Machine Learning and AI concepts and techniques

For CyBOK funded project:
Development of an active learning lesson plan and laboratory materials for AI for Security

Dr Hossein Abroshan

Senior Lecturer in Cyber Security

Anglia Ruskin University, Cambridge, UK

Artificial intelligence (AI)

AI refers to systems or machines that are programmed to think like humans and mimic their actions to perform tasks and can iteratively improve themselves based on the information they collect.[1]

Computer systems can be differentiated on the basis of rationality and thinking vs. acting[2]:

Human approach:

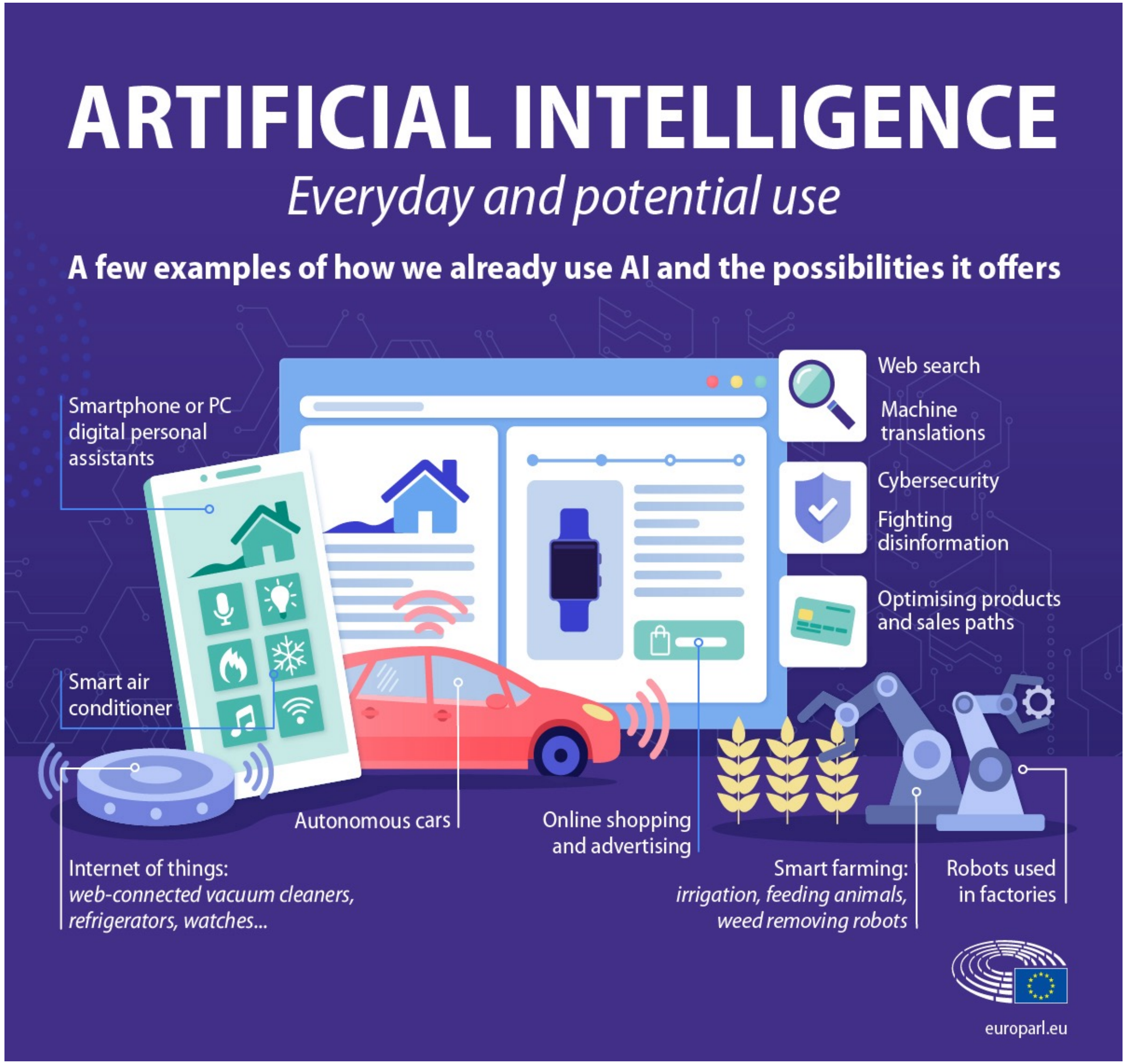
- Systems that think like humans
- Systems that act like humans

Ideal approach:

- Systems that think rationally
- Systems that act rationally

[1] Oracle
[2] Artificial Intelligence: A Modern Approach

Artificial intelligence (AI)


ARTIFICIAL INTELLIGENCE
Everyday and potential use
A few examples of how we already use AI and the possibilities it offers
Smartphone or PC digital personal assistants
Smart air conditioner
Internet of things: web-connected vacuum cleaners, refrigerators, watches...
Autonomous cars
Online shopping and advertising
Web search
Machine translations
Cybersecurity
Fighting disinformation
Optimising products and sales paths
Smart farming: irrigation, feeding animals, weed removing robots
Robots used in factories
europarl.eu

Source: European Parliament

AI Terms

ALGORITHM A set of step-by-step instructions. Computer algorithms can be simple (if it's 3 p.m., send a reminder) or complex (identify pedestrians).

BACKPROPAGATION The way many neural nets learn. They find the difference between their output and the desired output, then adjust the calculations in reverse order of execution.

BLACK BOX A description of some deep learning systems. They take an input and provide an output, but the calculations that occur in between are not easy for humans to interpret.

DEEP LEARNING How a neural network with multiple layers becomes sensitive to progressively more abstract patterns. In parsing a photo, layers might respond first to edges, then paws, then dogs.

EXPERT SYSTEM A form of AI that attempts to replicate a human's expertise in an area, such as medical diagnosis. It combines a knowledge base with a set of hand-coded rules for applying that knowledge. Machine-learning techniques are increasingly replacing hand coding.

GENERATIVE ADVERSARIAL NETWORKS A pair of jointly trained neural networks that generates realistic new data and improves through competition. One net creates new examples (fake Picassos, say) as the other tries to detect the fakes.

MACHINE LEARNING The use of algorithms that find patterns in data without explicit instruction. A system might learn how to associate features of inputs such as images with outputs such as labels.

NEURAL NETWORK A highly abstracted and simplified model of the human brain used in machine learning. A set of units receives pieces of an input (pixels in a photo, say), performs simple computations on them, and passes them on to the next layer of units. The final layer represents the answer.

Source: : https://www.sciencemag.org/news/2017/07/ai-revolution-science

AI Terms

PERCEPTRON An early type of neural network, developed in the 1950s. It received great hype but was then shown to have limitations, suppressing interest in neural nets for years.

NATURAL LANGUAGE PROCESSING A computer's attempt to "understand" spoken or written language. It must parse vocabulary, grammar, and intent, and allow for variation in language use. The process often involves machine learning.

TRANSFER LEARNING A technique in machine learning in which an algorithm learns to perform one task, such as recognizing cars, and builds on that knowledge when learning a different but related task, such as recognizing cats.

TURING TEST A test of AI's ability to pass as human. In Alan Turing's original conception, an AI would be judged by its ability to converse through written text.

NEUROMORPHIC CHIP A computer chip designed to act as a neural network. It can be analogue, digital, or a combination.

TENSORFLOW A collection of software tools developed by Google for use in deep learning. It is open source, meaning anyone can use or improve it. Similar projects include Torch and Theano.

REINFORCEMENT LEARNING A type of machine learning in which the algorithm learns by acting toward an abstract goal, such as "earn a high video game score" or "manage a factory efficiently." During training, each effort is evaluated based on its contribution toward the goal.

SUPERVISED LEARNING A type of machine learning in which the algorithm compares its outputs with the correct outputs during training. In unsupervised learning, the algorithm merely looks for patterns in a set of data.

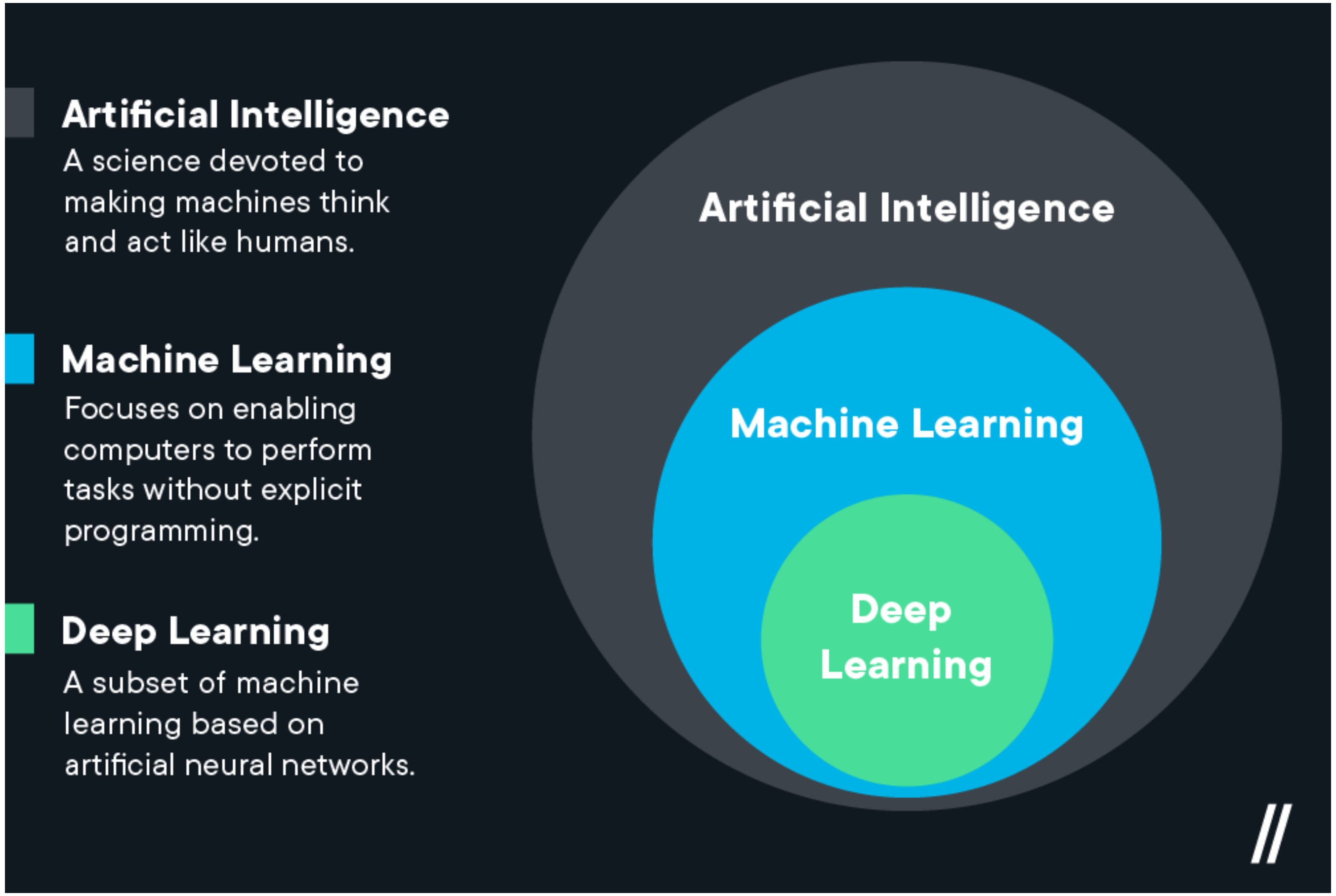
STRONG AI AI that is as smart and well-rounded as a human. Some say it's impossible. Current AI is weak, or narrow. It can play chess or drive but not both, and lacks common sense.

Source: : https://www.sciencemag.org/news/2017/07/ai-revolution-science

Machine Learning

Machine learning (ML) is a branch of AI and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.[1]

ML systems learn how to combine input to produce useful predictions on never-before-seen data.



Artificial Intelligence
A science devoted to making machines think and act like humans.

Machine Learning
Focuses on enabling computers to perform tasks without explicit programming.

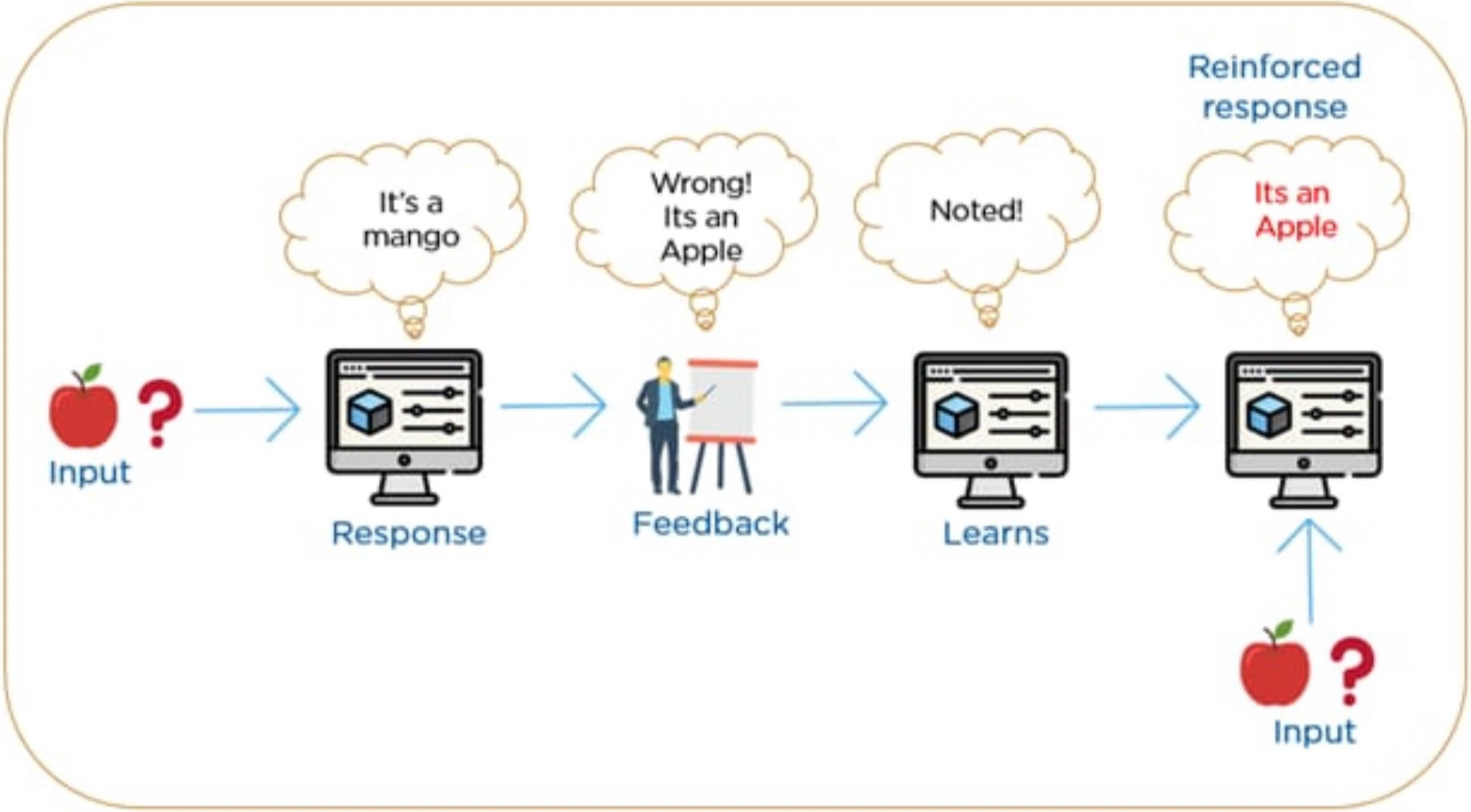
Deep Learning
A subset of machine learning based on artificial neural networks.

Source: Flatironschool

[1] IBM

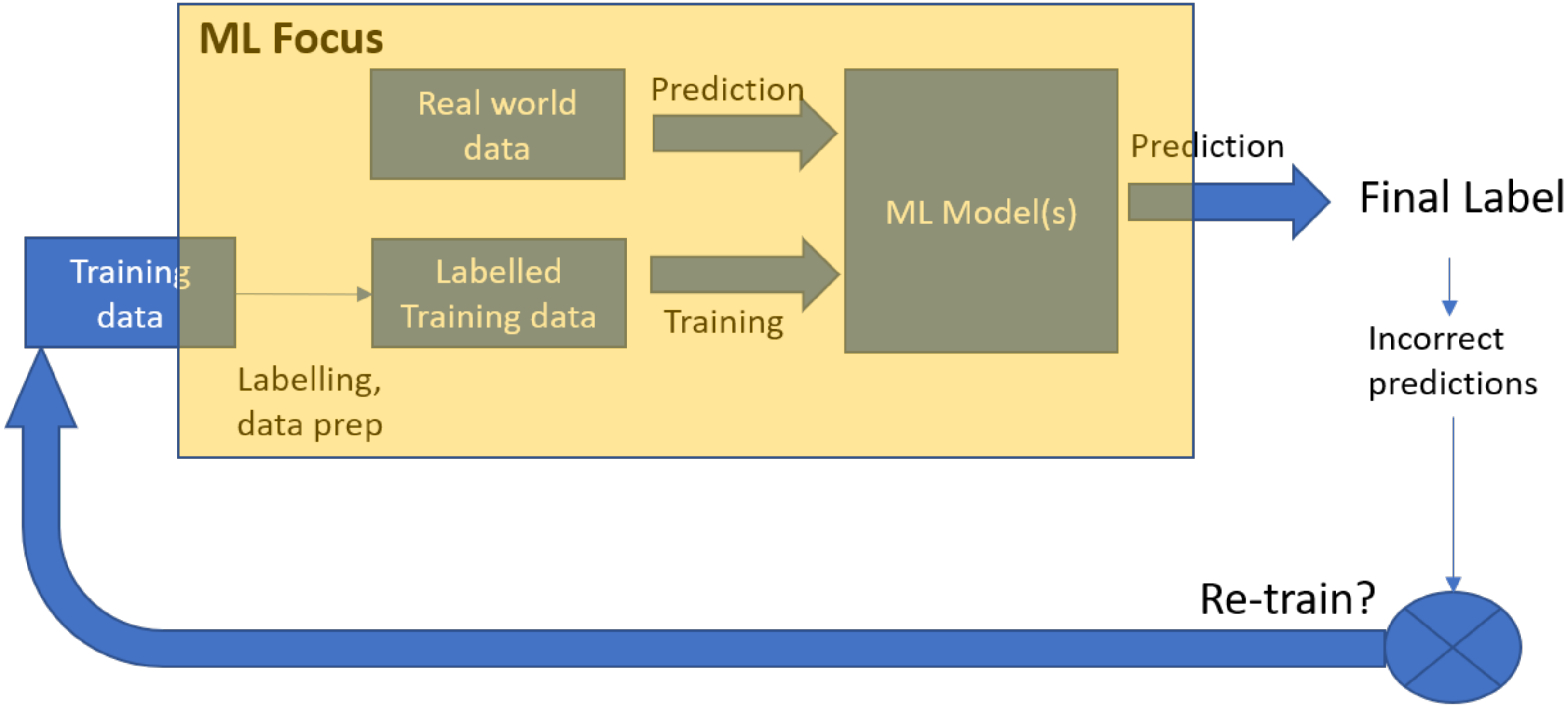
Machine Learning

ML systems learn how to combine input to produce useful predictions on never-before-seen data.



Source: IEEE PES DSCE

Machine Learning and AI



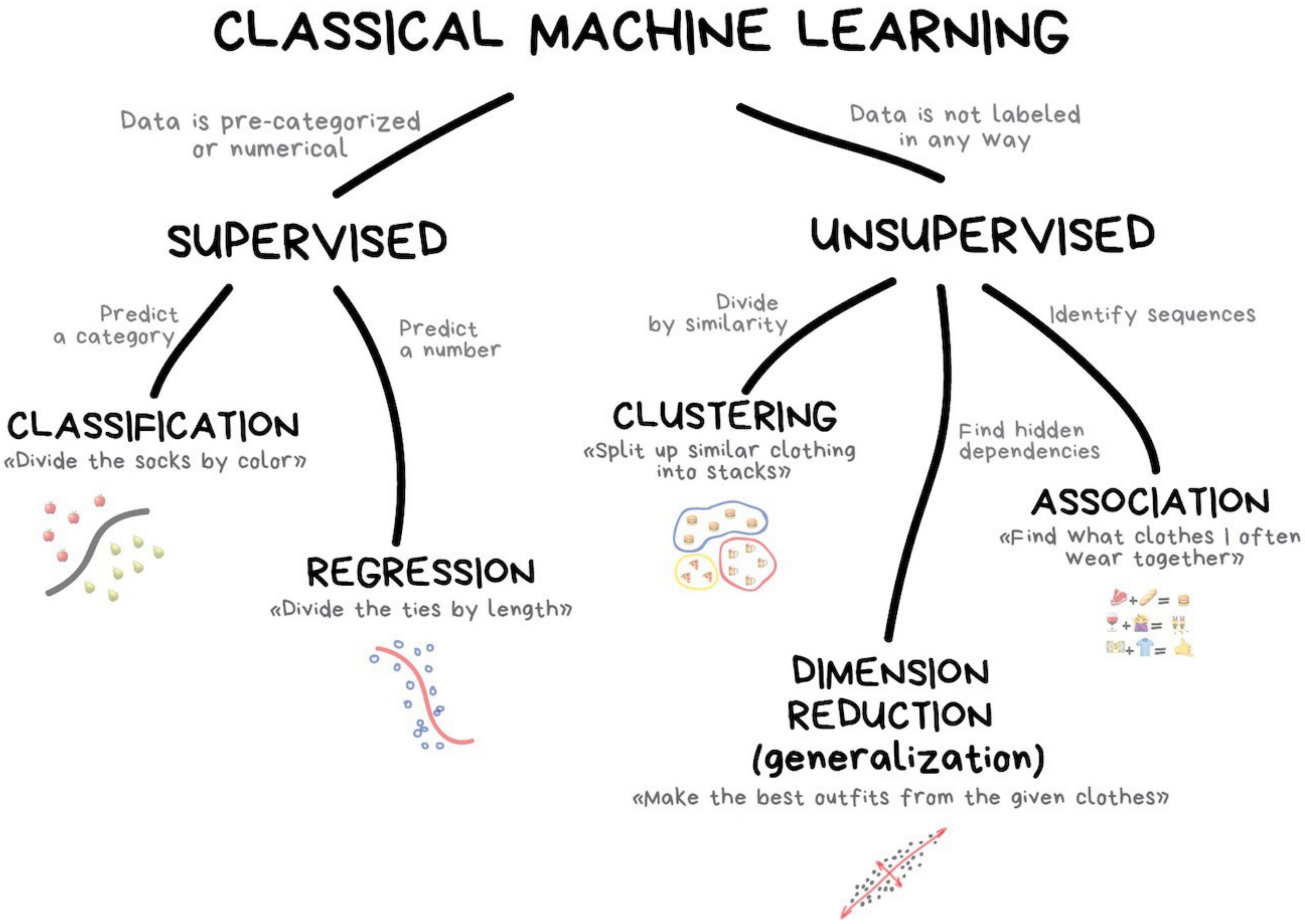
AI Focus

- System design
- Model Splitting
- Deployment
- Re-architecting upon practical limitations of training data
- Re-training strategy
- Reducing/increasing labels
- Handling incorrect predictions
- …

AI Focus : Overall system

Source: Towards Data Science

Classical ML

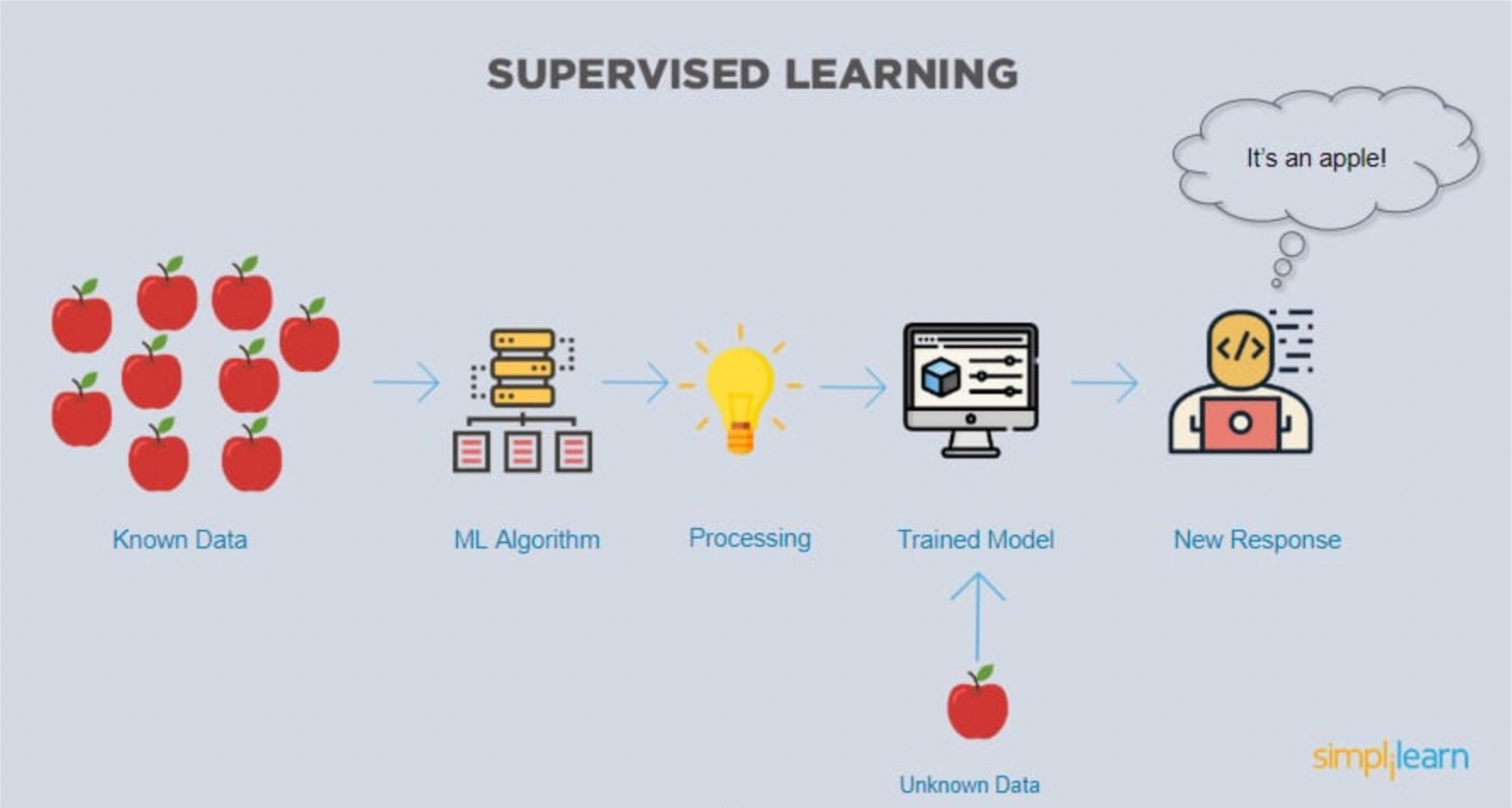

CLASSICAL MACHINE LEARNING
Data is pre-categorized or numerical
Data is not labeled in any way
SUPERVISED
UNSUPERVISED
Predict a category
Predict a number
CLASSIFICATION
«Divide the socks by color»
REGRESSION
«Divide the ties by length»
Divide by similarity
Identify sequences
Find hidden dependencies
CLUSTERING
«Split up similar clothing into stacks»
ASSOCIATION
«Find what clothes I often wear together»
DIMENSION REDUCTION (generalization)
«Make the best outfits from the given clothes»

Source: Machine Learning for Everyone

ML Learning Algorithms

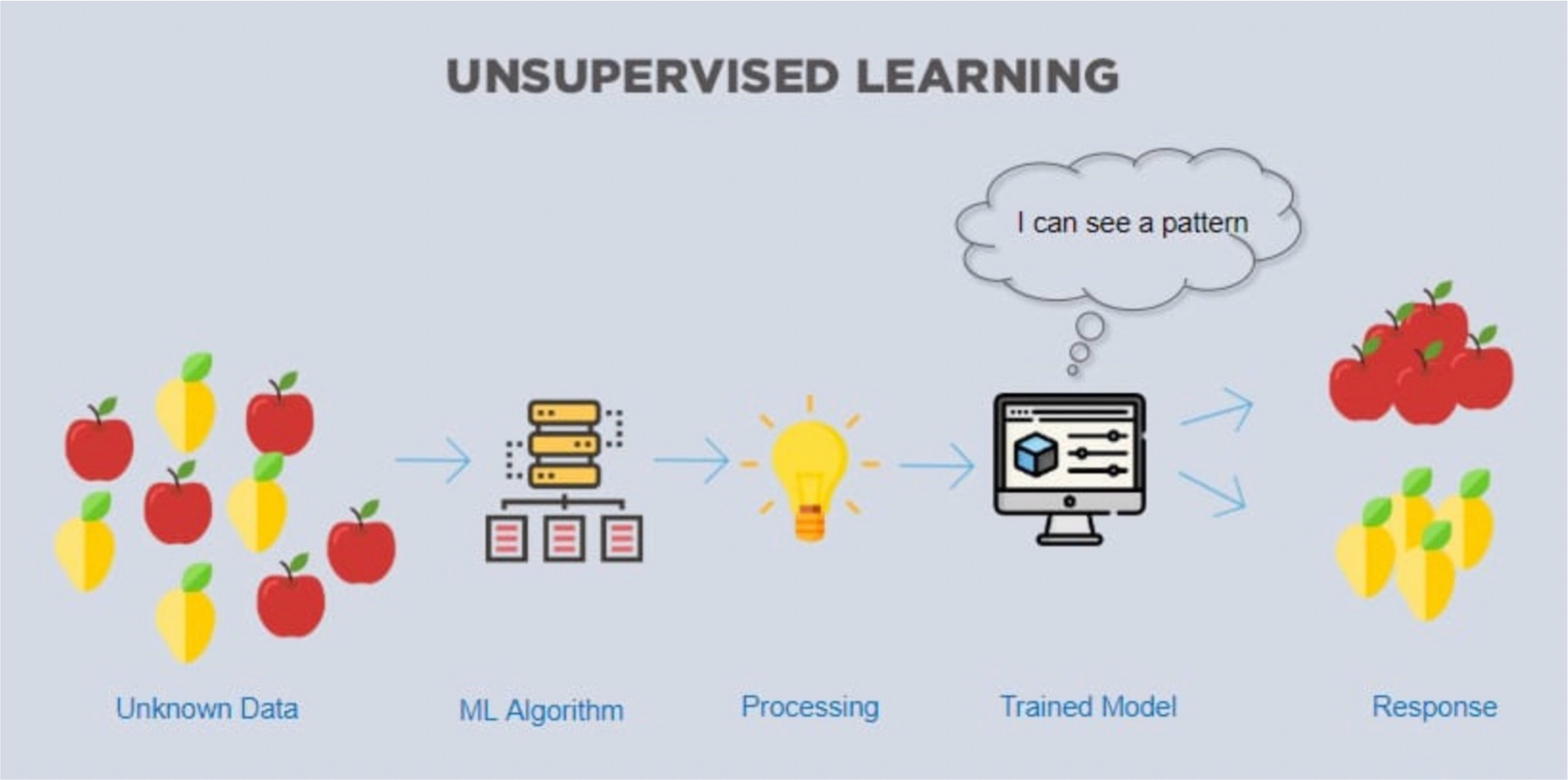
- **Supervised Learning:** learning to future outcome (**Prediction**)
  - Regression: Linear and Logistic Regression,
  - Classification, Decision Tree, Random Forests, SVM, Perceptron, kNN
  - Neural Network, Naive Bayes
- **Unsupervised Learning:** Learning to detect structures in data (**Pattern**/Structure **Detection**/Discovery)
  - K-means/BFR
  - Clustering
- **Reinforcement Learning:** Learning Series of action which can results in maximising rewards (**Optimisation**)
  - Online Learning
  - Control Learning

Supervised ML



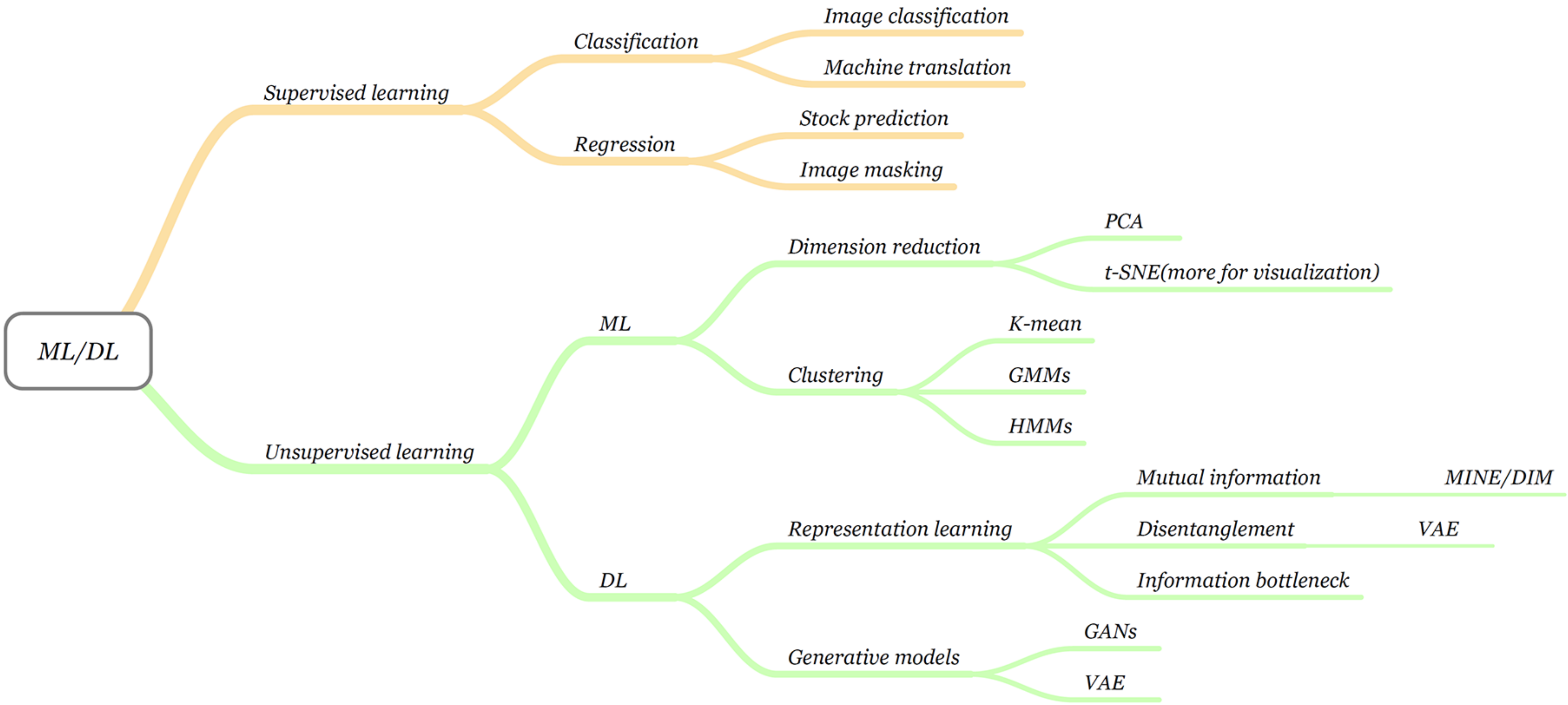
Source: Simplilearn

UnSupervised ML



Source: Simplilearn

Supervised and Unsupervised ML Algorithms Tree



Source: Towards AI

ML Terms

- **Example (sample data)**: or observation is a single data item (a particular instance of data or sample data). This is the input of the system. E.g. an image of an animal. A sample email.

- **Feature**: is an attribute of an example. An example might have many features/dimensions. Examples of features in the spam detector program:

    - words in the email text
    - sender's address
    - time of day the email was sent
    - email contains the phrase "one weird trick."

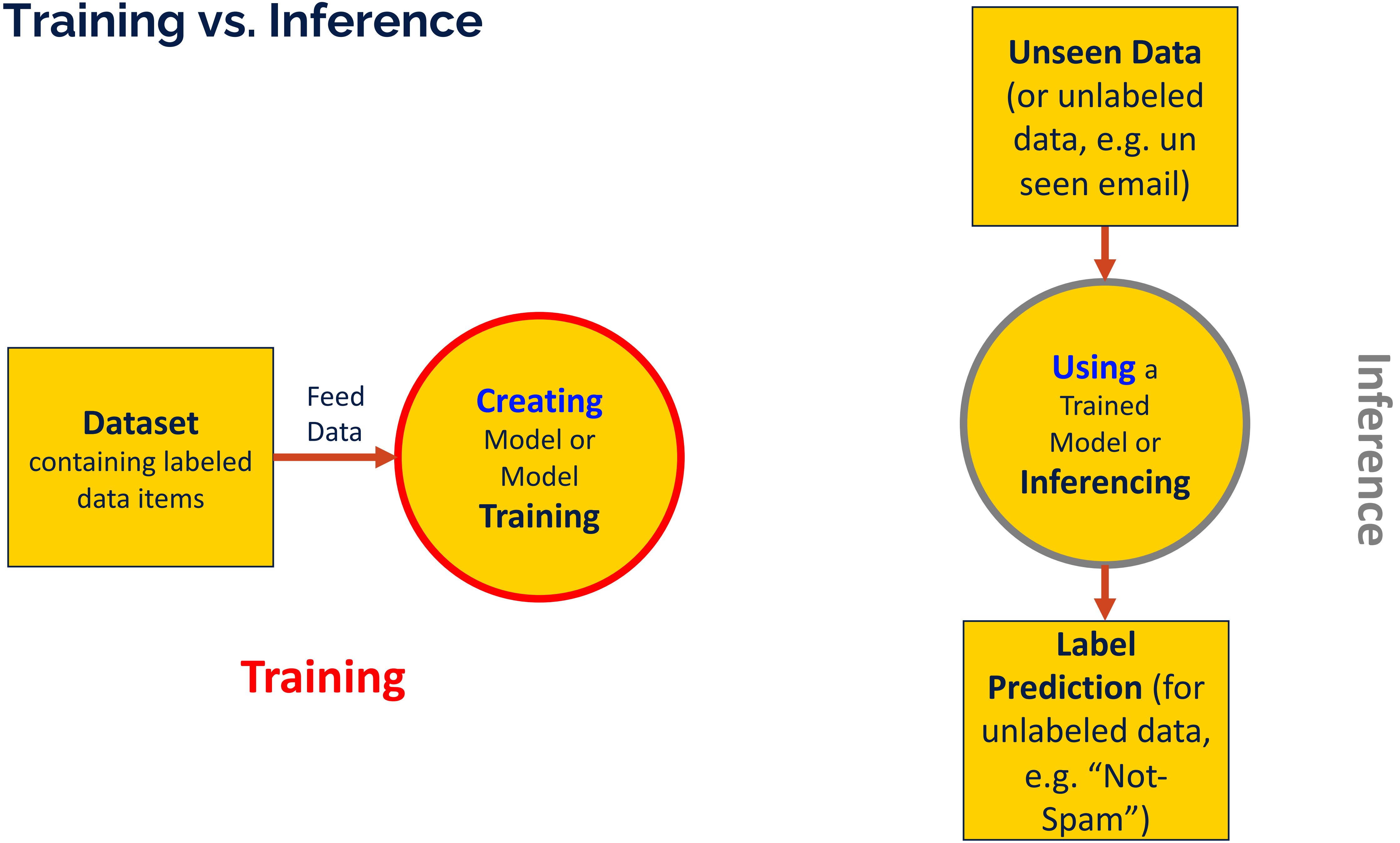
ML Terms

- **Label**: is **label** is the thing we're predicting. This is the output of the system. For example, a label “Dog” attached to an example (animal image) specifies that the image shows a dog. Spam vs None-Spam
- **Labelled Example (sample)**: is a complete pair of an input and an output (i.e. an example with a label or labelled data items). Content of an email together with its label either ”spam” or “none-spam”
- Labelled examples: {features, label}: (x, y)
- x={x1,x2,x3,..xn}
- **Data Set:** A data set is a collection of either unlabelled examples or labelled examples.

ML Terms

- **Training**: means creating or learning the model. A data set can be used to train a model.
- **Model**: defines the relationship between features and label. A model can be used to predict labels for unseen data items by running the model.
- **Training Set:** is a fraction of a dataset (generally the major part) which is only used for the purpose of model training.
- **Test Set:** is a fraction of a dataset (generally the minor part) which is only used for the purpose of assessing and evaluating the trained model (e.g. measuring the accuracy of the model).
- **Inference:** means applying the trained model to unlabelled examples

Training vs. Inference



Dataset containing labeled data items

Feed Data

Creating Model or Model **Training**

Training

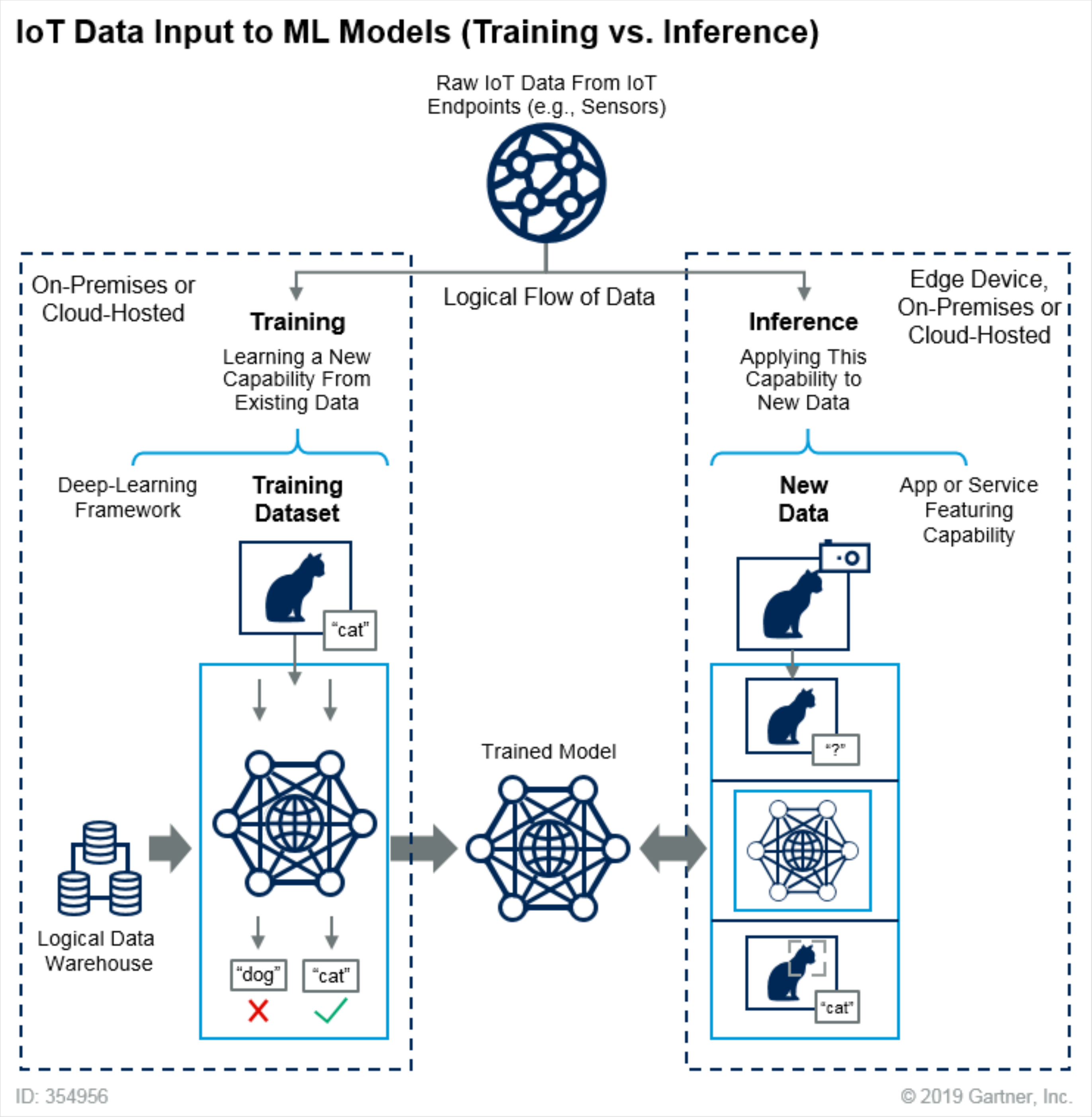
Unseen Data (or unlabeled data, e.g. un seen email)

Using a Trained Model or **Inferencing**

Label Prediction (for unlabeled data, e.g. “Not-Spam”)

Inference

Training vs. Inference
(IoT Example)


IoT Data Input to ML Models (Training vs. Inference)
Raw IoT Data From IoT Endpoints (e.g., Sensors)
On-Premises or Cloud-Hosted
Training
Learning a New Capability From Existing Data
Logical Flow of Data
Inference
Applying This Capability to New Data
Edge Device, On-Premises or Cloud-Hosted
Deep-Learning Framework
Training Dataset
"cat"
New Data
App or Service Featuring Capability
"?"
Trained Model
Logical Data Warehouse
"dog"
"cat"
"cat"
ID: 354956
© 2019 Gartner, Inc.

ML and Big Data Challenges

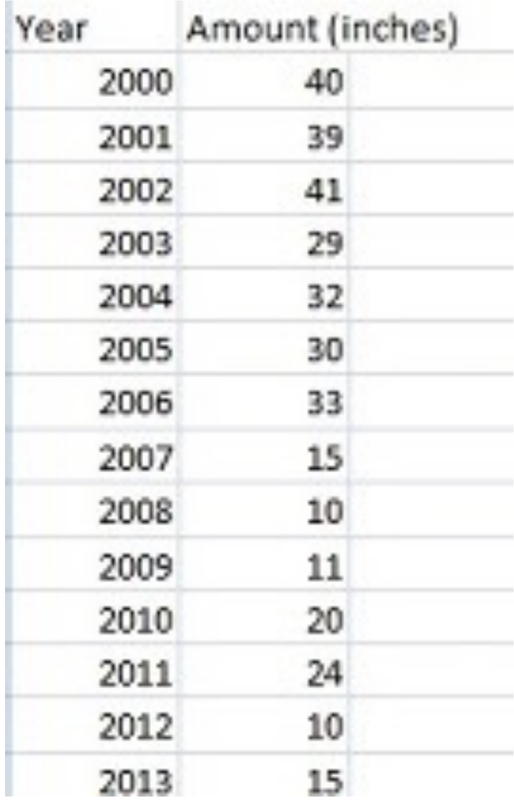
- **Veracity:** Data is in doubt (due to inconsistency, incompleteness, latency, missing, noisy and corrupted), Data Lacks properly understand structure
- **Variety:** Data is in heterogenous types
- **Volume:** High dimensional data, Large data sets often can't be accessed easily and directly, and Large data sets can't be processed centrally.
- **Velocity:** Data arrives at different speeds, and they may change over time (e.g. data streaming)

ML Solutions/Algorithms for Challenges

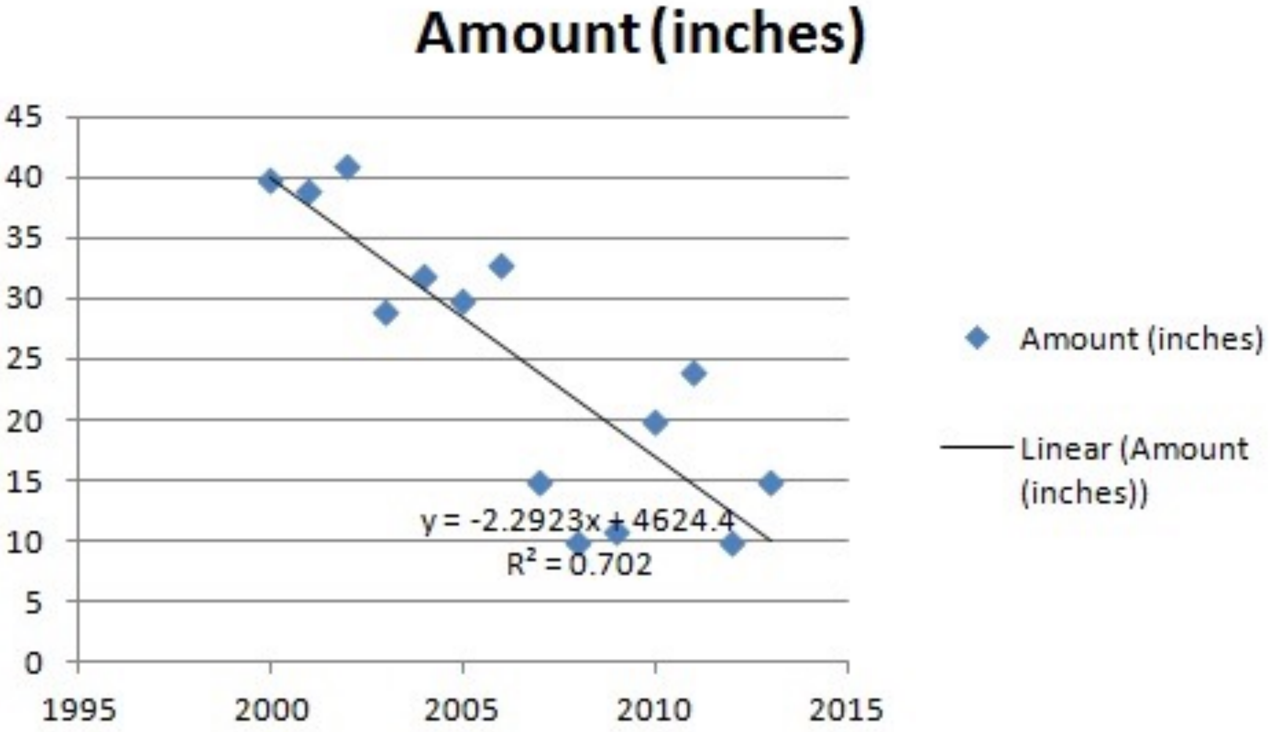
Solution	Algorithms	Veracity Issues	Variety Issues	Volume Issues	Velocity Issues
Preprocessing	Regression/Substitution , Dim. Reduction	noisy, corrupted, incomplete, missing		high-dimensional, can’t be accessed directly, processed centrally	
Geometric Techniques	SVMs, perceptron, kNN, clustering	structure, incomplete, missing	heterogeneous	high-dimensional	
Network Algorithms	Random Walks, Neural Networks	structure, incomplete, missing	heterogeneous	high-dimensional	
Online Learning	Sublinear, Streaming Algorithms, Regret Minimisation			can’t be accessed directly, processed centrally	change over time

Regression vs. Classification

Regression: aims to predict trends of data (in a quantitative manner). Also, predicts continuous values. For example, *rainfall* prediction using a simple linear regression



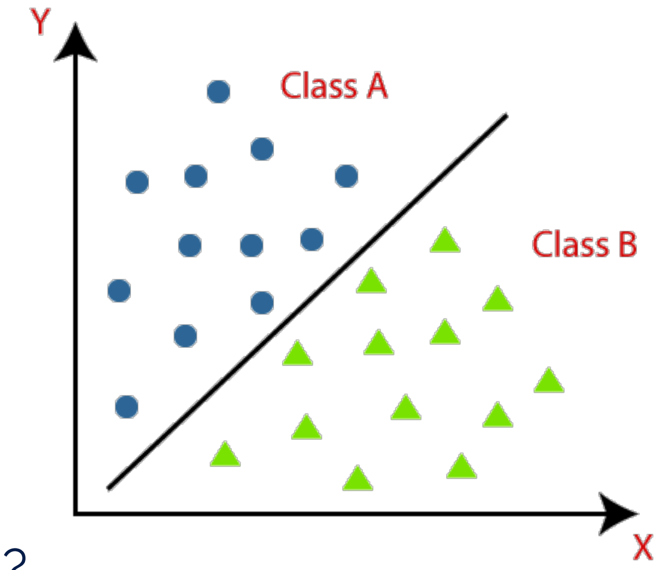
Year	Amount (inches)
2000	40
2001	39
2002	41
2003	29
2004	32
2005	30
2006	33
2007	15
2008	10
2009	11
2010	20
2011	24
2012	10
2013	15



Source: Statisticshowto

Regression vs. Classification

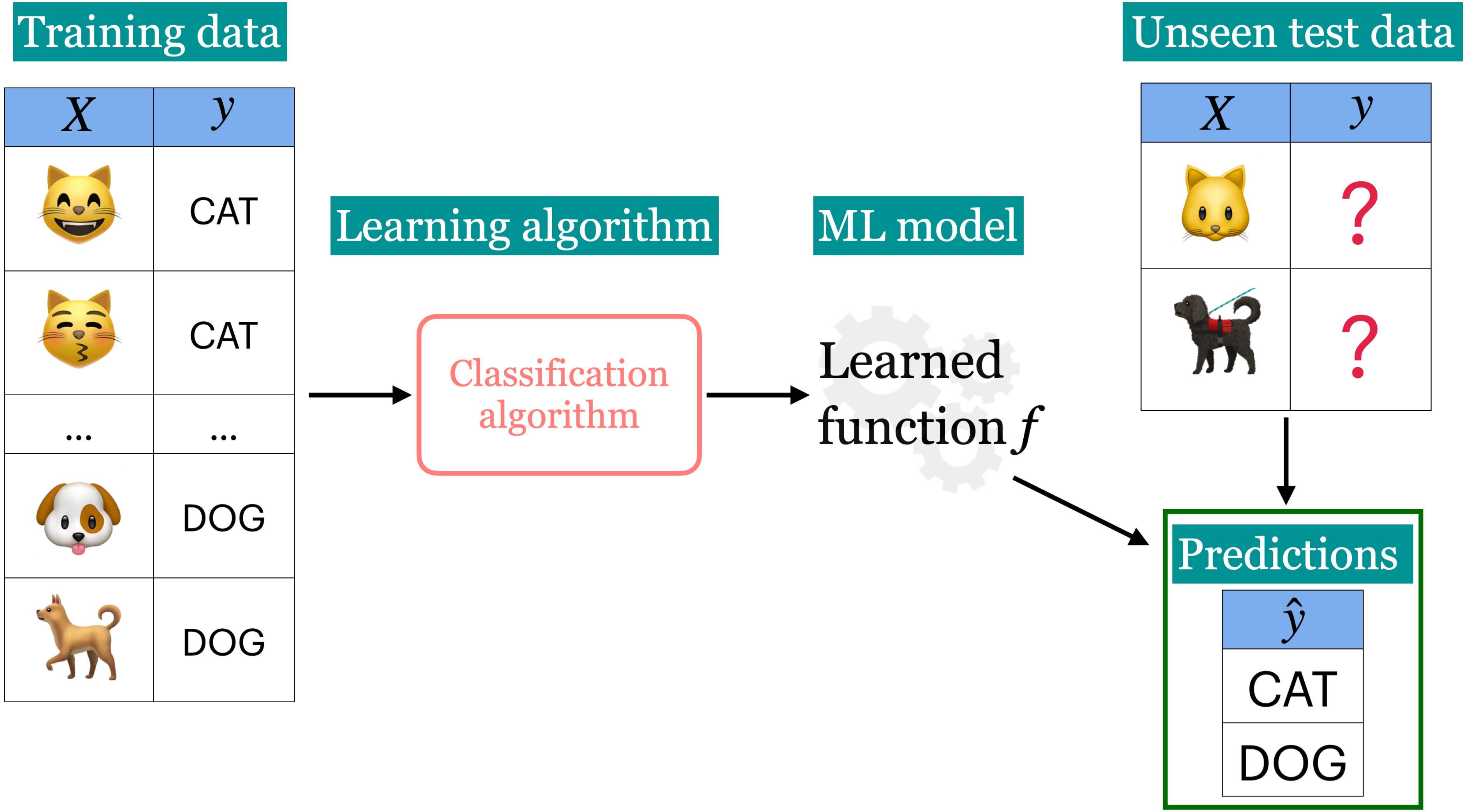
Classification: provides qualitative prediction for an observation (or classifying an observation by assigning a class or category to the observation), predicts discrete values.



- Examples of classification questions:
  - Is a given email spam or not spam?
  - Is this an image of a cat, a dog, horse, elephant, or others?
  - Is a given software malicious, or normal software?
  - Is a given website trustable or not trustable?
  - Is a given program infected, or not-infected?
  - Is a given malware adware, spyware, ransomware, or others?
  - Is a given fruit (i.e. fruit description/data) Apple, Mandarin, Lemon, or Orange?

Source: Towards Data Science

Example of ML



Training data

X	y
😸	CAT
😽	CAT
...	...
🐶	DOG
🐕	DOG

Learning algorithm

Classification algorithm

ML model

Learned function f

Unseen test data

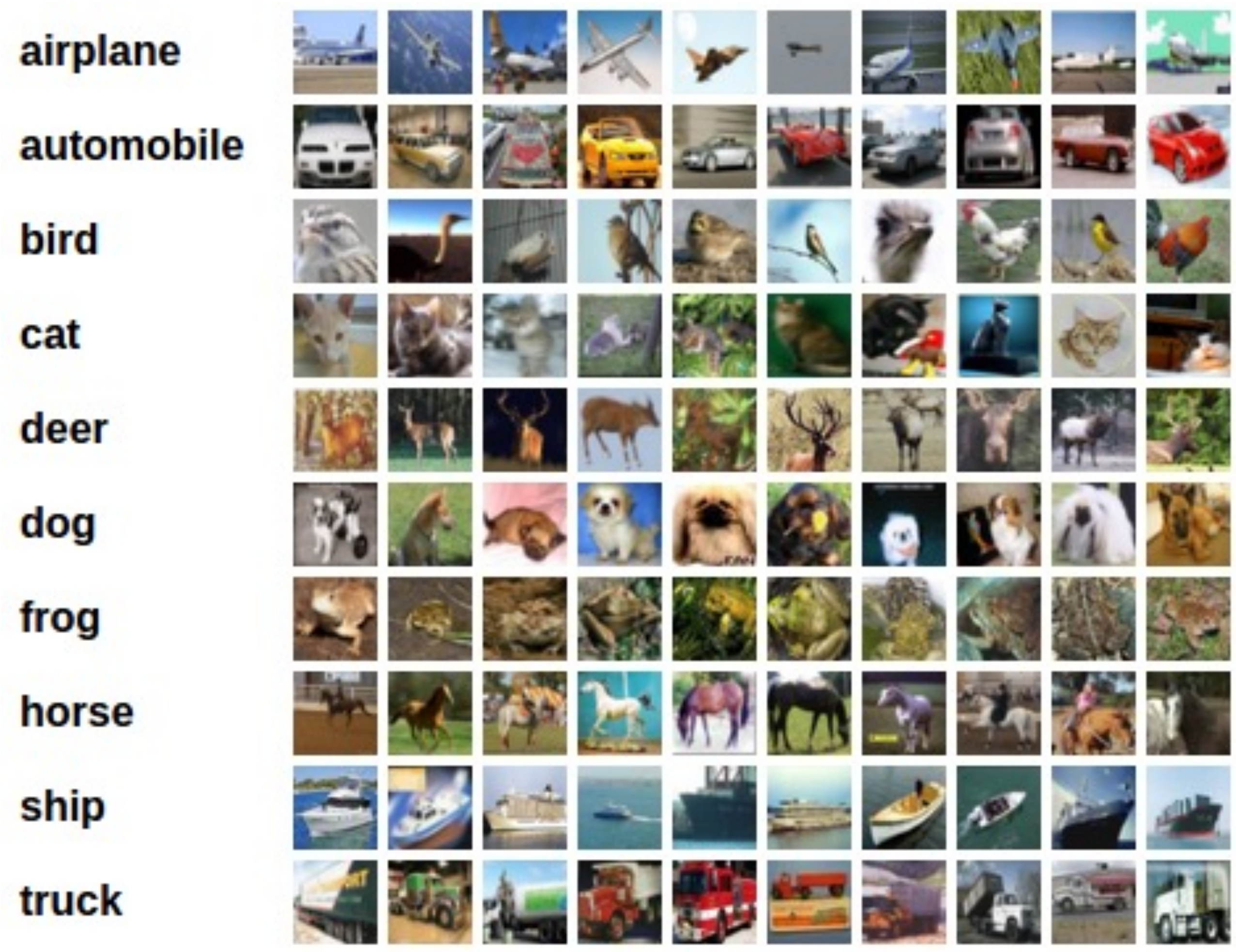
X	y
🐱	?
🐕‍🦺	?

Predictions

$\hat{y}$
CAT
DOG

Source: https://ubc-cs.github.io/cpsc330/lectures/02_decision-trees.html

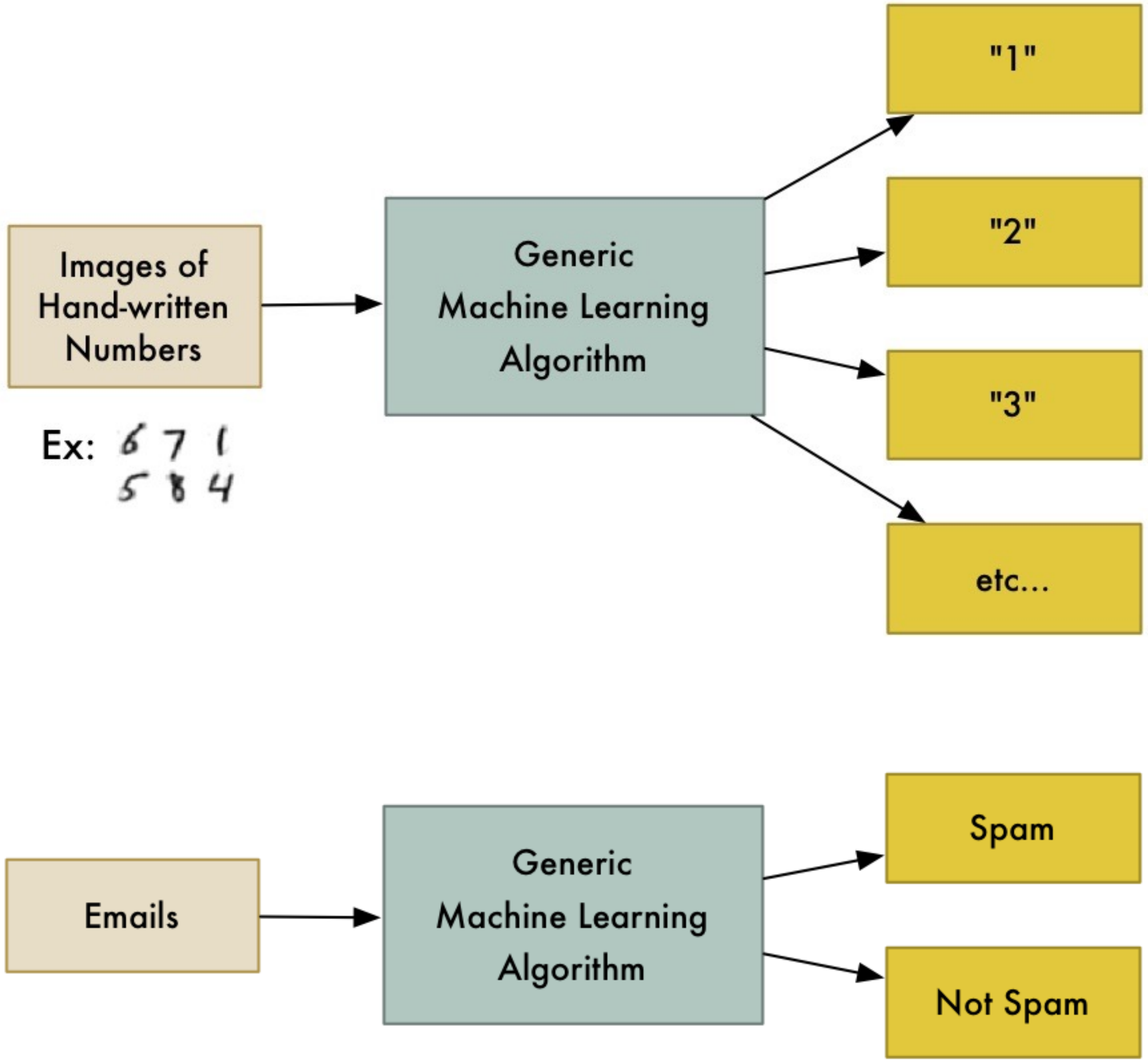
Example of ML



Classified Training Data

Source: https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html

Spam Detection



Images of Hand-written Numbers

Ex: 6 7 1
5 8 4

Generic Machine Learning Algorithm

"1"

"2"

"3"

etc...

Emails

Generic Machine Learning Algorithm

Spam

Not Spam

Source: Towards AI