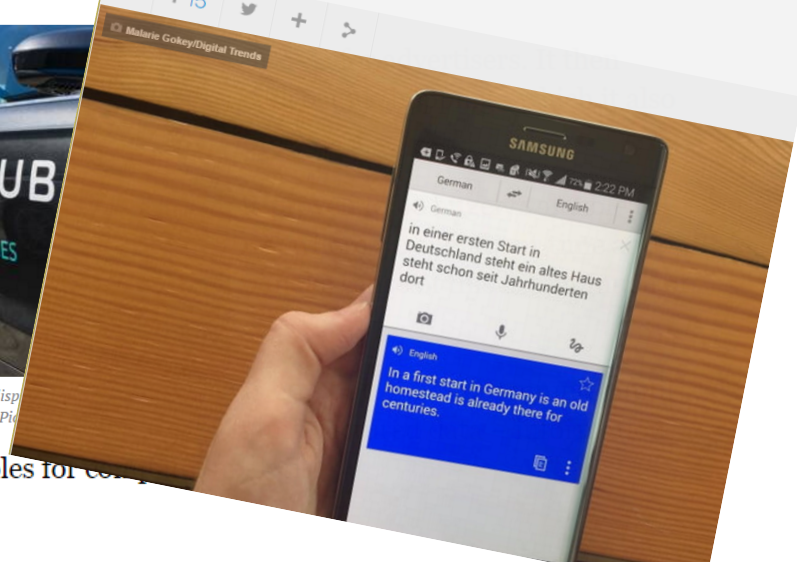
Now Reading: Machine learning improvements for Google Translate expand to more languages

Menu

MACHINE LEARNING IMPROVEMENTS FOR GOOGLE TRANSLATE EXPAND TO MORE LANGUAGES

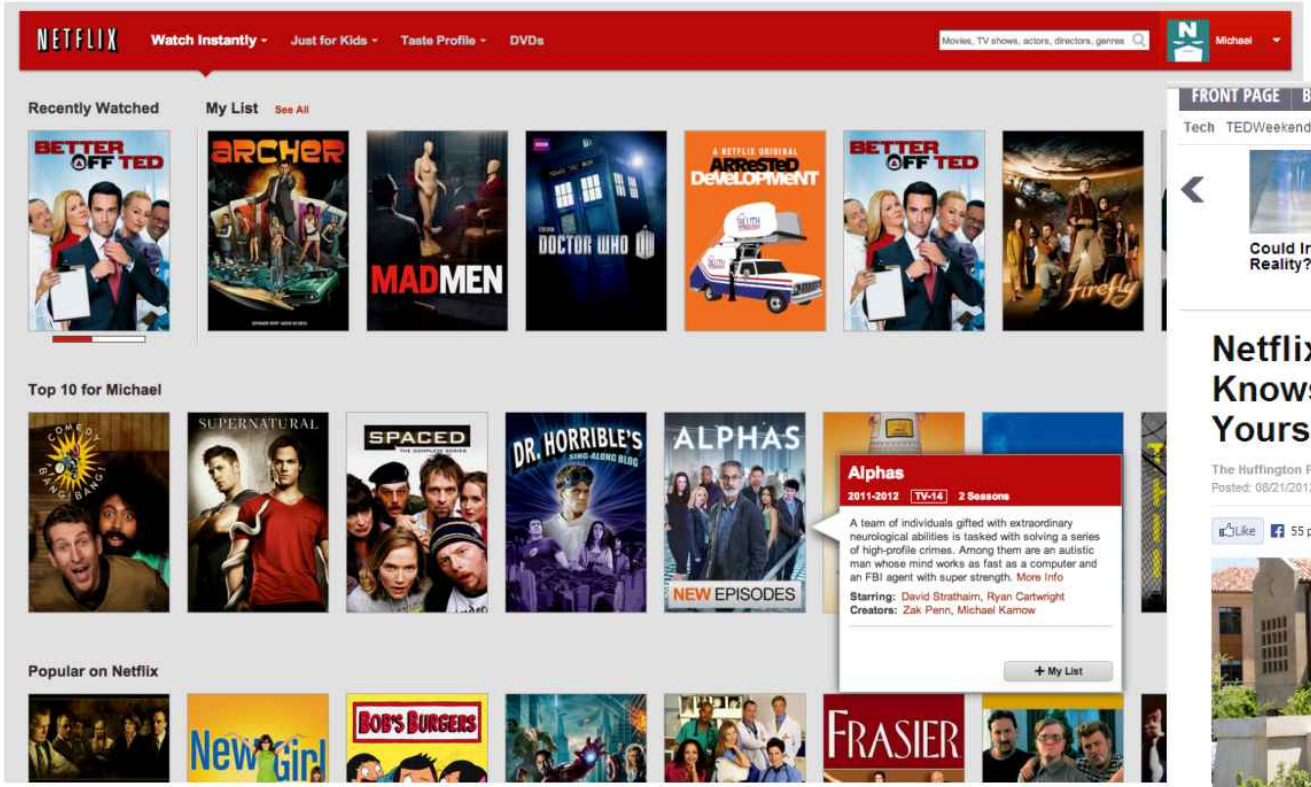
By Christian de Looper — Updated March 7, 2017 8:35 am

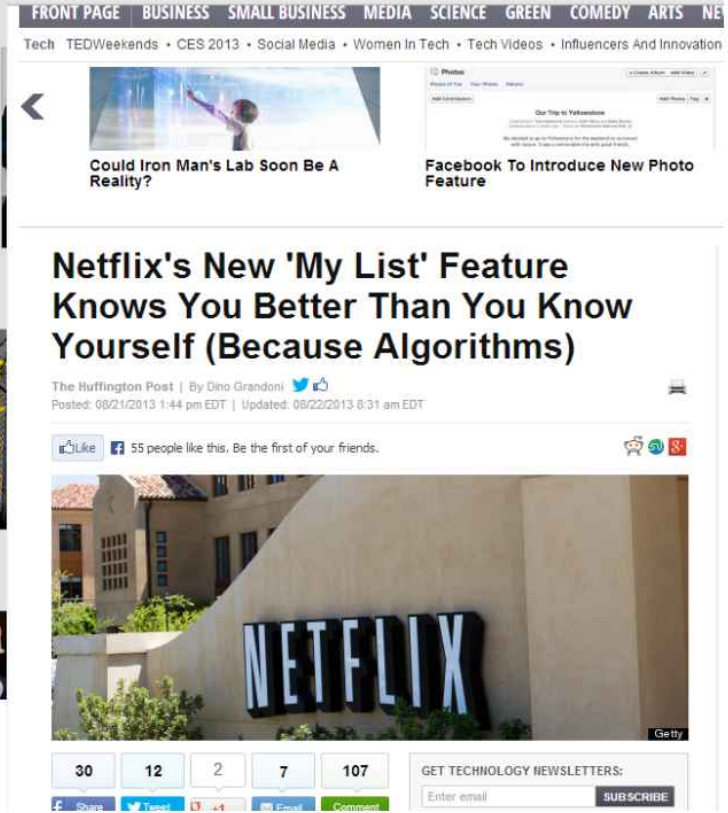
f 15



…rows and tables for co…

EVERYTHING is a Recommendation

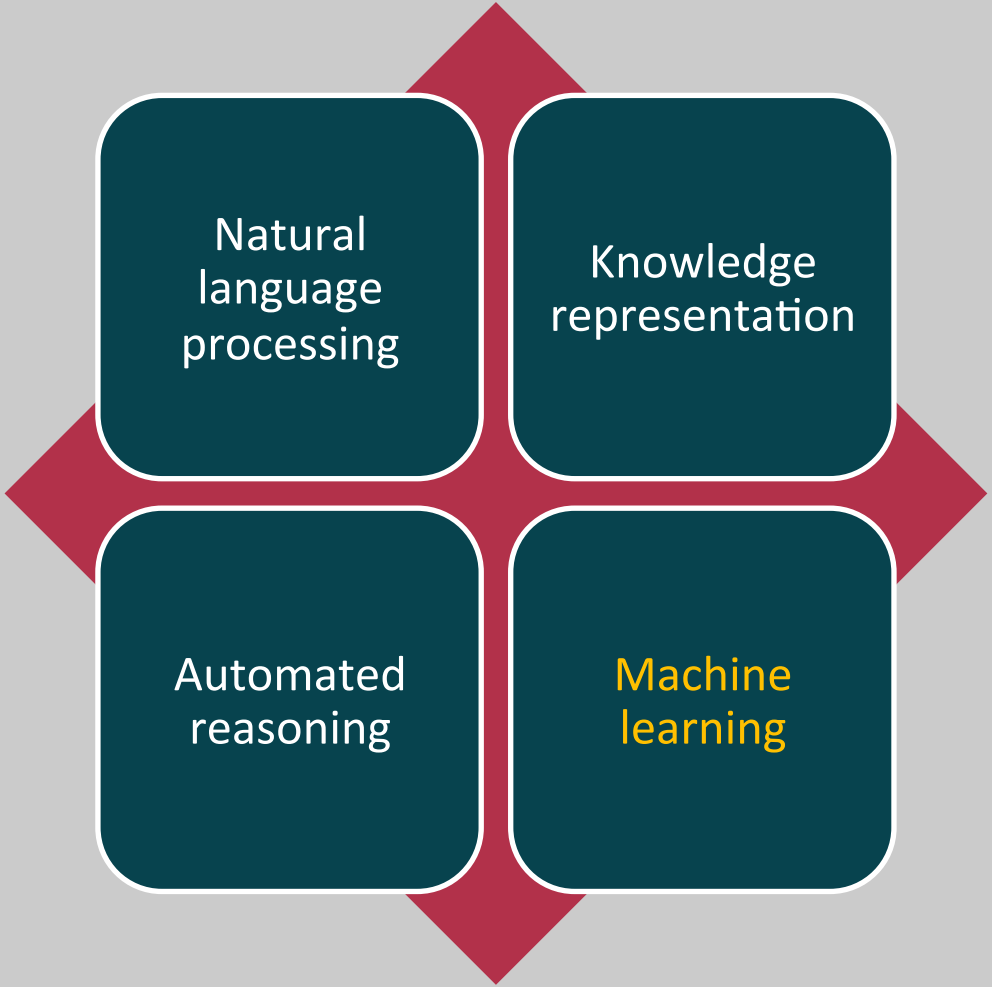




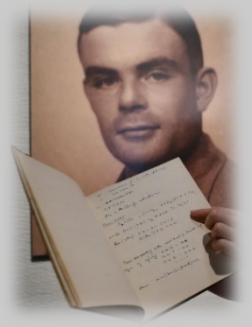
NETFLIX™

ASSISTANT

Artificial intelligence



ARTIFICIAL INTELLIGENCE



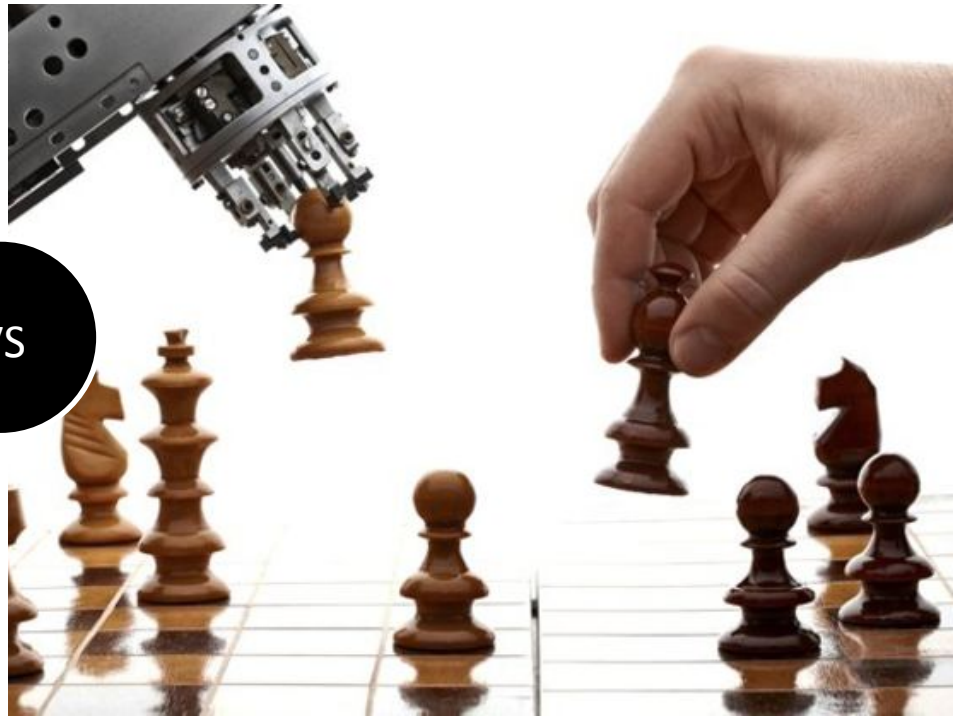
Machine Learning

The field of study that gives computers the ability to learn without being explicitly programmed

- Arthur Samuel -

1950 1960 1970 1980 1990 2000 2010

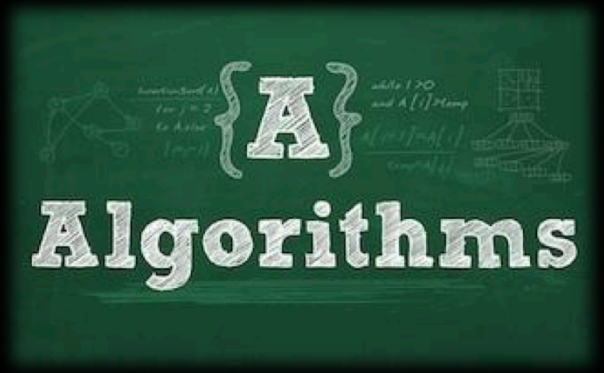
VS

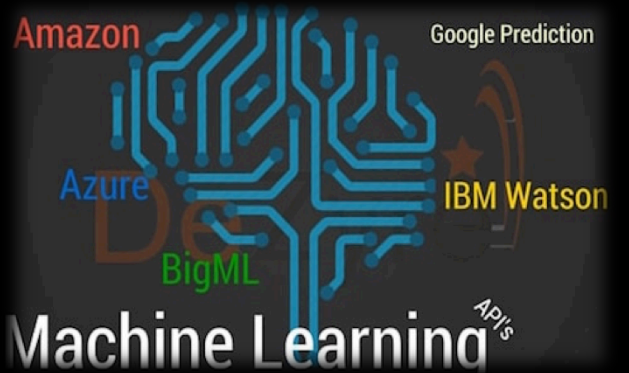


Strong AI
Artificial Generic Intelligence

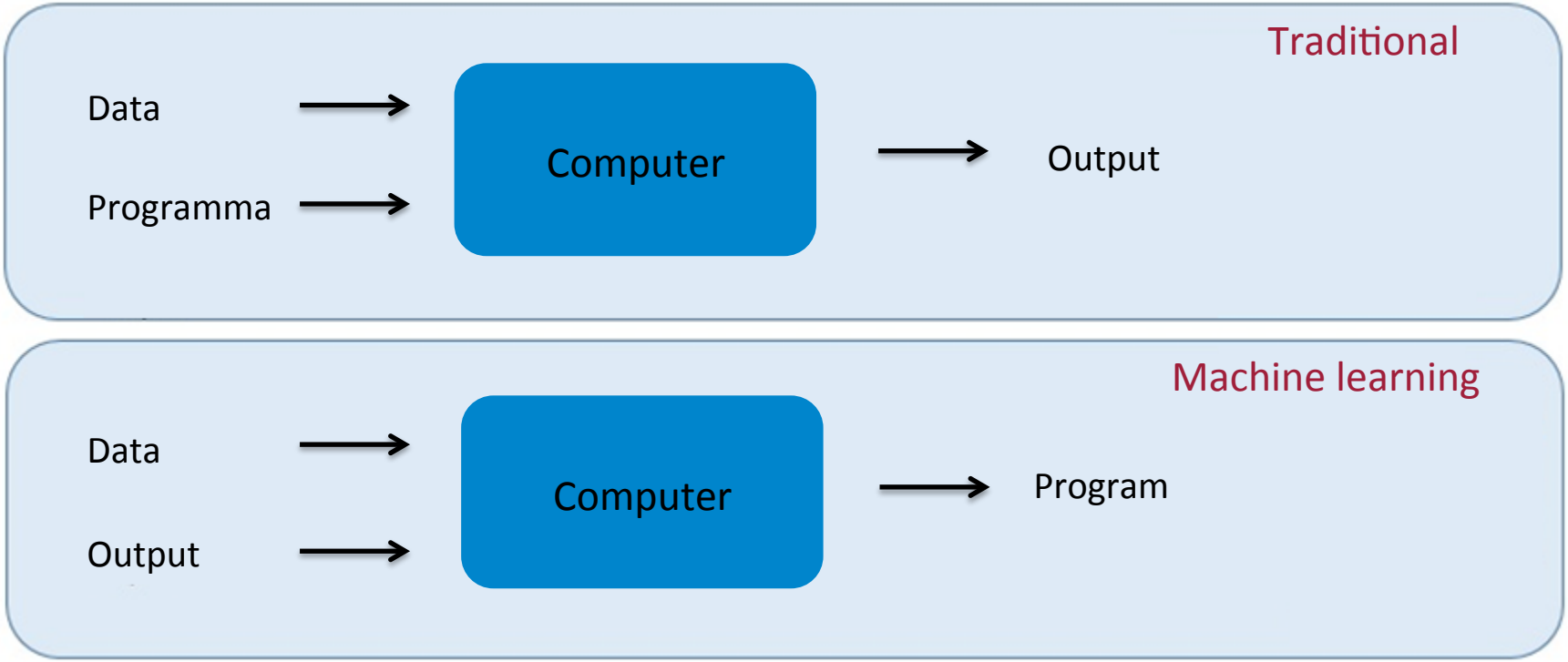
Weak AI

BIG
DATA


{A}
Algorithms

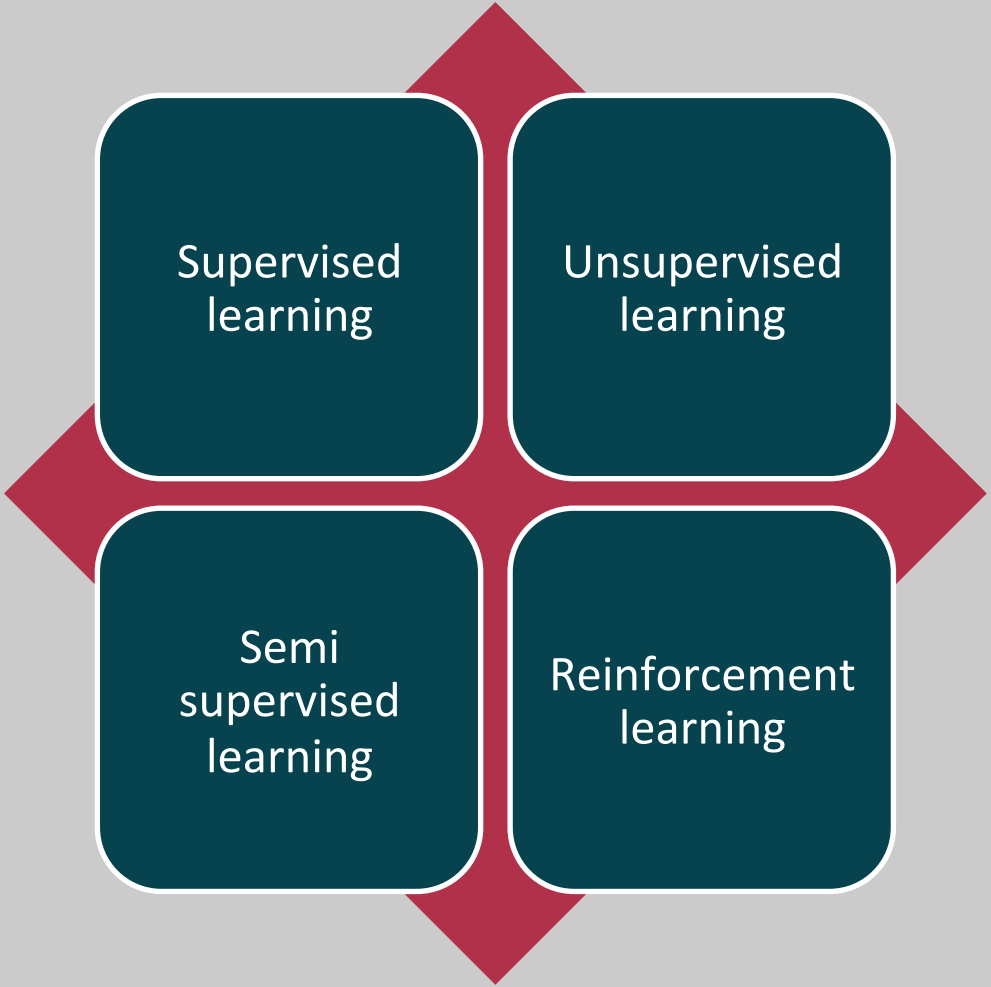

Amazon
Google Prediction
Azure
IBM Watson
BigML
Machine Learning
API's

Machine learning

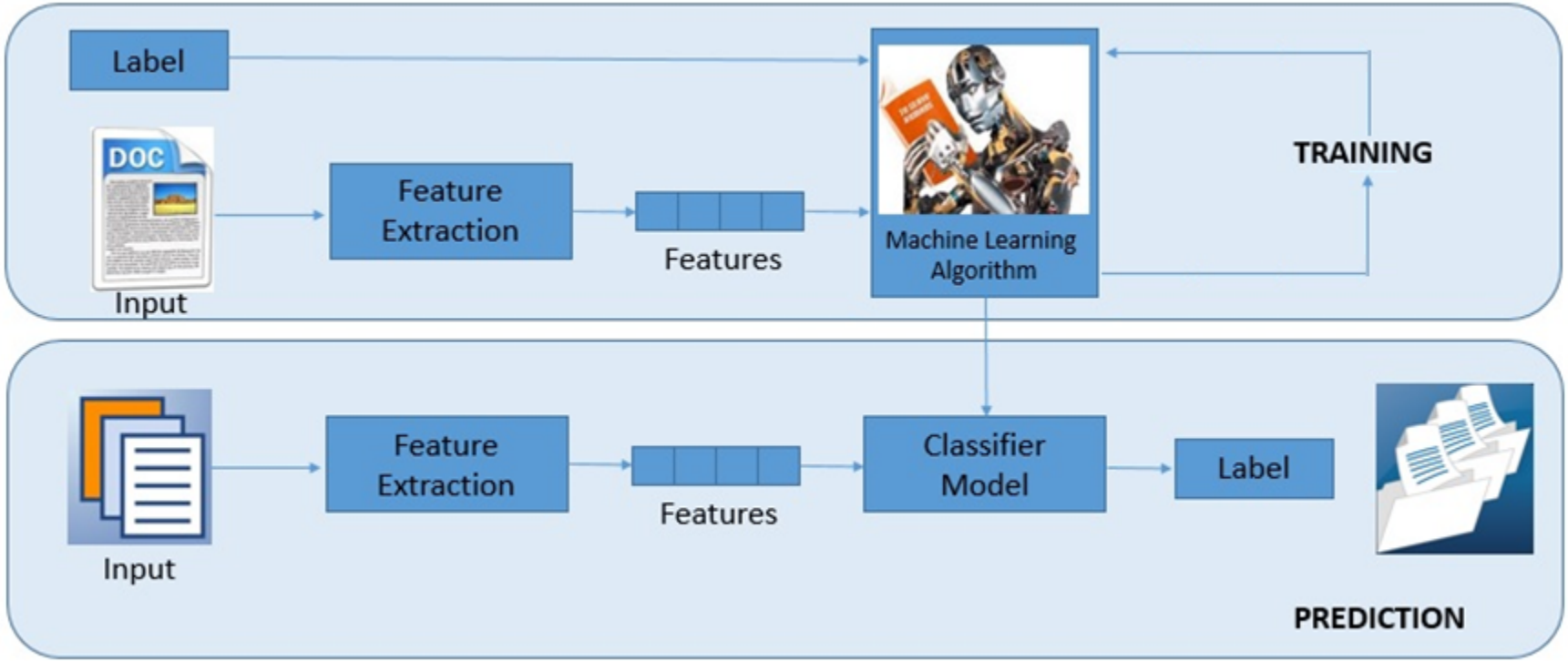


def detect_colors(image):
# lots of code
def detect_edges(image):
# lots of code
def analyze_shapes(image):
# lots of code
def guess_texture(image):
# lots of code
def define_fruit():
# lots of code
def handle_probability():
# lots of code

Methods

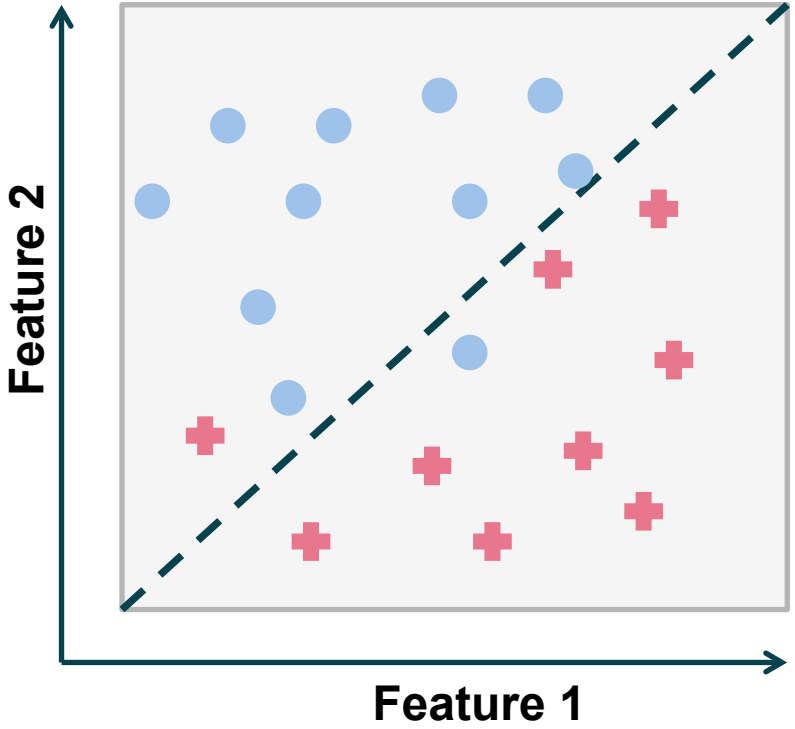

Supervised learning
Unsupervised learning
Semi supervised learning
Reinforcement learning

Supervised learning

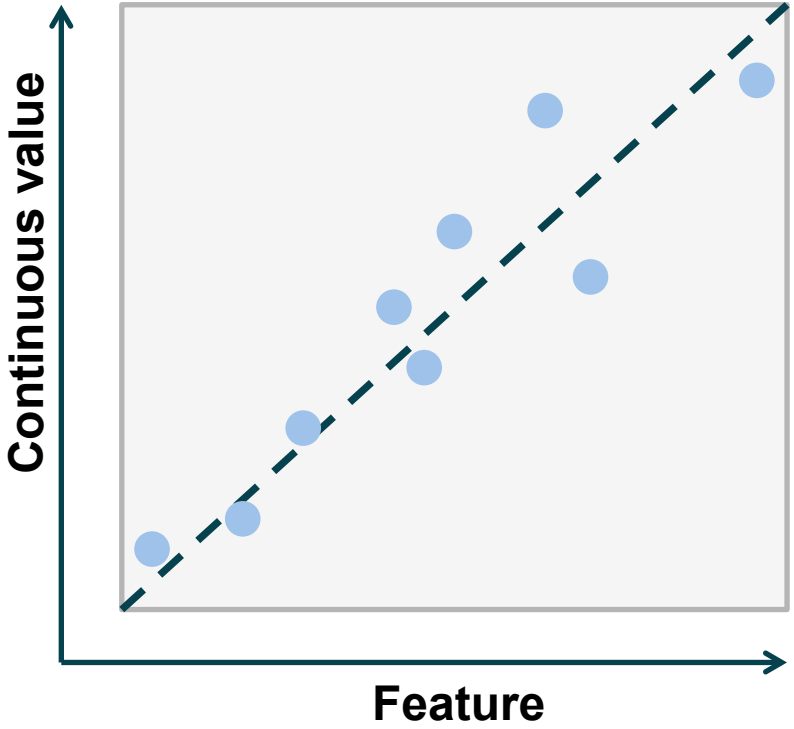


Algorithms Supervised learning

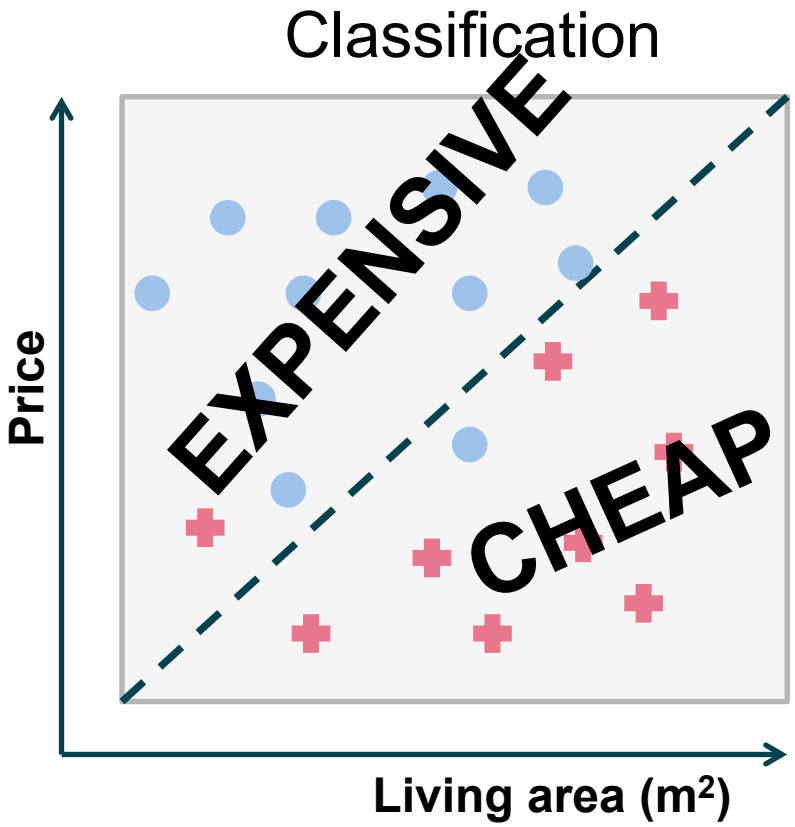
Classification

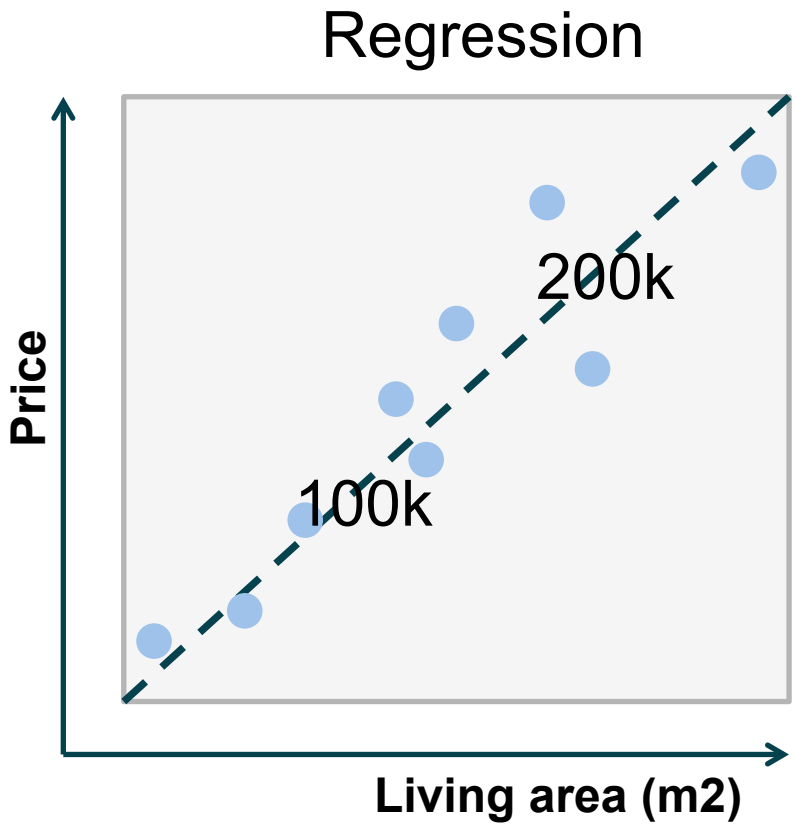


Regression

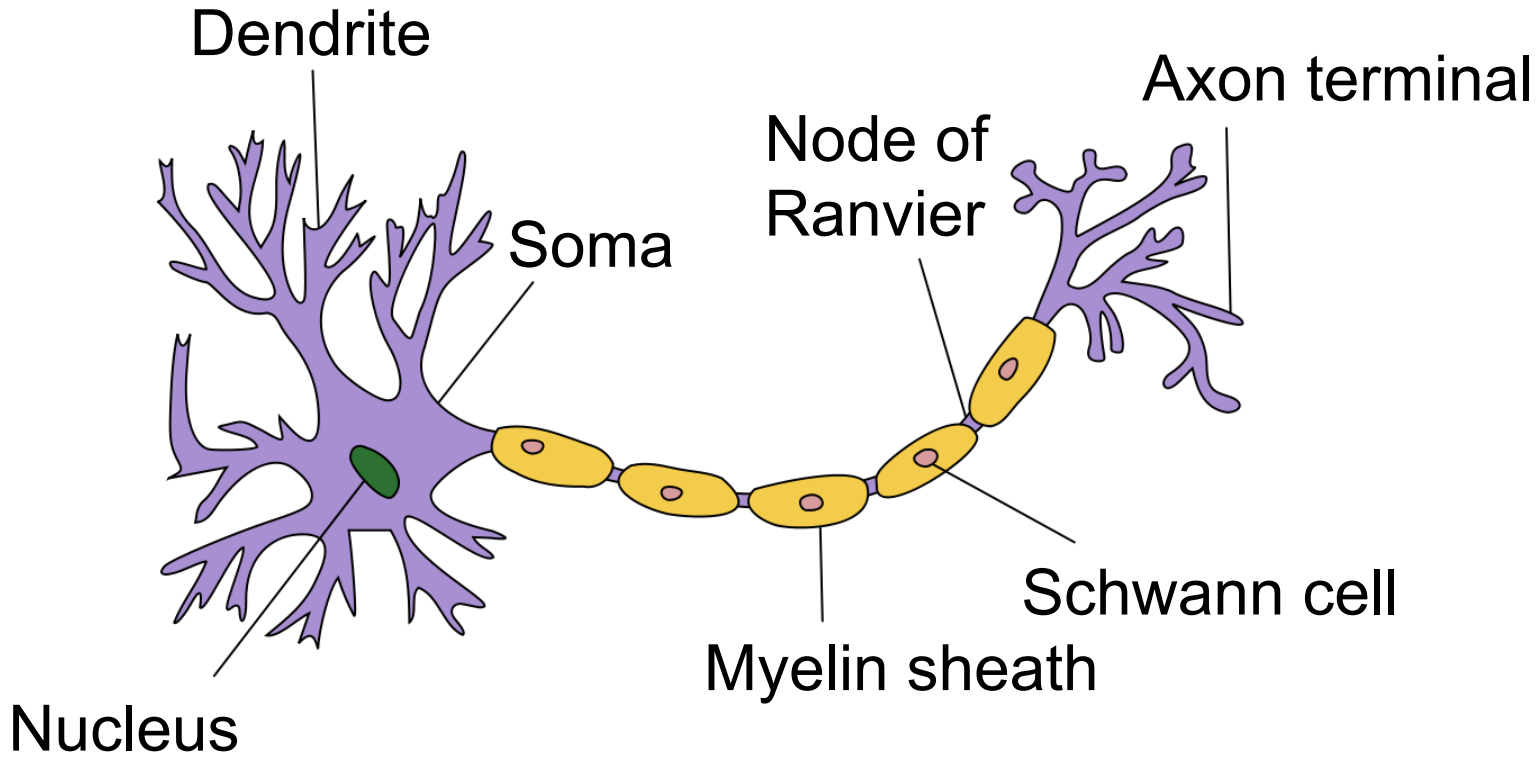


Algorithms Supervised learning

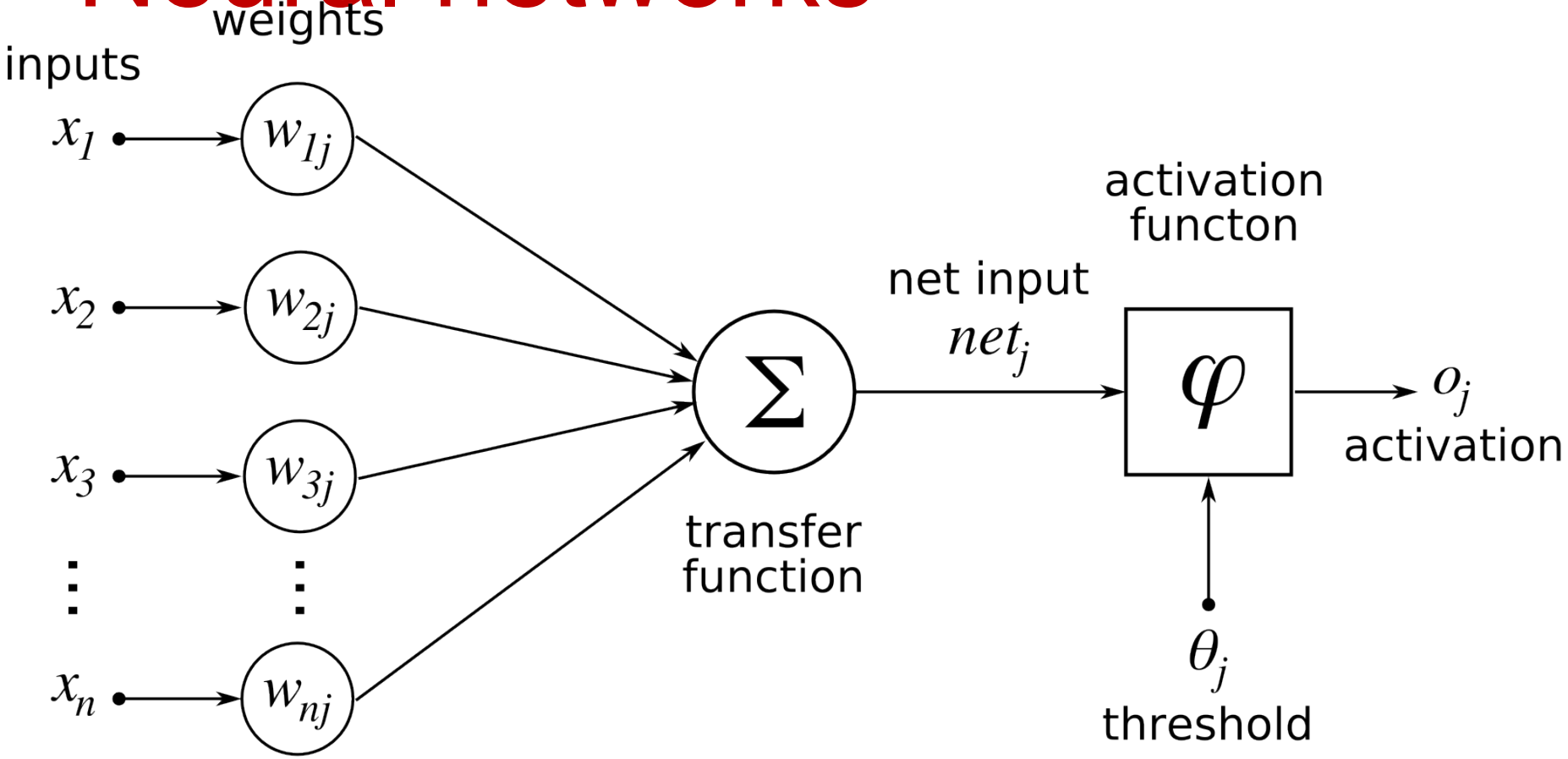




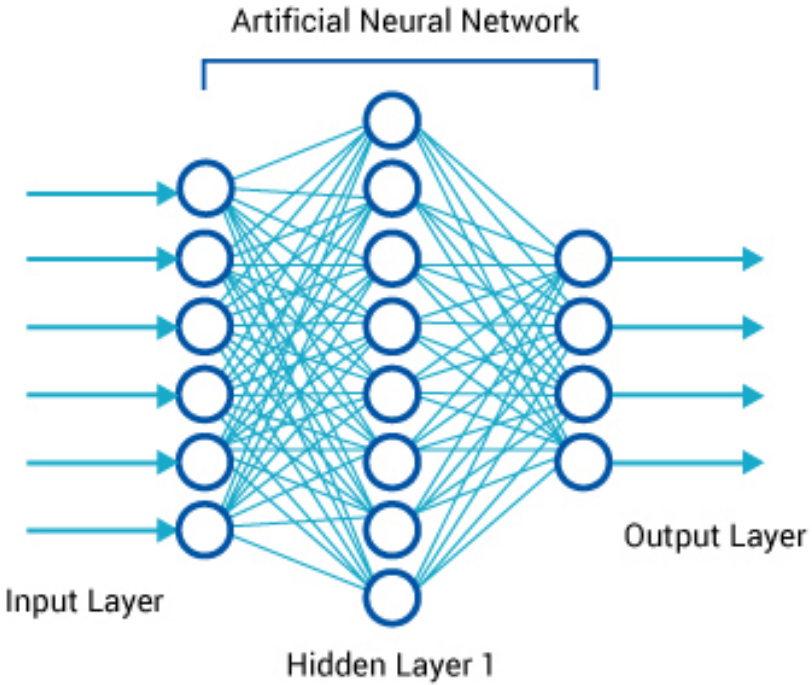
Neural networks



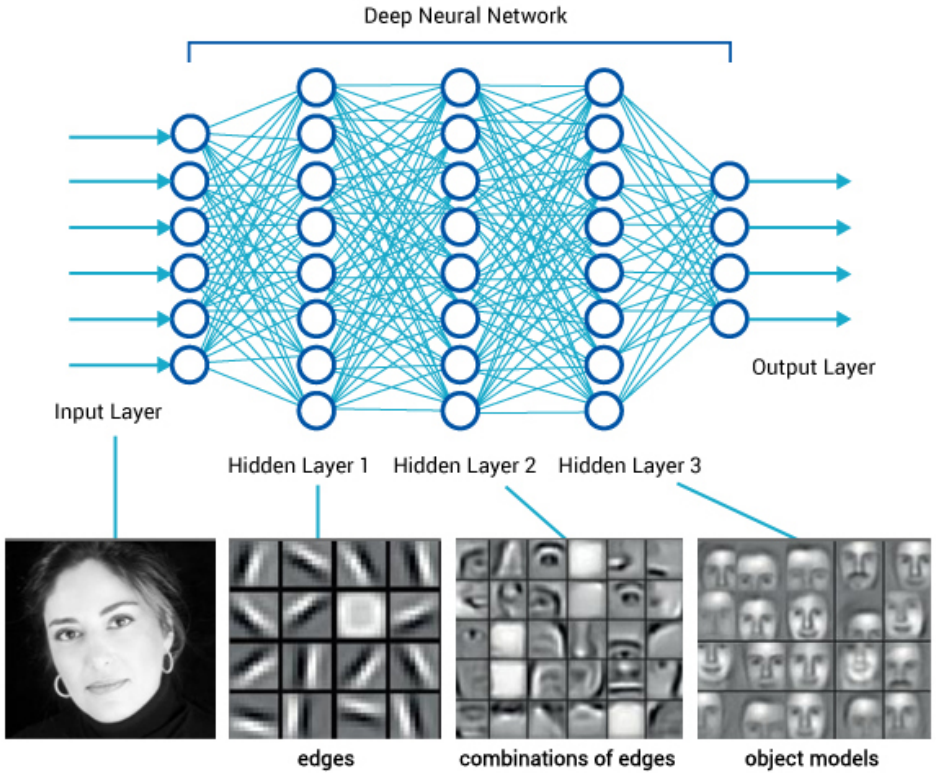
Neural networks



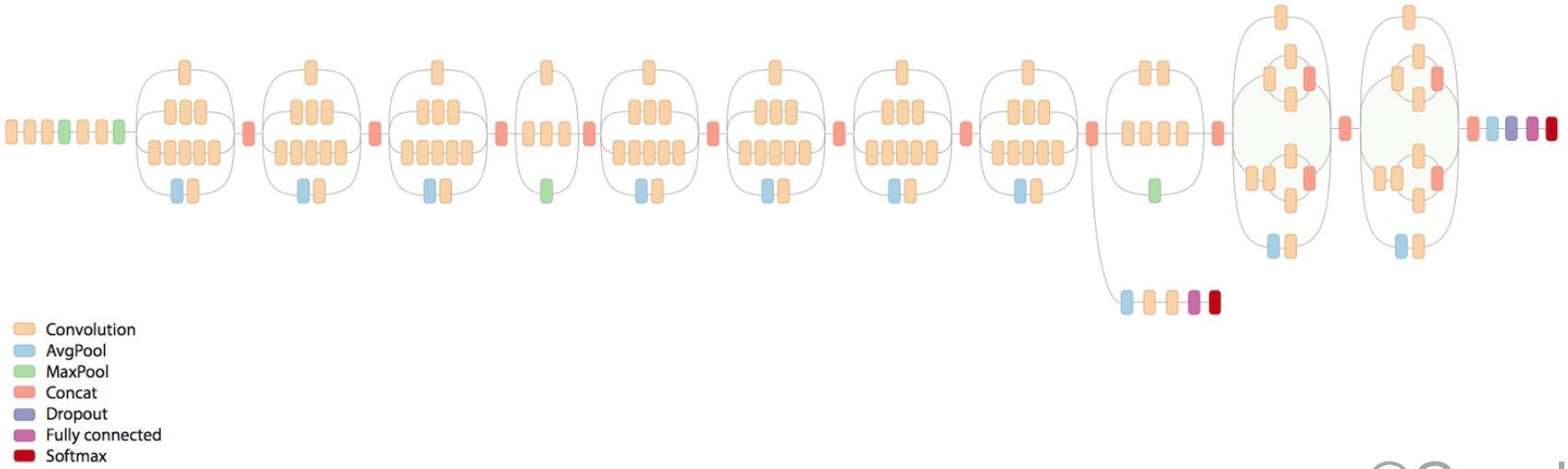
Neurale networks



Deep learning



Inception



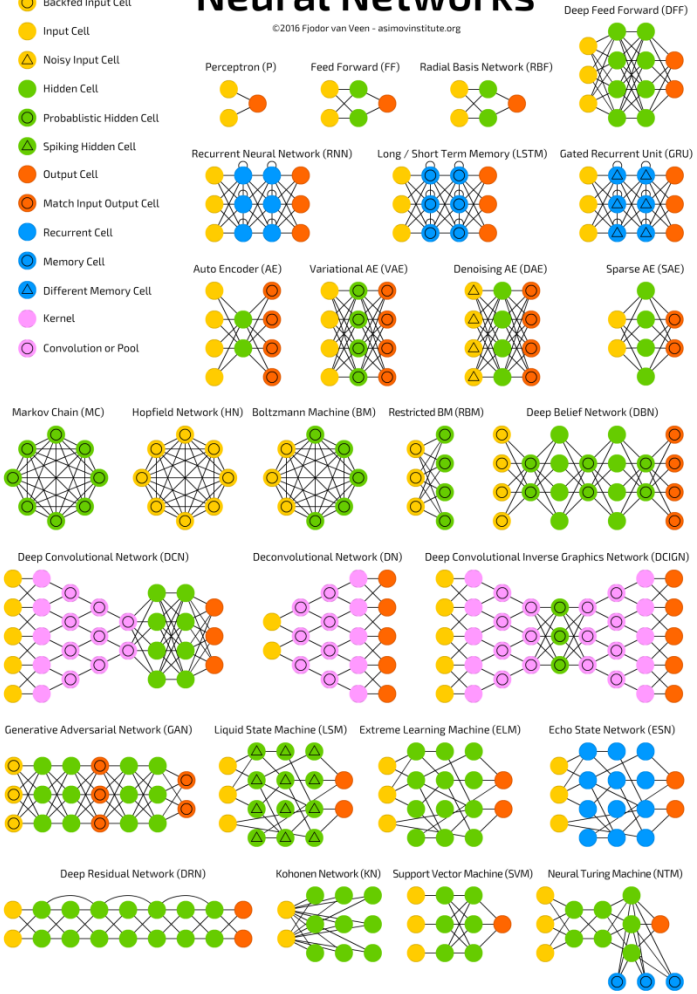
©Google

A mostly complete chart of

Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

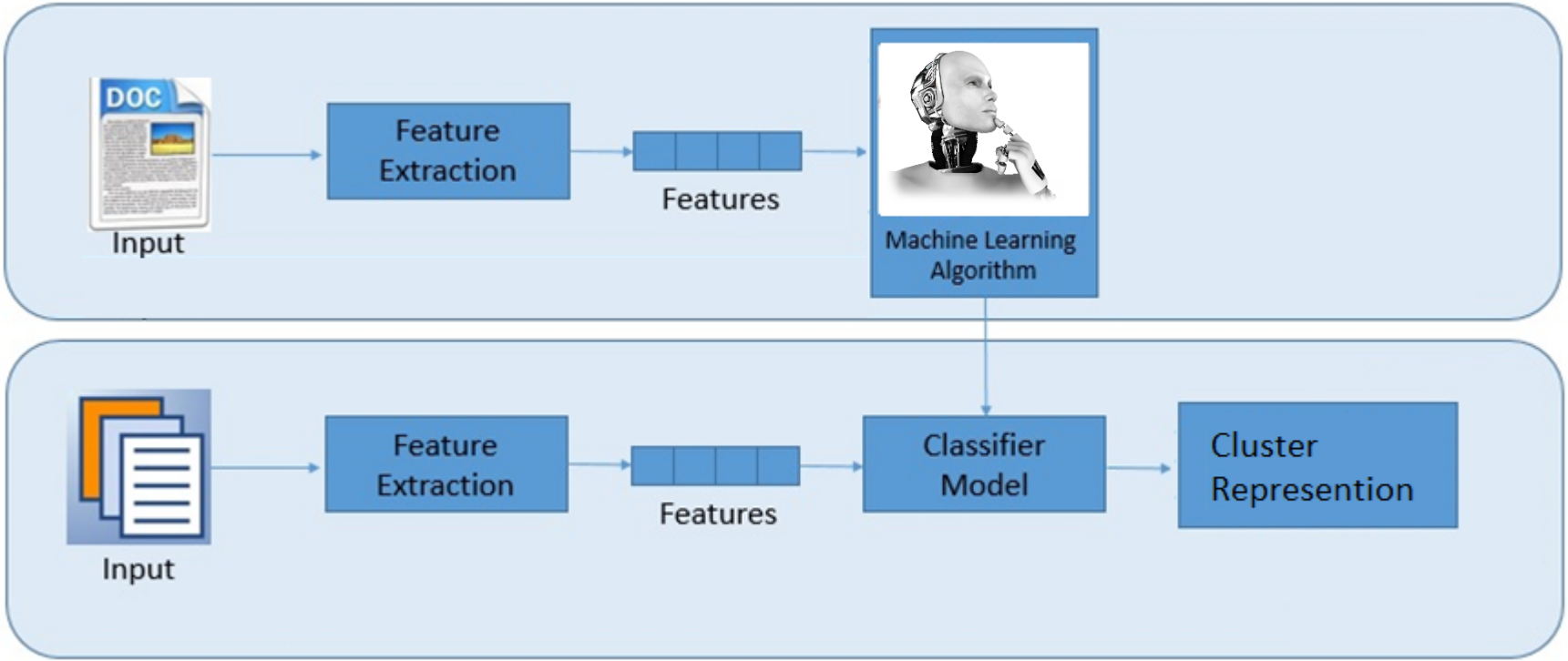
- Backfed Input Cell
- Input Cell
- Noisy Input Cell
- Hidden Cell
- Probablistic Hidden Cell
- Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- Different Memory Cell
- Kernel
- Convolution or Pool



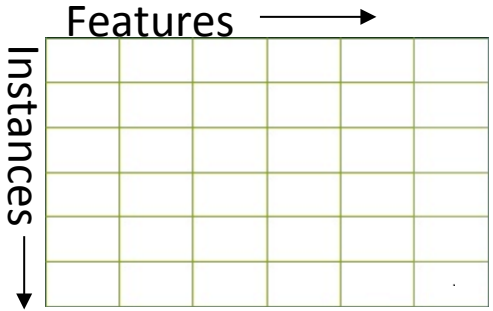
Frameworks providing trained

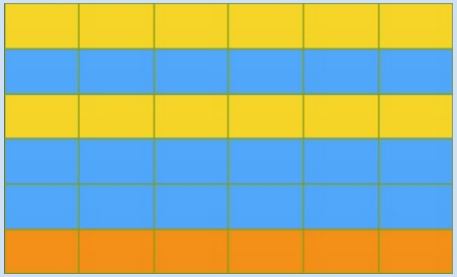
- Google Cloud Machine Learning Engine
- IBM Watson Machine Learning
- Microsoft Azure Machine Learning
- Apple Core ML
- Caffe2 Model Zoo
- and more …

Unsupervised learning

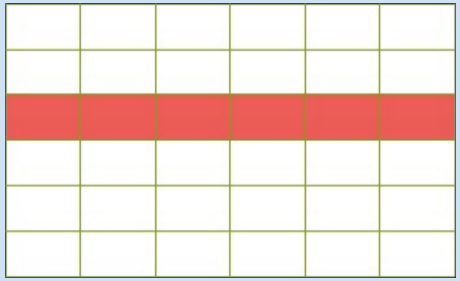


Algoritmen unsupervised learning

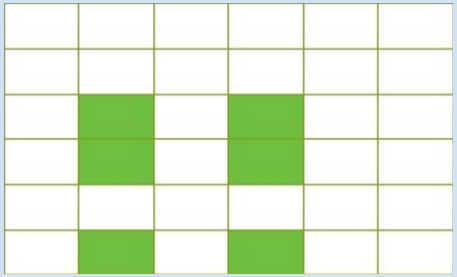




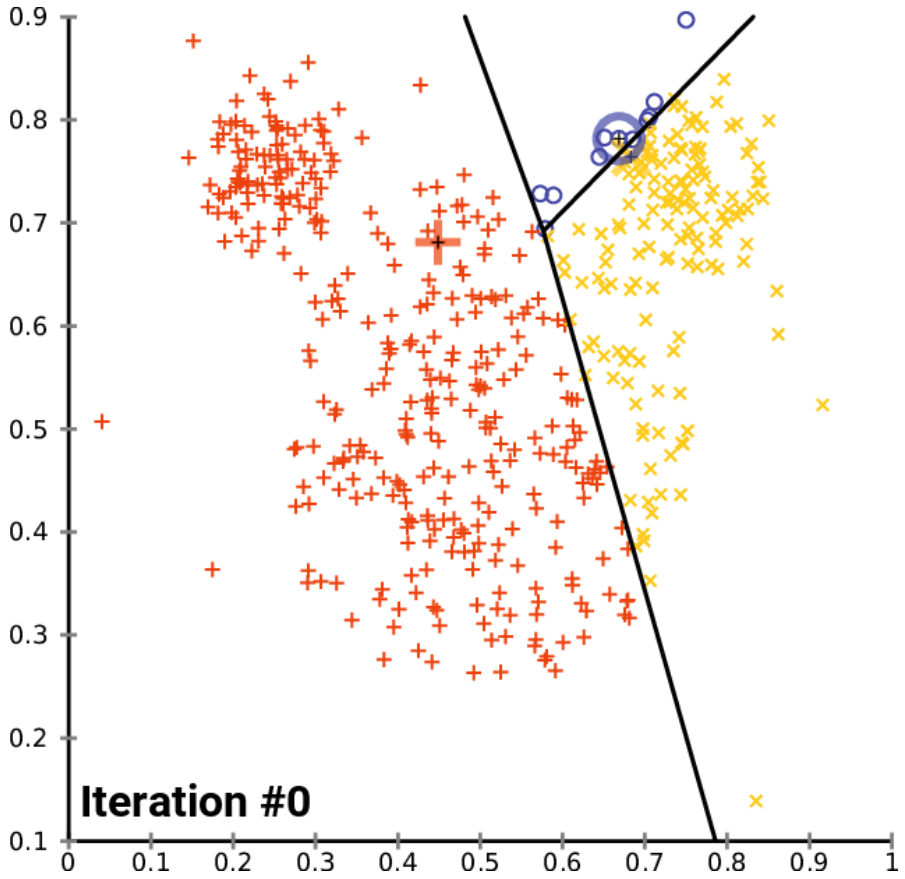
Clustering
Find simular instances



Anomaly detection
Find unusual instances



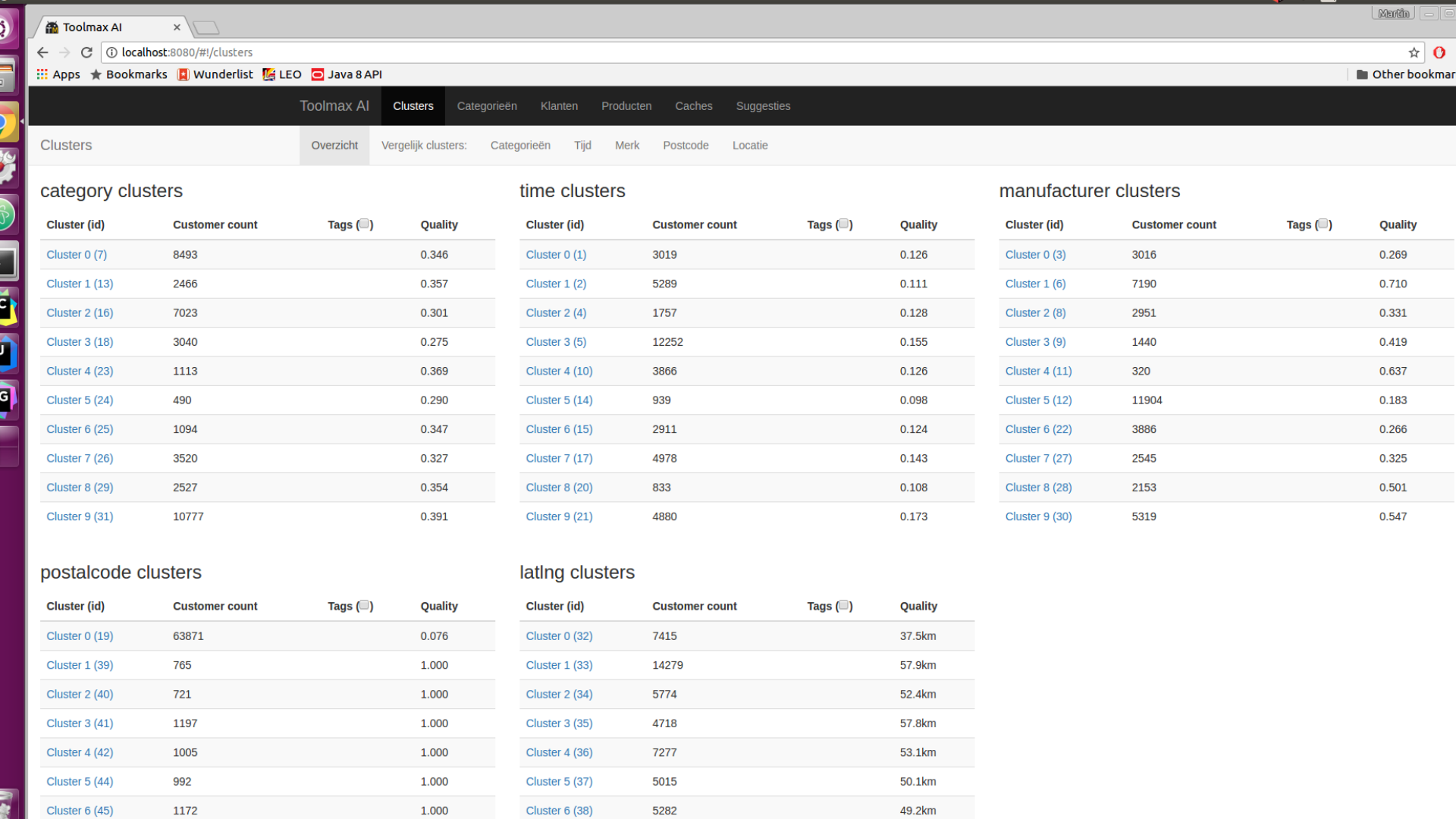
Association discovery
Find feature rules



By chire @ wikimedia

ToolMax
Professioneel Gereedschap & Accessoires - Ruim Assortiment

Offer products by customer profile



category clusters

Cluster (id)	Customer count	Tags ( )	Quality
Cluster 0 (7)	8493		0.346
Cluster 1 (13)	2466		0.357
Cluster 2 (16)	7023		0.301
Cluster 3 (18)	3040		0.275
Cluster 4 (23)	1113		0.369
Cluster 5 (24)	490		0.290
Cluster 6 (25)	1094		0.347
Cluster 7 (26)	3520		0.327
Cluster 8 (29)	2527		0.354
Cluster 9 (31)	10777		0.391

time clusters

Cluster (id)	Customer count	Tags ( )	Quality
Cluster 0 (1)	3019		0.126
Cluster 1 (2)	5289		0.111
Cluster 2 (4)	1757		0.128
Cluster 3 (5)	12252		0.155
Cluster 4 (10)	3866		0.126
Cluster 5 (14)	939		0.098
Cluster 6 (15)	2911		0.124
Cluster 7 (17)	4978		0.143
Cluster 8 (20)	833		0.108
Cluster 9 (21)	4880		0.173

manufacturer clusters

Cluster (id)	Customer count	Tags ( )	Quality
Cluster 0 (3)	3016		0.269
Cluster 1 (6)	7190		0.710
Cluster 2 (8)	2951		0.331
Cluster 3 (9)	1440		0.419
Cluster 4 (11)	320		0.637
Cluster 5 (12)	11904		0.183
Cluster 6 (22)	3886		0.266
Cluster 7 (27)	2545		0.325
Cluster 8 (28)	2153		0.501
Cluster 9 (30)	5319		0.547

postalcode clusters

Cluster (id)	Customer count	Tags ( )	Quality
Cluster 0 (19)	63871		0.076
Cluster 1 (39)	765		1.000
Cluster 2 (40)	721		1.000
Cluster 3 (41)	1197		1.000
Cluster 4 (42)	1005		1.000
Cluster 5 (44)	992		1.000
Cluster 6 (45)	1172		1.000

latlng clusters

Cluster (id)	Customer count	Tags ( )	Quality
Cluster 0 (32)	7415		37.5km
Cluster 1 (33)	14279		57.9km
Cluster 2 (34)	5774		52.4km
Cluster 3 (35)	4718		57.8km
Cluster 4 (36)	7277		53.1km
Cluster 5 (37)	5015		50.1km
Cluster 6 (38)	5282		49.2km

Chrome
Toolmax AI
localhost:8080/#!/customer/65906/products
Apps Bookmarks Wunderlist LEO Java 8 API
Other bookmarks

Aantal suggesties 10 Suggest

Product suggesties

ID	Naam
11807	0.19116371189686707
2143	0.16708896858254116
9520	0.15727100961257262
15551	0.1252443198225216
6123	0.11358069719174357
2142	0.09420548449891594
2052	0.09253856001782608
5119	0.08204529375151812
16430	0.07396950883297149
3493	0.07136500531840409

Voorgestelde producten

SDS-Plus borenset | Robustline S4L | 5-Delig | 2607019928

Certainty: 0,1912

GBH 2-28 DFV Boorhamer | 3.2J 850w | in L-boxx

Certainty: 0,1671

Accessoire L-boxx I 102mm | 6 dlg SET | 2608438035

Certainty: 0,1573

GBH 2-28 DFV Boorhamer | 3.2J 850w | in koffer

Certainty: 0,0942

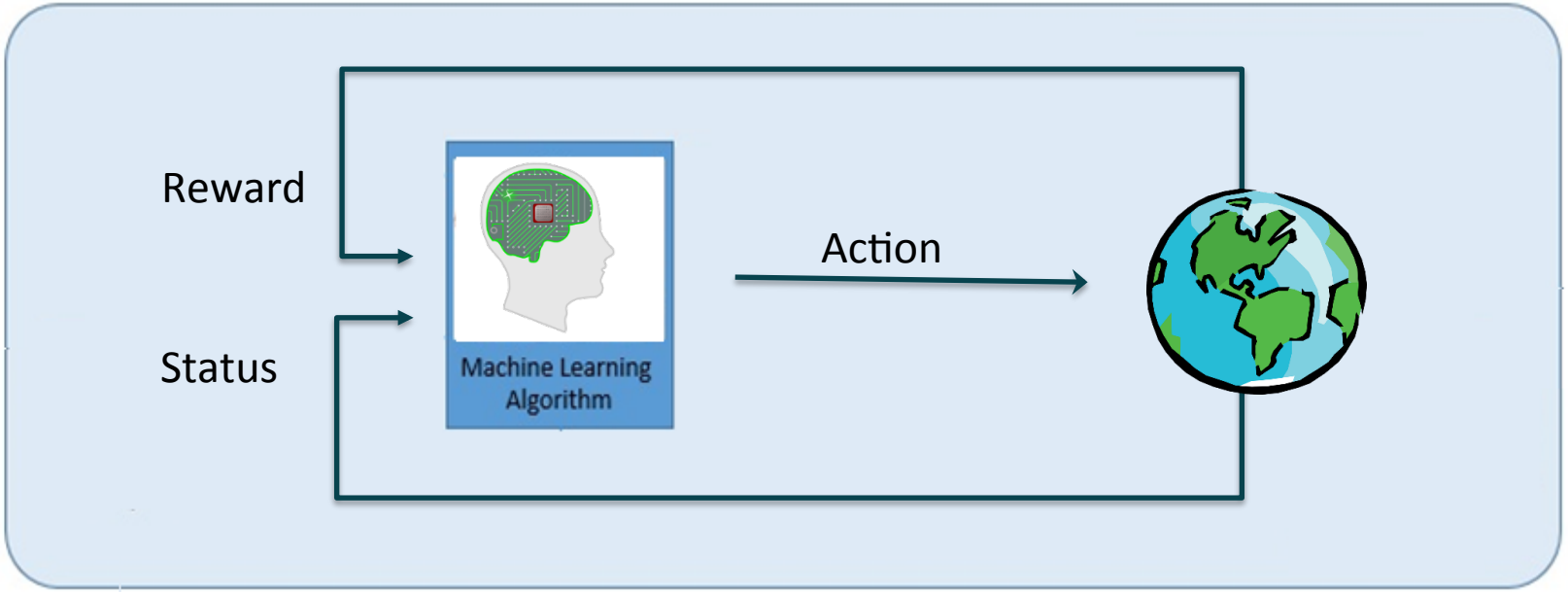
GBH 2-28 DFV Boorhamer | 3.2J 850w | + GWS 7-125 Slijper

Certainty: 0,1252

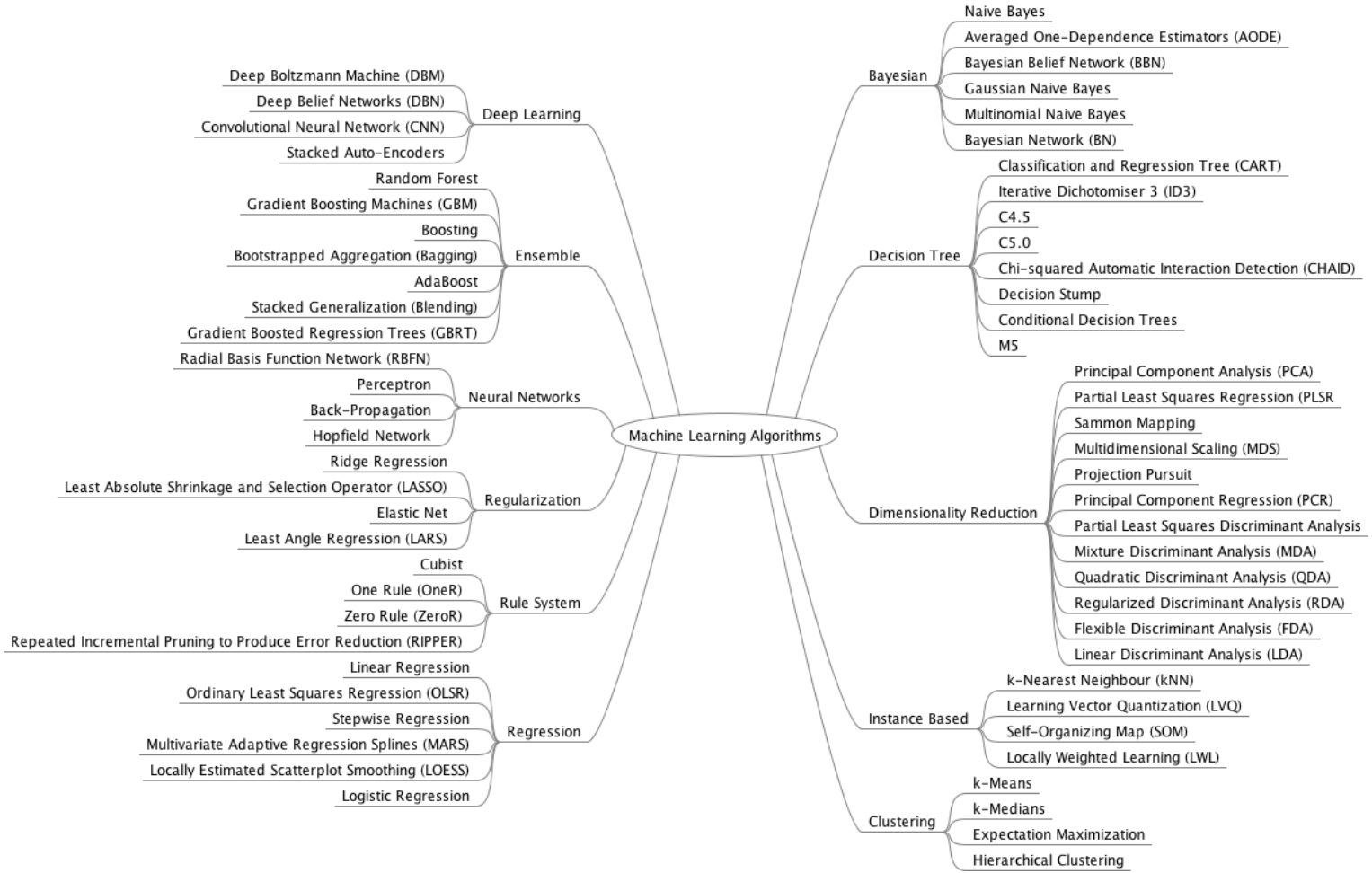
GMS 120 Multidetector

Certainty: 0,1136

Reinforcement learning



Demo reinforcement


Machine Learning Algorithms
Bayesian
Naive Bayes
Averaged One-Dependence Estimators (AODE)
Bayesian Belief Network (BBN)
Gaussian Naive Bayes
Multinomial Naive Bayes
Bayesian Network (BN)
Decision Tree
Classification and Regression Tree (CART)
Iterative Dichotomiser 3 (ID3)
C4.5
C5.0
Chi-squared Automatic Interaction Detection (CHAID)
Decision Stump
Conditional Decision Trees
M5
Dimensionality Reduction
Principal Component Analysis (PCA)
Partial Least Squares Regression (PLSR)
Sammon Mapping
Multidimensional Scaling (MDS)
Projection Pursuit
Principal Component Regression (PCR)
Partial Least Squares Discriminant Analysis
Mixture Discriminant Analysis (MDA)
Quadratic Discriminant Analysis (QDA)
Regularized Discriminant Analysis (RDA)
Flexible Discriminant Analysis (FDA)
Linear Discriminant Analysis (LDA)
Instance Based
k-Nearest Neighbour (kNN)
Learning Vector Quantization (LVQ)
Self-Organizing Map (SOM)
Locally Weighted Learning (LWL)
Clustering
k-Means
k-Medians
Expectation Maximization
Hierarchical Clustering
Deep Learning
Deep Boltzmann Machine (DBM)
Deep Belief Networks (DBN)
Convolutional Neural Network (CNN)
Stacked Auto-Encoders
Ensemble
Random Forest
Gradient Boosting Machines (GBM)
Boosting
Bootstrapped Aggregation (Bagging)
AdaBoost
Stacked Generalization (Blending)
Gradient Boosted Regression Trees (GBRT)
Neural Networks
Radial Basis Function Network (RBFN)
Perceptron
Back-Propagation
Hopfield Network
Regularization
Ridge Regression
Least Absolute Shrinkage and Selection Operator (LASSO)
Elastic Net
Least Angle Regression (LARS)
Rule System
Cubist
One Rule (OneR)
Zero Rule (ZeroR)
Repeated Incremental Pruning to Produce Error Reduction (RIPPER)
Regression
Linear Regression
Ordinary Least Squares Regression (OLSR)
Stepwise Regression
Multivariate Adaptive Regression Splines (MARS)
Locally Estimated Scatterplot Smoothing (LOESS)
Logistic Regression

ML applied in everyday life

14-Day Weather Forecast for

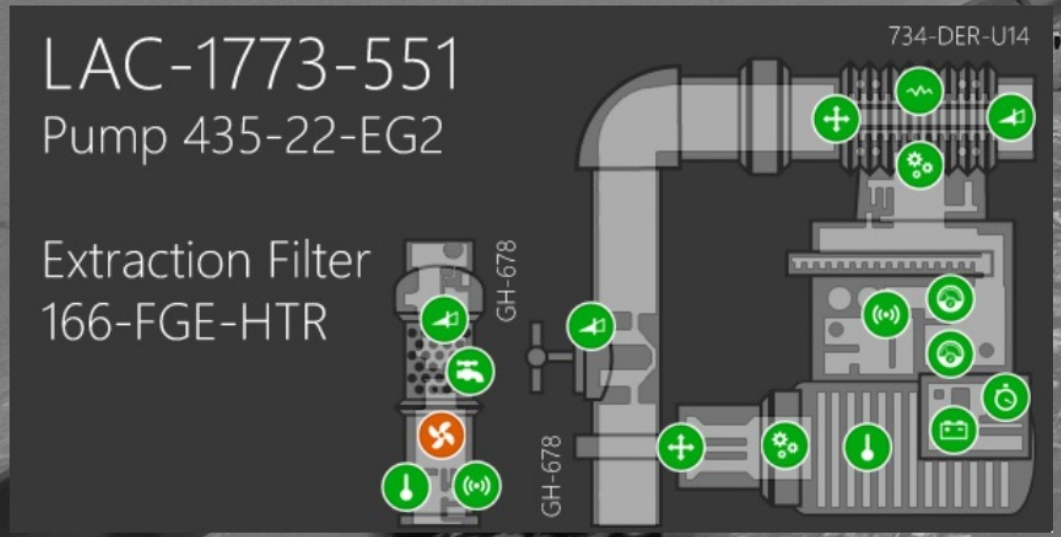
Amsterdam

Netherlands

Extraction

Nort

Asset Sensor Details



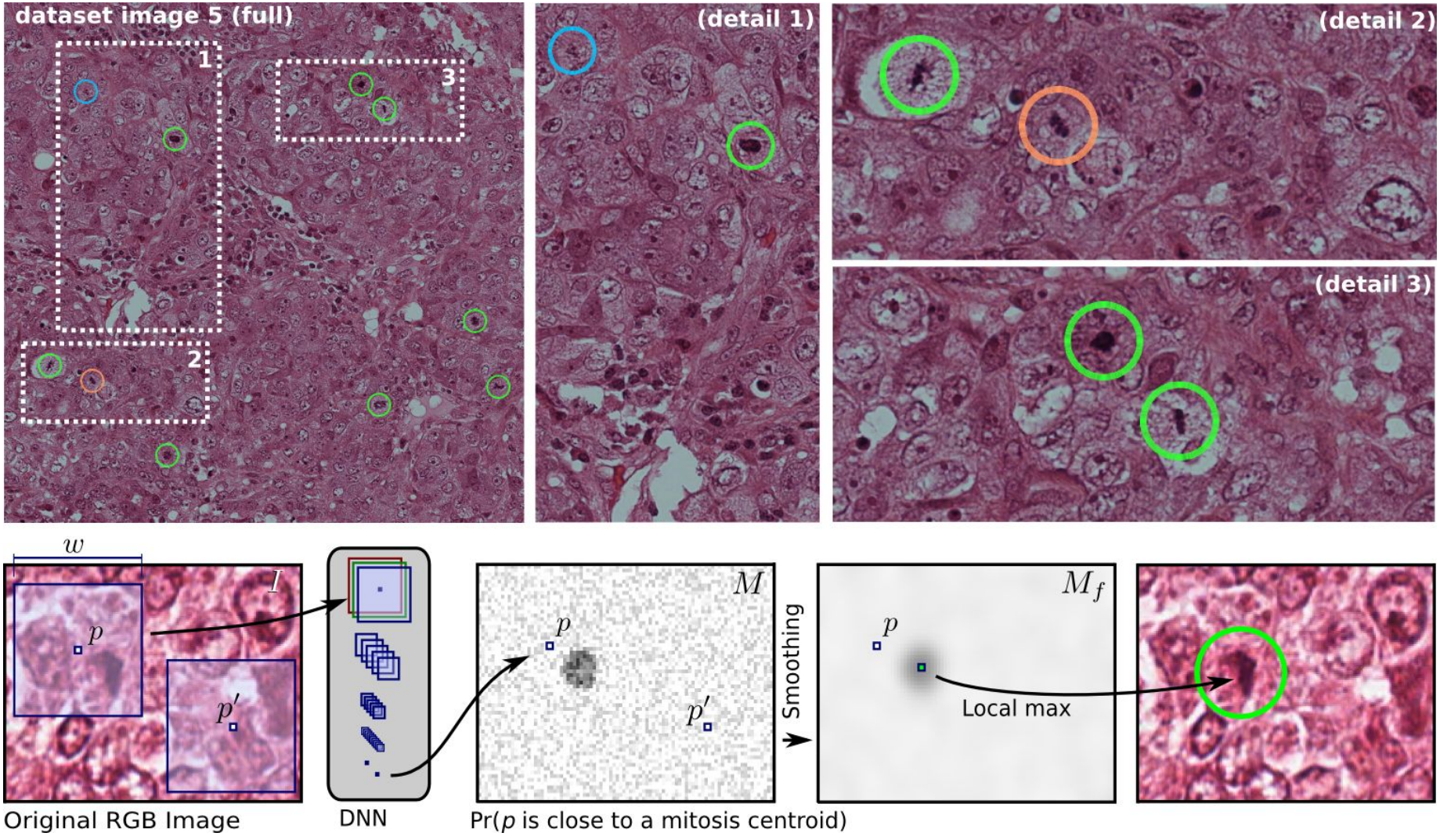
Product Quality Detail

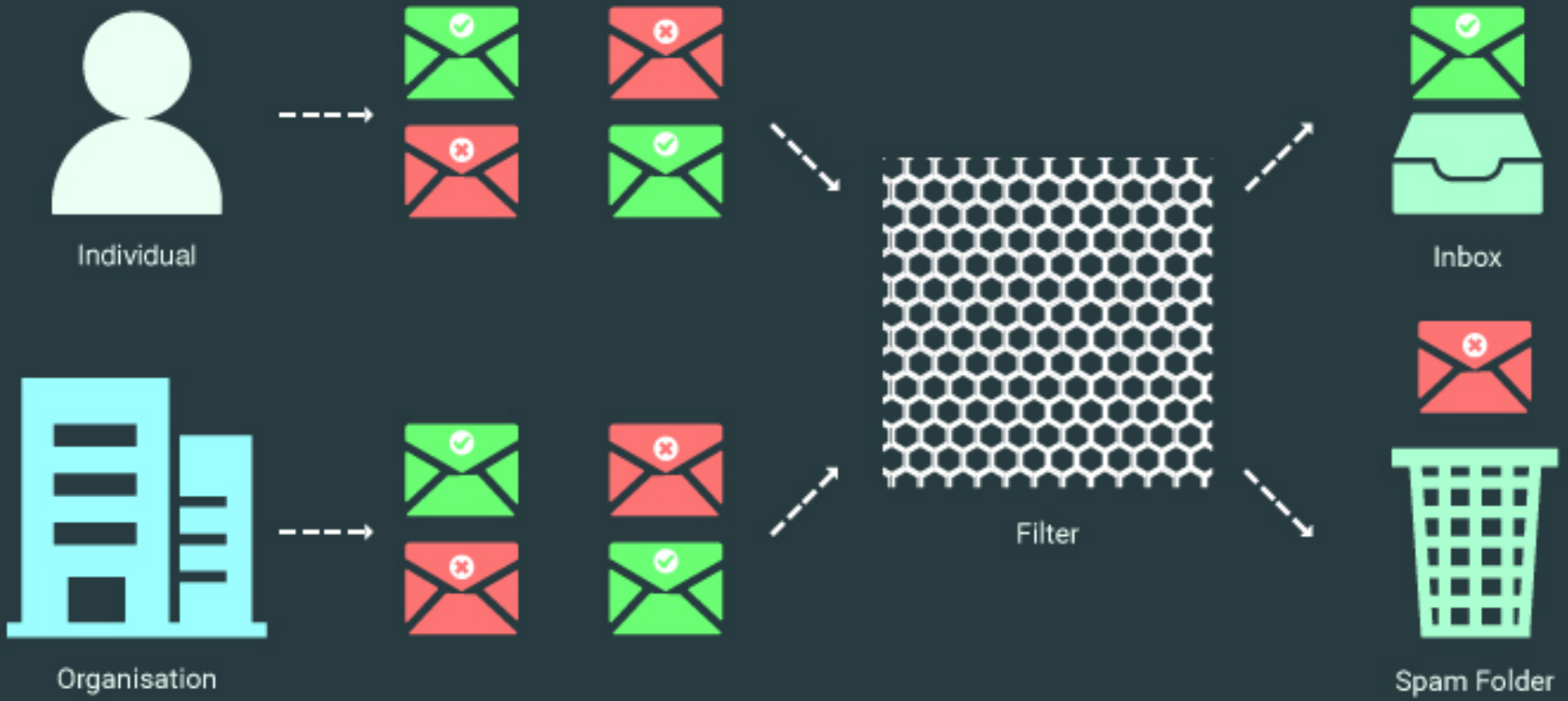
747.5 Barrels/hour	26.8 BS&W (%) Purity
39.2 API Gravity (%)	237k Barrel count (month)
26.1 H2S (%)	19.6M Barrel count (total)

Asset Health Status

89% Overall Status	85% Tank level (%)	100 AMPS/Rated (%)	176.1 Mean temperature

ВЫХОД В ГОРОД
Russian English
ACCESS TO CITY
IMPORT
PAUSE
SCAN


dataset image 5 (full)
1
2
3
(detail 1)
(detail 2)
(detail 3)
w
I
p
p′
DNN
M
p
p′
Smoothing
M_f
p
Local max
Original RGB Image
Pr(p is close to a mitosis centroid)


Individual
Organisation
Filter
Inbox
Spam Folder

LEFT REARWARD VEHICLE CAMERA
MEDIUM RANGE VEHICLE CAMERA
RIGHT REARWARD VEHICLE CAMERA
MOTION FLOW
LANE LINES
LANE LINES
ROAD FLOW
IN-PATH OBJECTS
ROAD LIGHTS
OBJECTS
ROAD SIGNS

How to get started

Getting started

APIs for pretrained models

- Google Cloud ML Engine
- IBM Bluemix
- Microsoft Azure ML
- Apple Core ML
- and more …

Build your own

1. Learn python
2. Install python notebook
3. Try the online guides

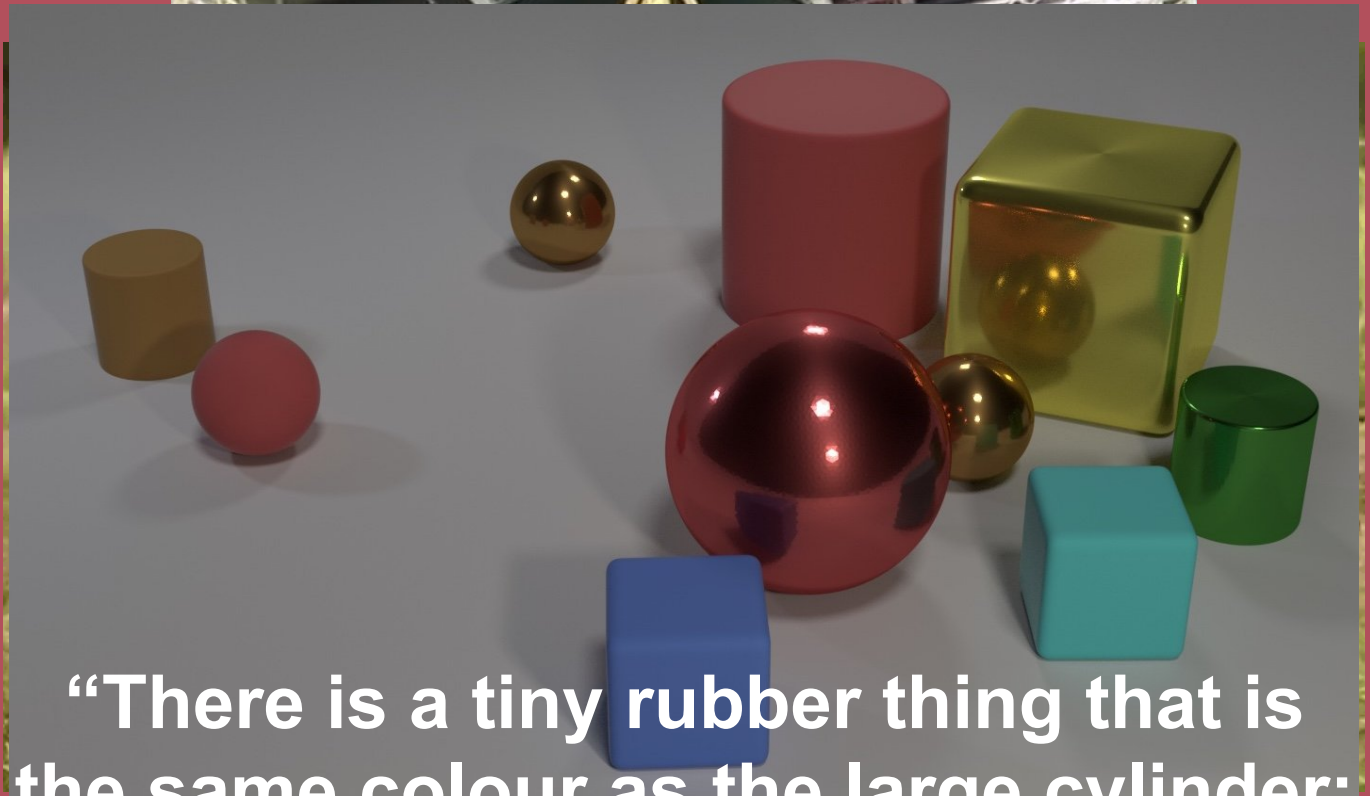
Machine learning libraries

- Tensorflow (TFLearn, Keras)
- Torch (Lua)
- Theano (Python)
- Deeplearning4j (Java)
- Caffe2 (Python)

Data analysis

- Frame the problem (what do you want to solve?)
- Look at the bigger picture
  (what happens before and after this?)
- Check assumptions
- Visualize!

Future



“There is a tiny rubber thing that is the same colour as the large cylinder; what shape is it?

Resources – online framewor

- Google Cloud ML: https://cloud.google.com/ml-engine/
- IBM Bluemix: https://console.ng.bluemix.net/
- Microsoft Azure ML: https://azure.microsoft.com/en-us/services/machine-learning/
- Apple Core ML: https://developer.apple.com/documentation/coreml

Resources – build your own

Libraries:
Tensorflow: https://www.tensorflow.org
Pyhon: https://docs.python.org/3/
NumPy: http://www.numpy.org/
Pandas: http://pandas.pydata.org/
SciKit-learn: http://scikit-learn.org/stable/

Utilities:
Demo site: http://playground.tensorflow.org/
Good hands-on (Aurélien Geron): https://github.com/ageron/handson-ml

Resources

Miscellaneous:

- ImageNet: http://www.image-net.org/
- OpenAI: https://openai.com/
  - Open AI Gym: https://gym.openai.com/
- Genetic Algorithm Walkers: http://rednuht.org/genetic_walkers/

Questions?