SWE30011 Internet of Things Programming

WEEK 11 Lecture

Advance Topics in IoT: Machine Learning

Dr Felip Marti

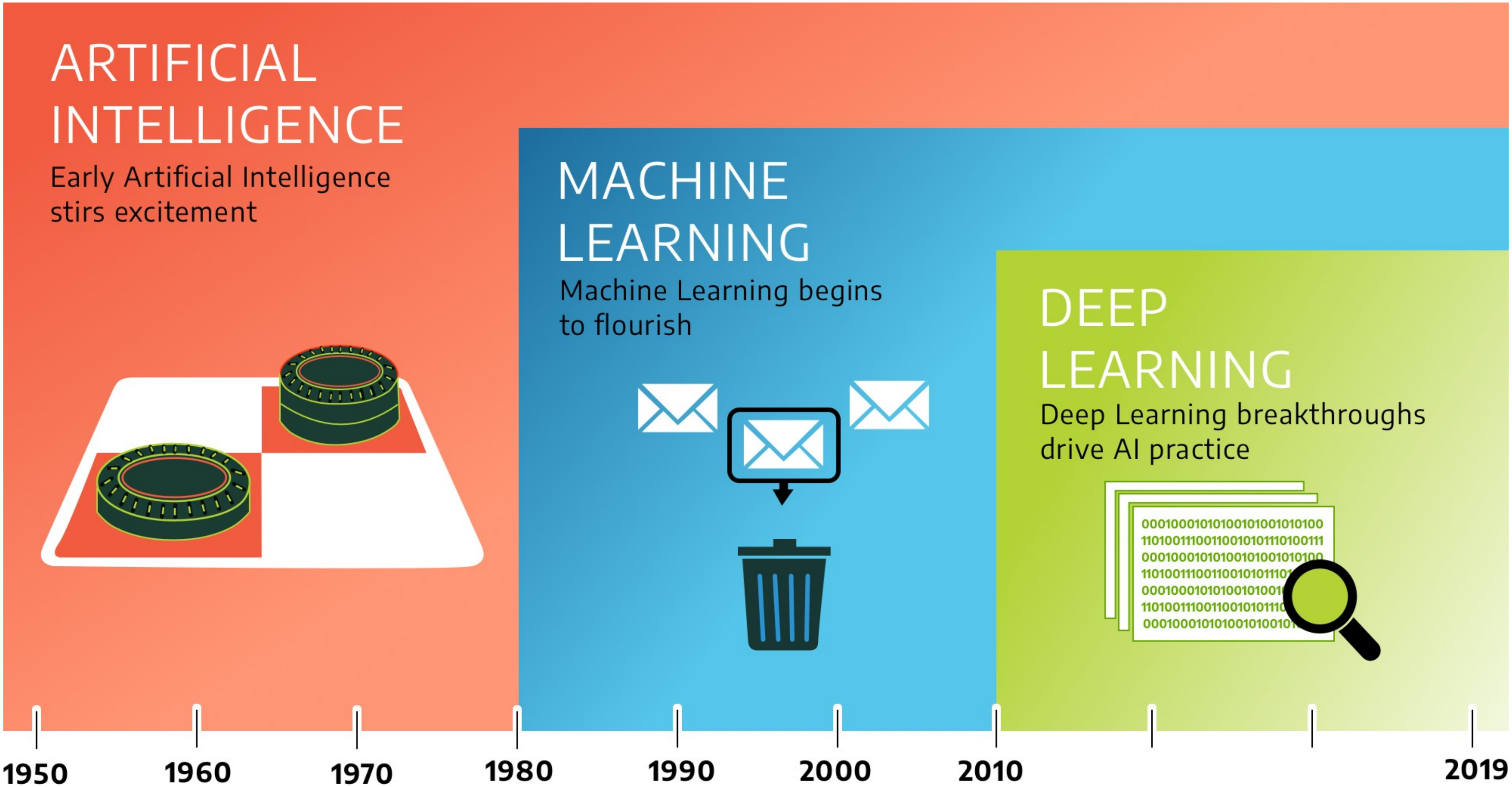
fmarti@swin.edu.au

Swinburne University of Technology

May 2021

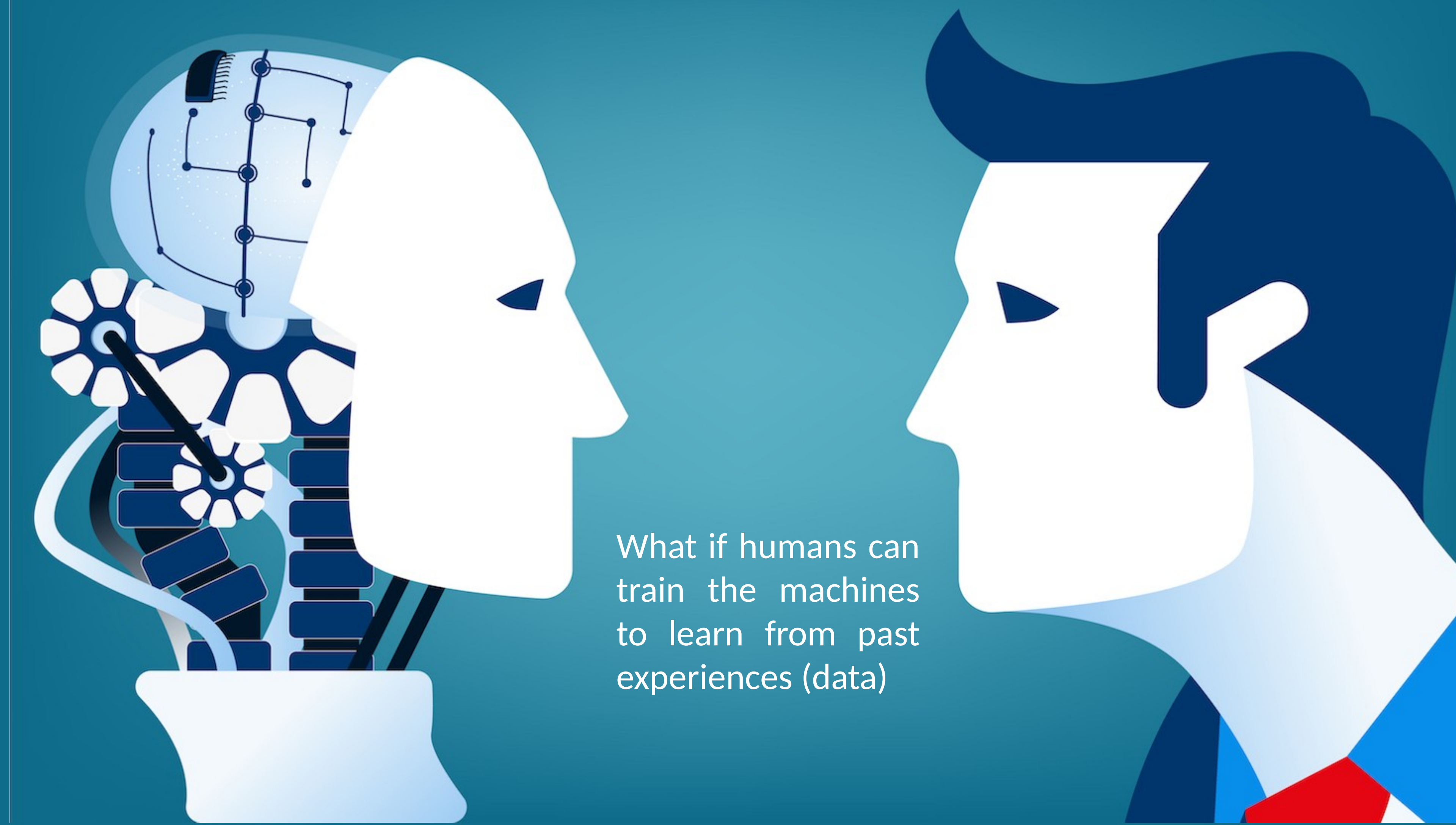
[Material originally prepared by Dr Ali Yavari]

PART-1 What is Machine Learning



https://www.wavelabs.ai/wavelabs-conducts-deep-learning-workshop-in-association-with-malai-club/

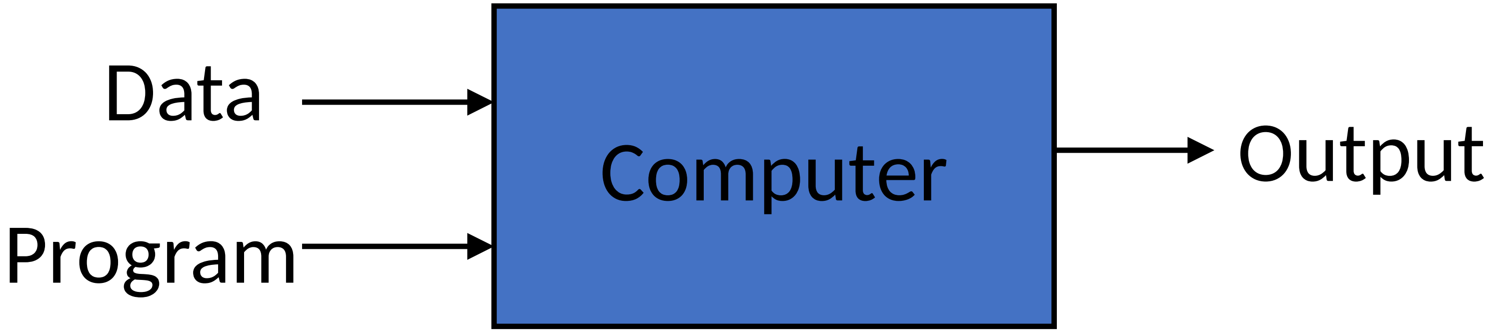
Beginning of Artificial Intelligence:
A. M. Turing (1950) Computing Machinery and Intelligence. Mind 49: 433-460.
https://www.csee.umbc.edu/courses/471/papers/turing.pdf



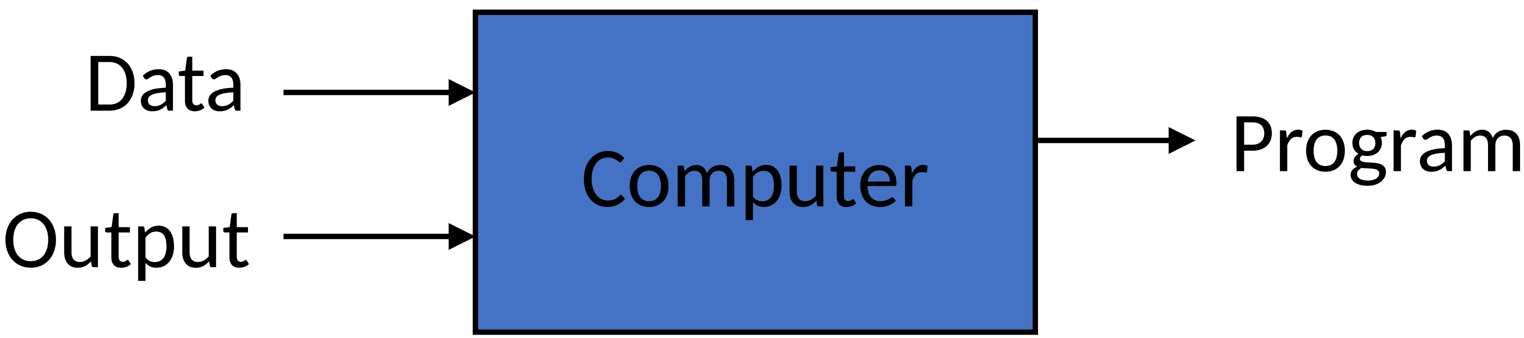
Machines follow instructions

Humans learn from past experiences

Traditional Programming



Machine Learning



Slide credit :Pedro Domingos

Definitions of machine learning

- “Field of study that gives computers the ability to learn without being explicitly programmed”

Arthur Samuel (1959)

- ”Computer program is said to learn from:
    - experience *E*
    - with respect to some class of tasks *T*
    - and performance measure *P*
    - if its performance at tasks in *T*, as measured by *P*, improves with experience *E*.”

– Tom M. Mitchell

Machine Learning, Tom Mitchell, McGraw Hill, 1997.

http://www.cs.cmu.edu/~tom/mlbook.html

Why Machine Learning

- Some tasks cannot be defined well, except by example
- Relationships and correlations can be hidden within large amount of data.
- Amount of knowledge available about a certain task might be too large for explicit encoding
- New knowledge about task is constantly being discovered
- Machine learning is a form of AI that enables a system to learn from data rather than through explicit programming.

Data everywhere!

1. **Google:** processes 24 peta bytes of data per day.
2. **Facebook:** 10 million photos uploaded every hour.
3. **Youtube:** 1 hour of video uploaded every second.
4. **Twitter:** 400 million tweets per day.
5. **Astronomy:** Satellite data is in hundreds of PB.
6. ...
7. **"By 2020 the digital universe will reach 44 zettabytes..."**

The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, April 2014.

That's 44 trillion gigabytes!

Some applications

- Robotics.
- Web search.
- Speech recognition.
- Translation.
- Space exploration.
- Genomics.

Machine Learning Systems

- Training is a critical step in the machine learning process
- When you're training a machine learning system, you know the inputs (for example customer income, buying history, location, and so on), and you know your desired goal (predicting a customer's propensity to churn).
- Training a machine learning algorithm to create an accurate model can be broken down into three steps:
  - Representation
  - Evaluation
  - Optimisation

Slide credit :Pedro Domingos https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf

Representation

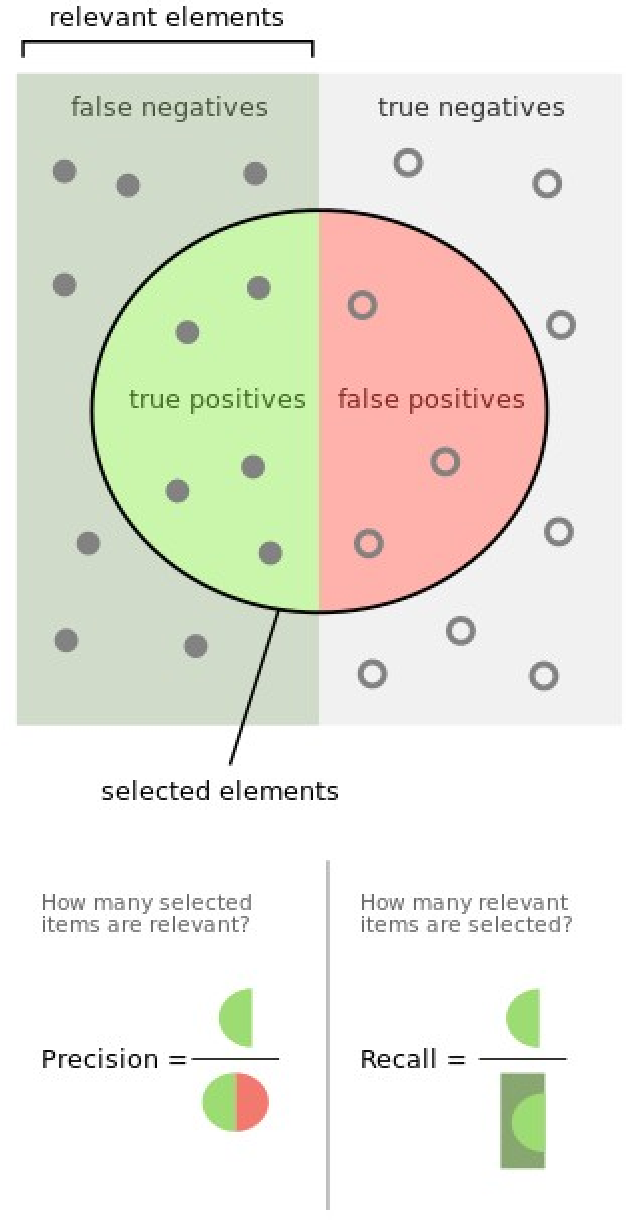
A classifier must be represented in some formal language that the computer can handle.

- Decision trees
- Sets of rules / Logic programs
- Instances
- Graphical models (Bayes/Markov nets)
- Support vector machines
- Model ensembles
- The algorithm creates a model to transform the inputted data into the desired results.

Slide credit :Pedro Domingos

Evaluation

- As the algorithm creates multiple models, either a human or the algorithm will need to evaluate and score the models based on which model produces the most accurate predictions.
- Accuracy
- Precision and recall
- Etc.

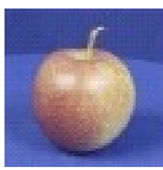


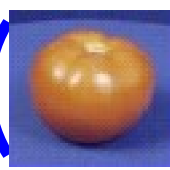
Optimization

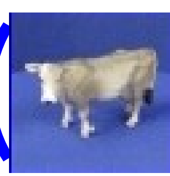
- After the algorithm creates and scores multiple models, select the best performing algorithm.
- As you expose the algorithm to more diverse sets of input data, select the most generalised model
- The most important part of the training process is to have enough data so you’re in a position to test your model

The machine learning framework

- Apply a prediction function to a feature representation of the image to get the desired output:

f() = “apple”

f() = “tomato”

f() = “cow”

SourceL. Lazebnik

The machine learning framework

$$\mathbf{y = f(x)}$$

output

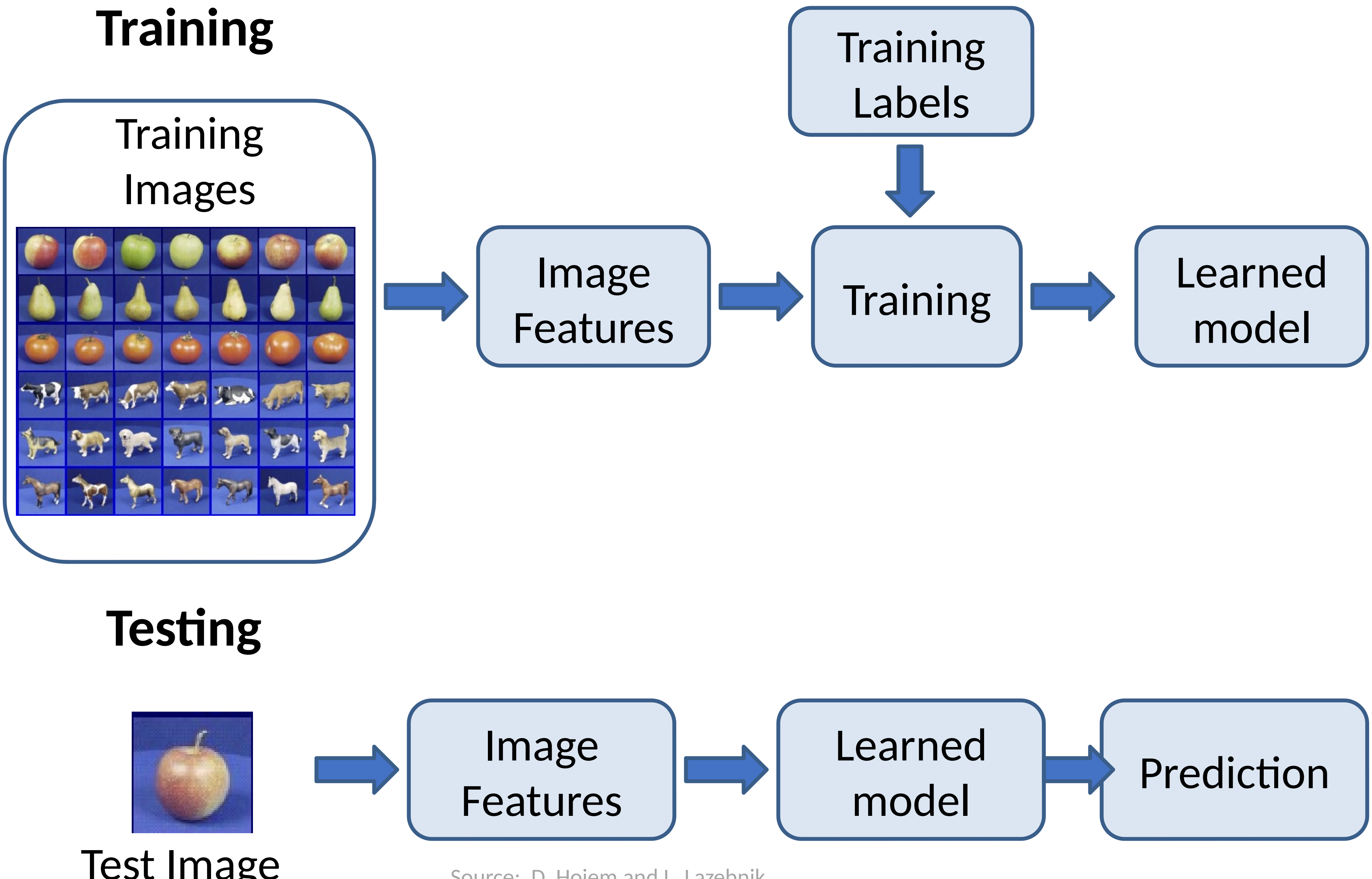
prediction function

Image feature

- **Training:** given a *training set* of labeled examples $\{(\mathbf{x}_1,y_1), \ldots, (\mathbf{x}_N,y_N)\}$, estimate the prediction function f by minimizing the prediction error on the training set
- **Testing:** apply f to a never before seen *test example* $\mathbf{x}$ and output the predicted value $y = f(\mathbf{x})$

Source: L. Lazebnik

Steps



Source: D. Hoiem and L. Lazebnik

PART-II Types of Machine Learning

Machine learning ⊆ artificial intelligence

ARTIFICIAL INTELLIGENCE

Design an intelligent agent that perceives its environment and makes decisions to maximize chances of achieving its goal.
Subfields: vision, robotics, machine learning, natural language processing, planning, …

MACHINE LEARNING

Gives "computers the ability to learn without being explicitly programmed" (Arthur Samuel, 1959)

SUPERVISED LEARNING

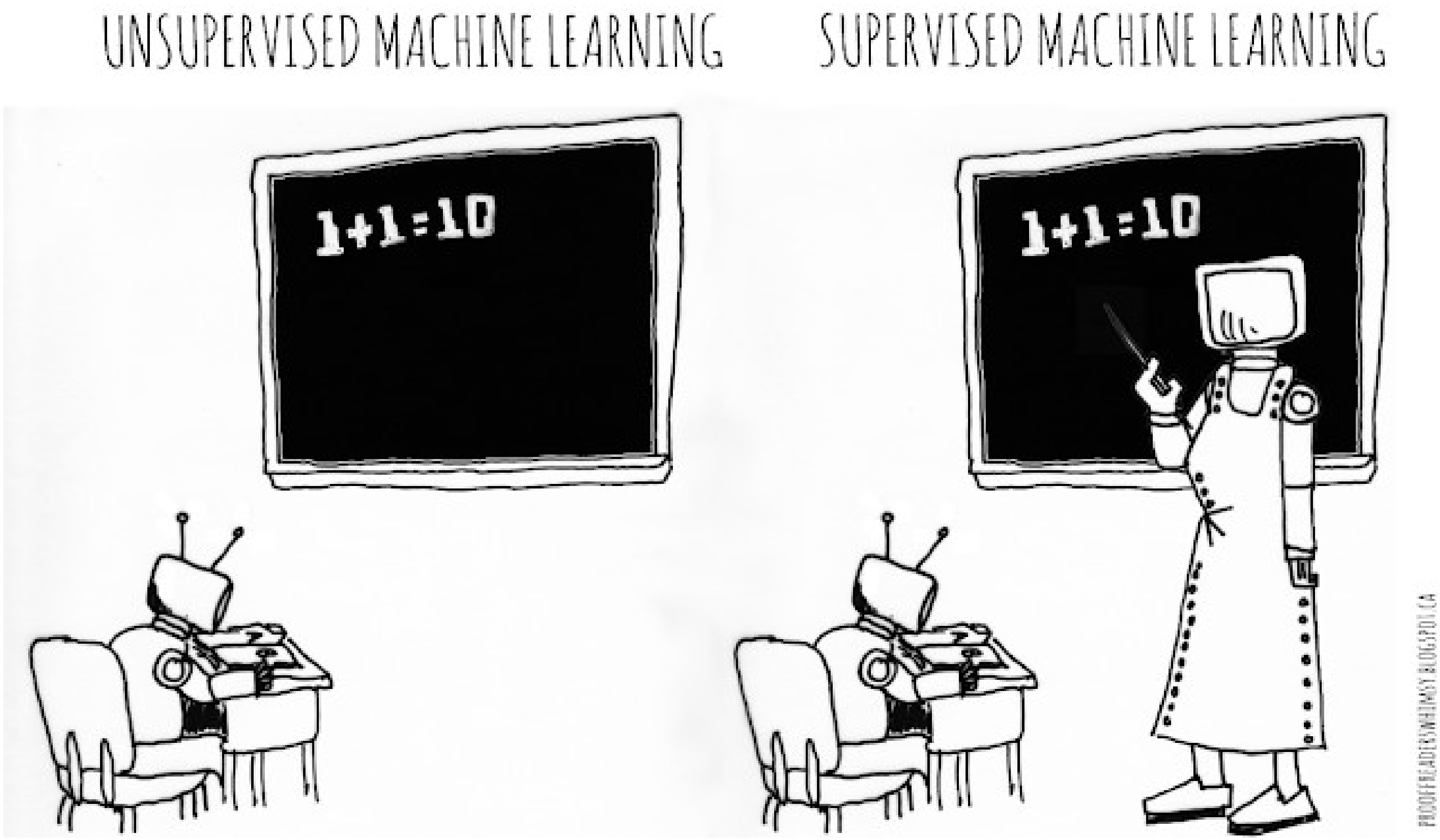
Classification, regression

UNSUPERVISED LEARNING

Clustering, dimensionality reduction, recommendation

REINFORCEMENT LEARNING

Reward maximization



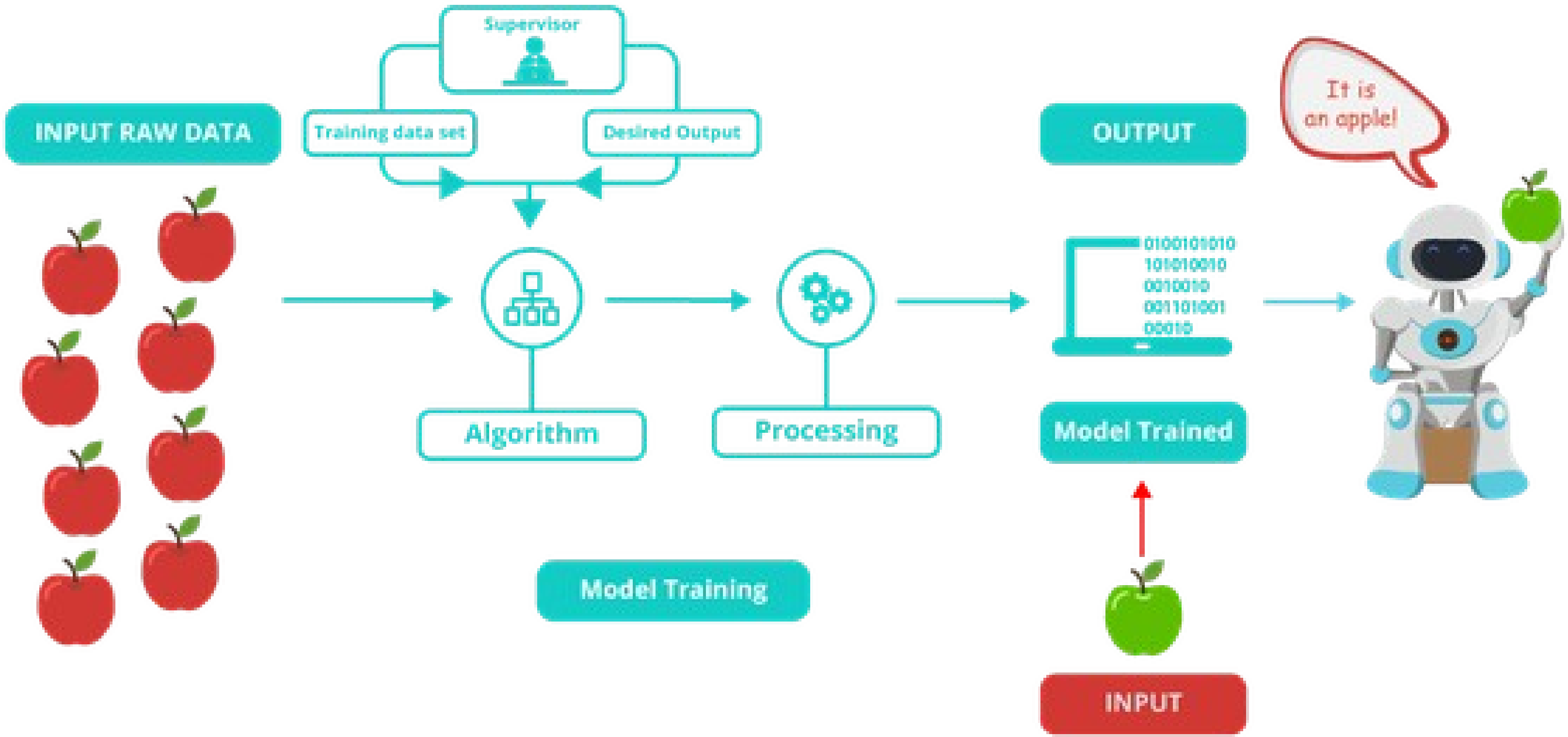
Source: http://prooffreaderswhimsy.blogspot.com/2014/11/machine-learning.html

Supervised Learning

Supervised learning typically begins with an established set of data and a certain understanding of how that data is classified

Supervised learning is intended to find patterns in data that can be applied to an analytics process

Supervised learning helps you understand the correlation between variables



https://medium.com/@gowthamy/machine-learning-supervised-learning-vs-unsupervised-learning-f1658e12a780

Classification

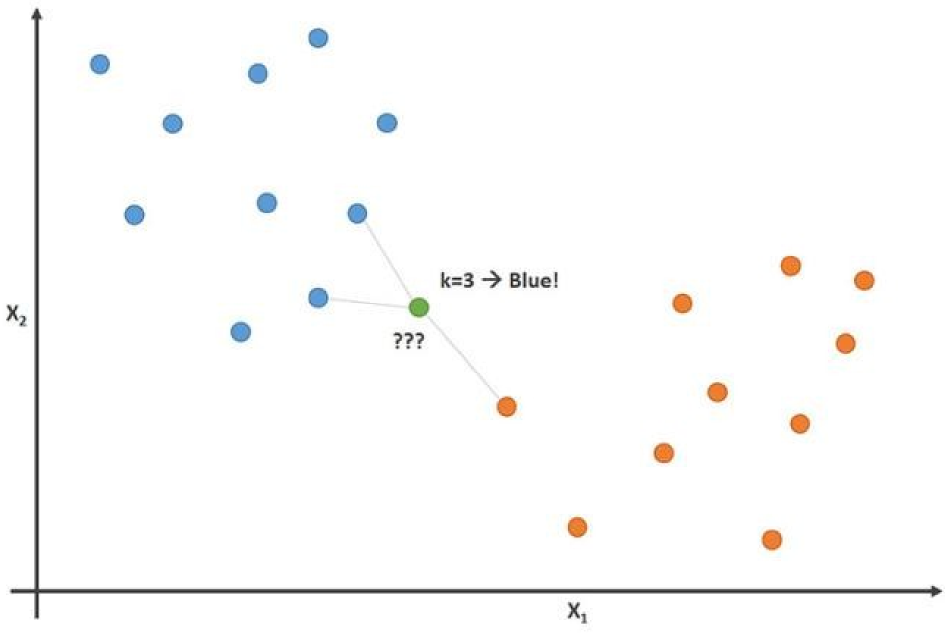
$$f = R^n \rightarrow \{1,2,\ldots,k\}$$

Example:
Object recognition: where the input is an image, and the output is a numeric code identifying the object in the image.

- In this type of task, the computer program is asked to specify which of k categories some input belongs to. To solve this task, the learning algorithm is usually asked to produce a function
.
- When y =f(x), the model assigns an input described by vector **x** to a category identified by numeric code y. There are other variants of classification task, for example where f outputs a probability distribution over classes.

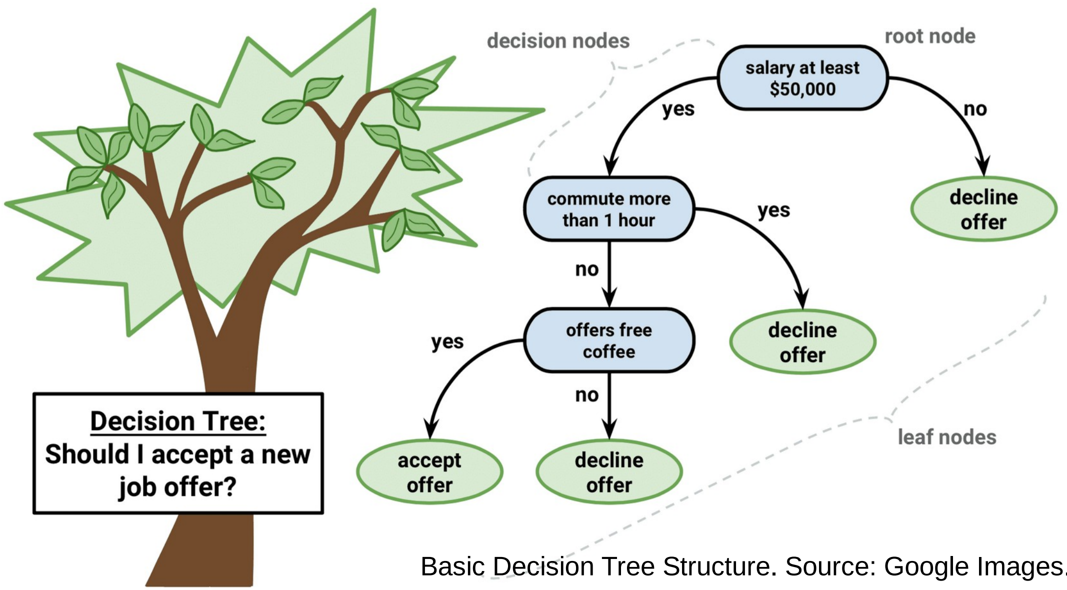
K-nearest Neighbours

- K-Nearest Neighbours (K-NN) is one of the simplest machine learning algorithms
- When a new situation occurs, it scans through all past experiences and looks up the k closest experiences
- Those experiences (or: data points) are what we call the k nearest neighbours.



Decision Tree

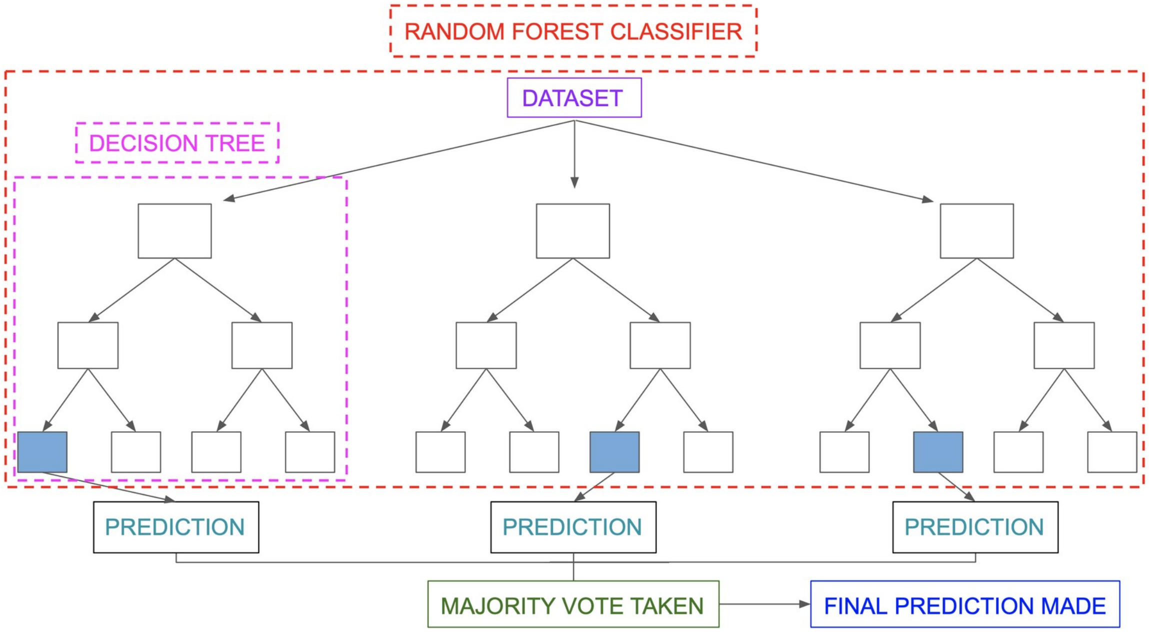
- Decision tree algorithms use a branching structure to illustrate the results of a decision.
- Decision trees can be used to map the possible outcomes of a decision.
- Each node of a decision tree represents a possible outcome.
- Percentages are assigned to nodes based on the likelihood of the outcome occurring



Basic Decision Tree Structure. Source: Google Images.

Random Forest

- The algorithm deducts the classification label for new documents from a set of decision trees where for each tree, a sample is selected from the training data, and a decision tree is created by choosing a random subset of all features (hence “Random”).
- The major drawback is performance as a large number of trees may make the method slow for real-time prediction.



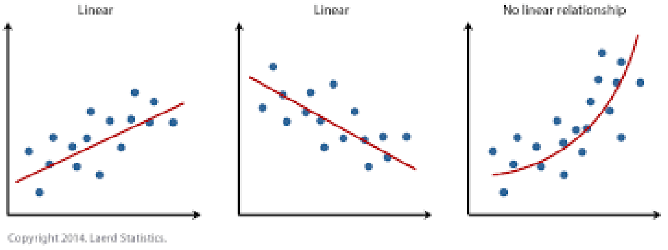
Regression

- In this type of task, the computer program is asked to predict a numeric value given some input. To solve this task, the learning algorithm is asked to output a function

$$f = \Re^n \rightarrow \Re$$

- This type of task is similar to classification, except that the format of output is different.

- **Example:** an example of regression task is the prediction of the expected claim amount that an insured person will make (used to set insurance premiums), or the predictions of future prices of securities.



Linear Regression

Regression algorithms are commonly used for statistical analysis and are key algorithms for use in machine learning

Regression algorithms help analysts model relationships between data points

Regression algorithms can quantify the strength of correlation between variables in a data set

Unsupervised Learning

Unsupervised learning is best suited when the problem requires a massive amount of data that is unlabelled

For example, social media applications, such as Twitter have large amounts of unlabelled data

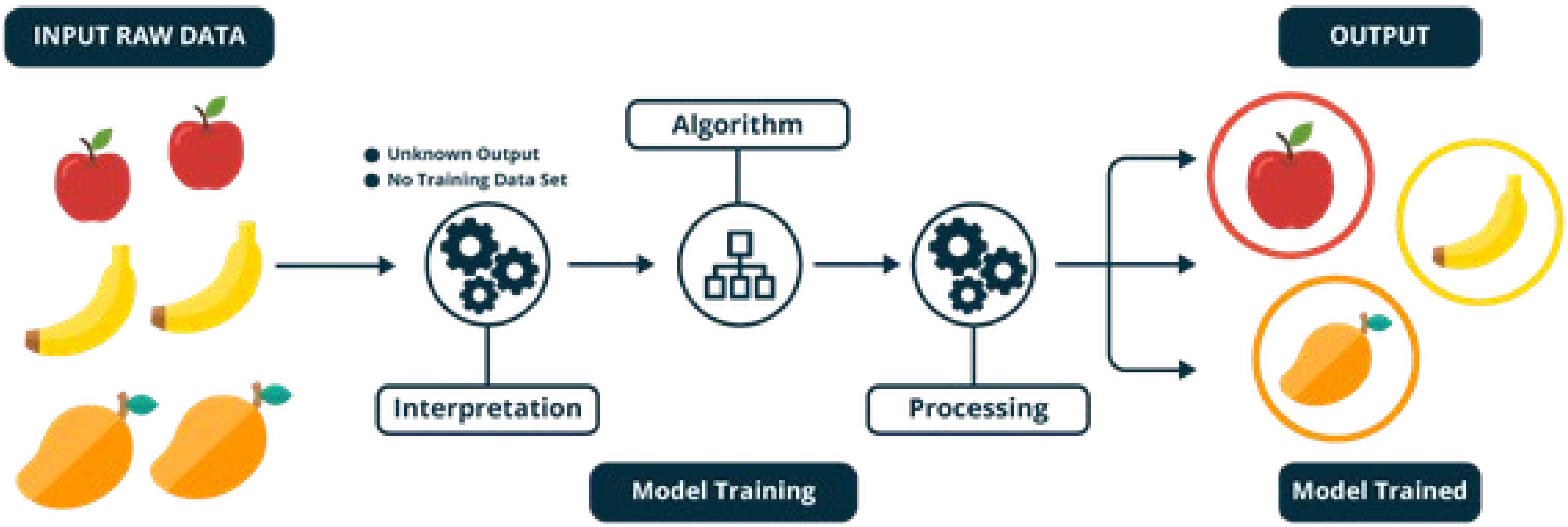
Unsupervised learning is used with email spam-detecting technology

There are far too many variables in spam emails for an analyst to flag

Unsupervised learning algorithms segment data into groups of examples (clusters) or groups of features

For example, in healthcare, collecting huge amounts of data about a specific disease can help practitioners gain insights into the patterns of symptoms and relate those to outcomes from patients

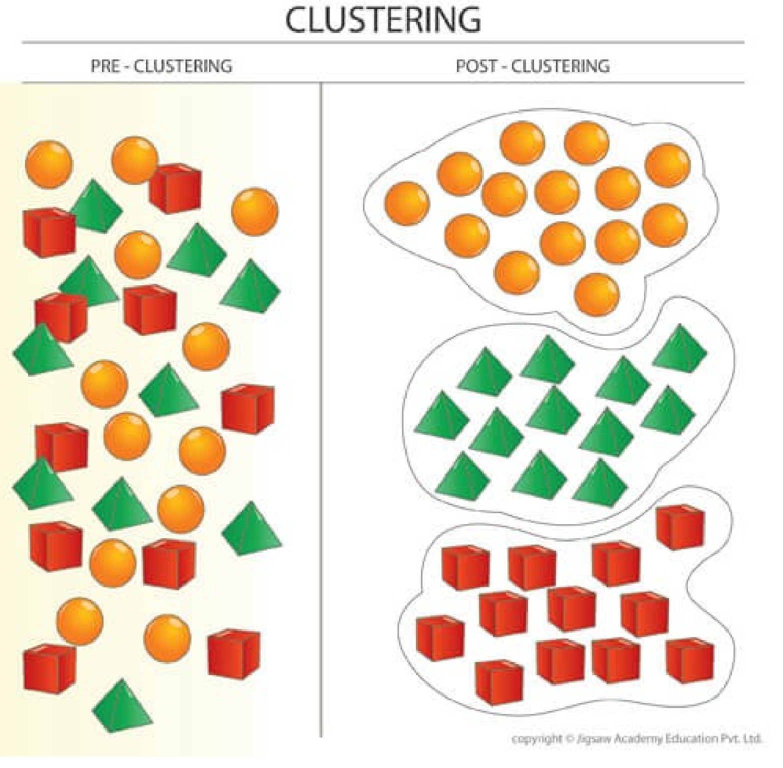
Unsupervised learning

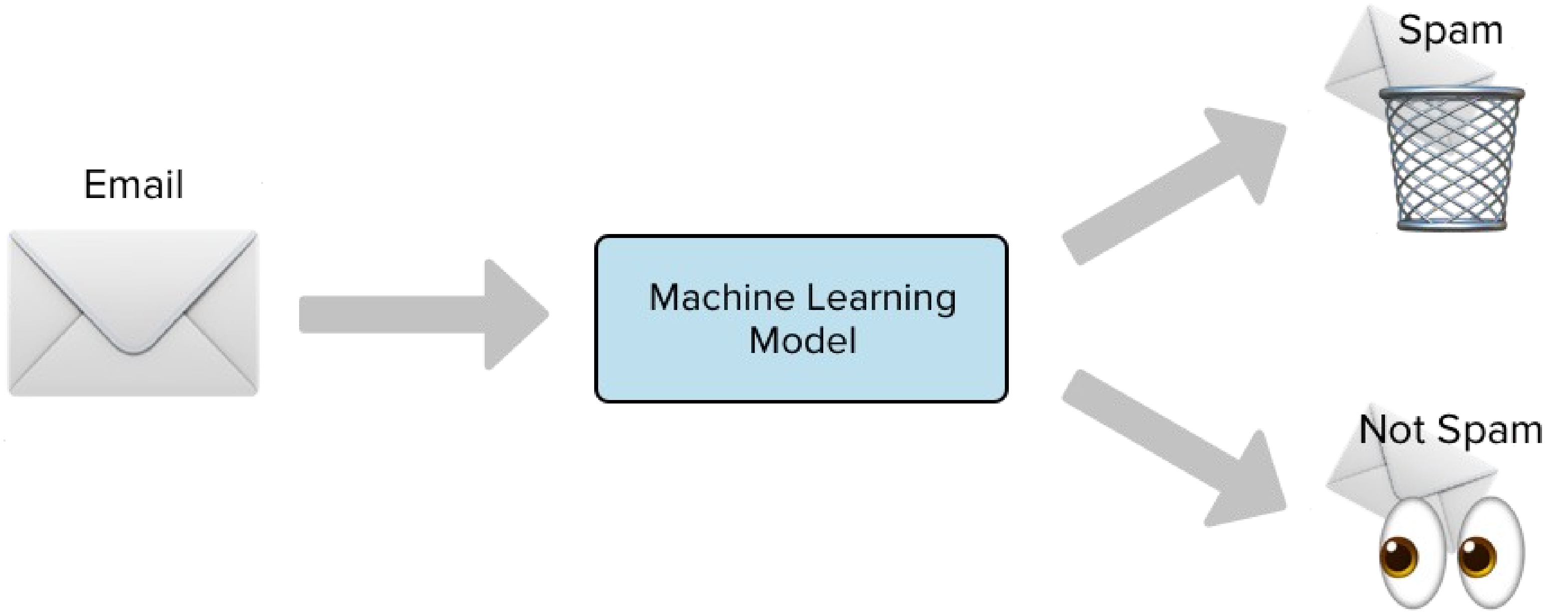


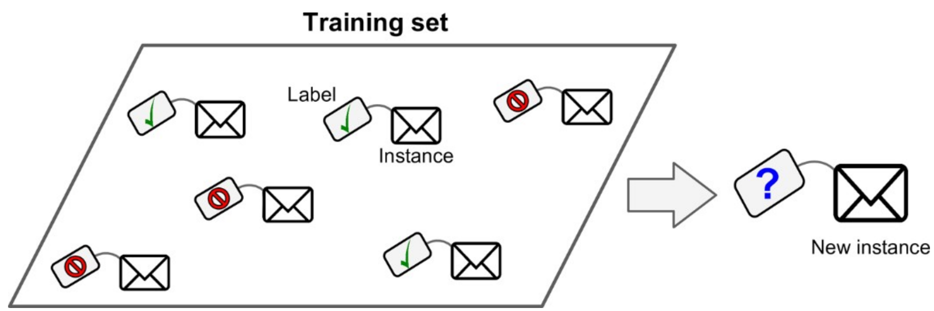
Source: https://medium.com/@gowthamy/machine-learning-supervised-learning-vs-unsupervised-learning-f1658e12a780

Clustering

- Clustering is a fairly straightforward technique to understand — objects with similar parameters are grouped together (in a cluster).
- All objects in a cluster are more similar to each other than objects in other clusters.
- Clustering is a type of unsupervised learning because the data is not labelled.
- The algorithm interprets the parameters that make up each item and then groups them accordingly.







A Case for Unsupervised-Learning-based Spam Filtering:
https://www.eecs.umich.edu/techreports/cse/2010/CSE-TR-561-10.pdf

O'Reilly Media

Anomaly detection

- In this type of the task, the computer program sifts through a set of events or objects and flags some of them as being unusual or anomaly.

- **Example**
  - **Credit-card fraud detection:** An example of an anomaly detection task is credit card fraud detection. By modeling your purchasing habits, a credit card company can detect misuse of your credit cards. If a thief steals your credit card or credit card information, the thief's purchases will often come from a different probability distribution over purchase types than your own. The credit card company can prevent the fraud by placing a hold on an account as soon as that card has been used for an uncharacteristic purchase.
  - **Intrusion detection system for the network traffic**.

Dimensionality reduction

Dimensionality reduction helps systems remove data that’s not useful for analysis.

This group of algorithms is used to remove redundant data, outliers, and other non-useful data.

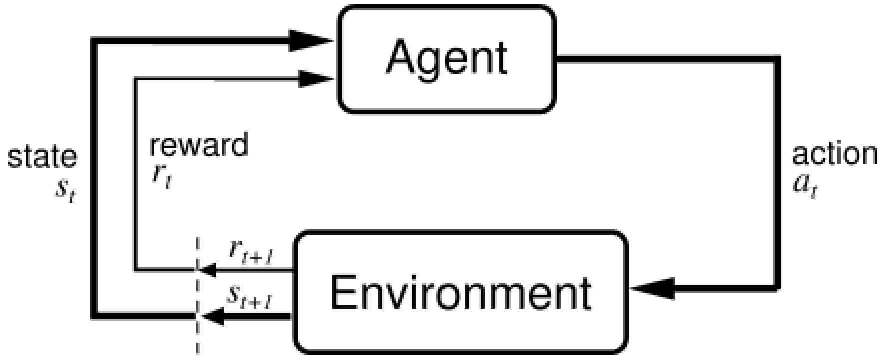
Dimensionality reduction can be helpful when analysing data from sensors and other Internet of Things (IoT) use cases

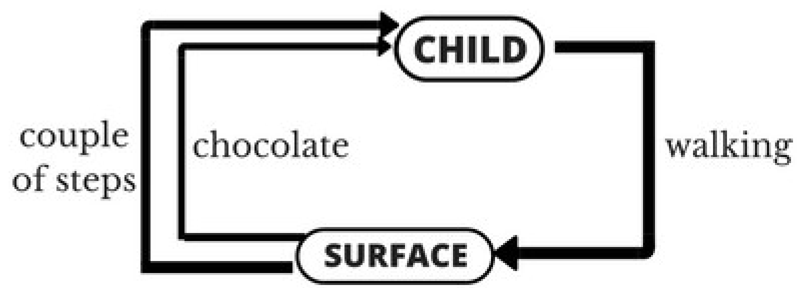
In IoT systems, there might be thousands of data points simply telling you that a sensor is turned on. Storing and analyzing that “on” data is not helpful and will occupy important storage space.

Reinforcement learning

- Reinforcement learning is a behavioural learning model
- The algorithm receives feedback from the analysis of the data so the user is guided to the best outcome
- Reinforcement learning differs from other types of supervised learning because the system isn't trained with the sample data set. Rather, the system learns through trial and error
- A sequence of successful decisions will result in the process being "reinforced" because it best solves the problem at hand
- One of the most common applications of reinforcement learning is in robotics or game playing

Reinforcement learning





Supervised vs. Reinforcement Learning

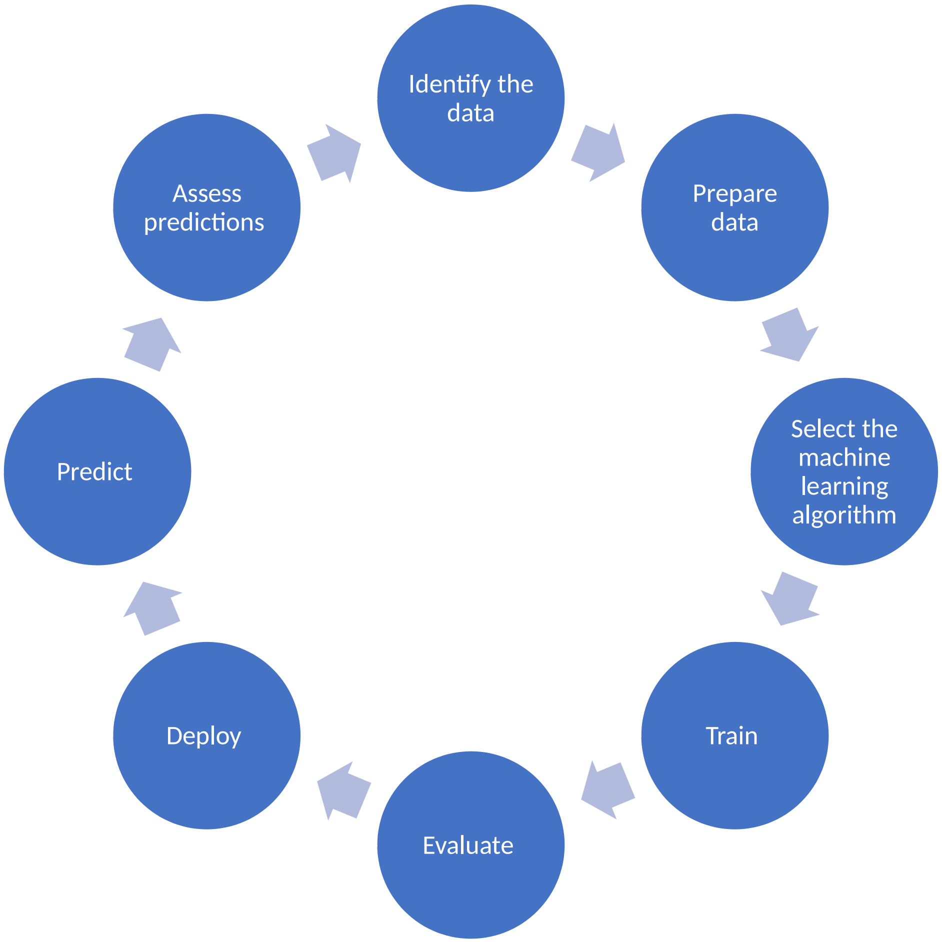
In supervised learning, there’s an external “supervisor”, which has knowledge of the environment and who shares it with the agent to complete the task.

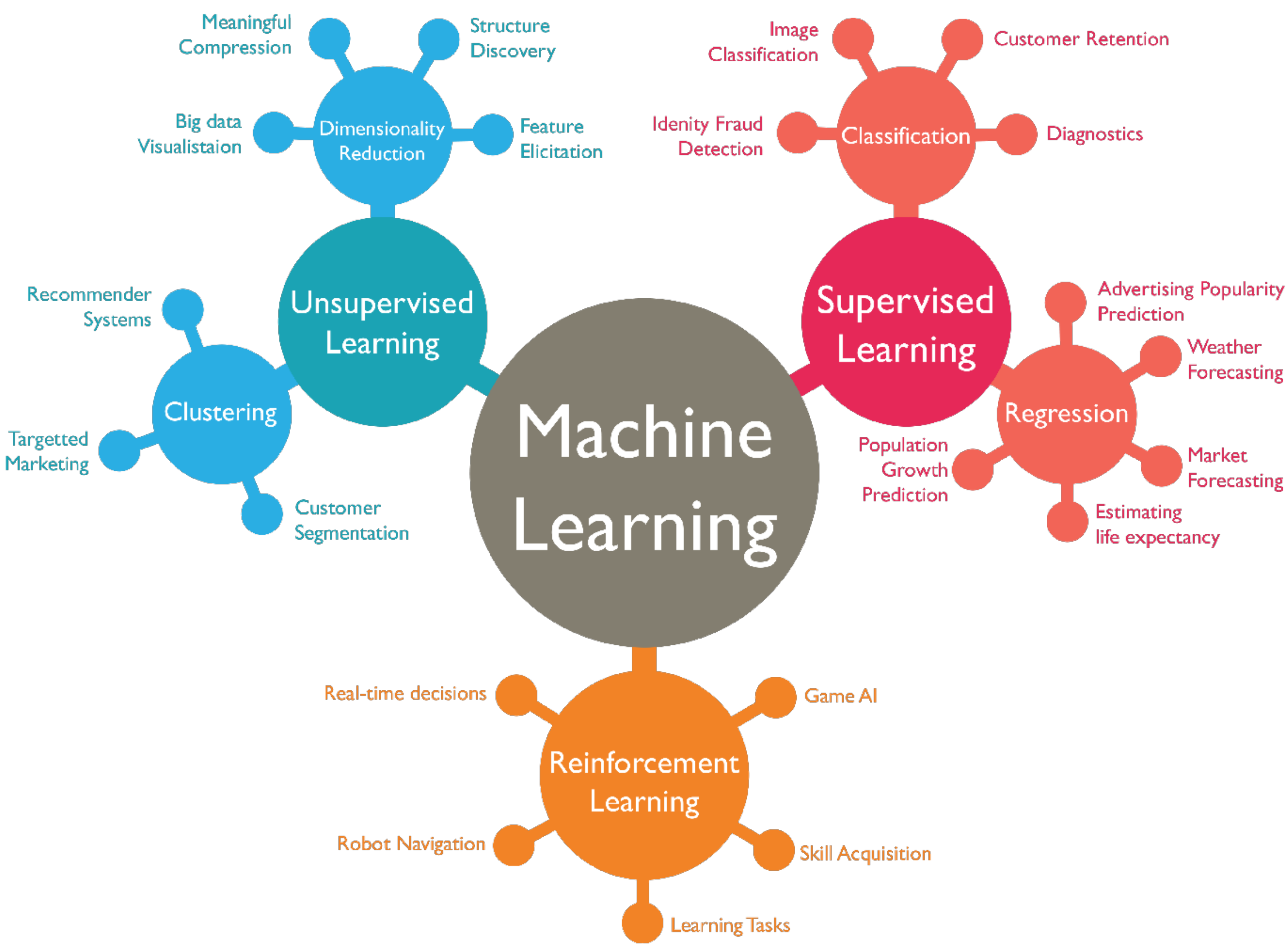
But there are some problems in which there are so many combinations of subtasks that the agent can perform to achieve the objective.

Unsupervised vs. Reinforcement Learning

- In reinforcement learning, there's a mapping from input to output which is not present in unsupervised learning. In unsupervised learning, the main task is to find the underlying patterns rather than the mapping.

The Machine Learning Cycle





https://www.slideshare.net/awahid/big-data-and-machine-learning-for-businesses

PART-III IoT and Machine Learning

IoT and Machine Learning

- 31 Billion connected device to the Internet 2018
- Filtering, sorting, querying?
- Can we use machine learning models and algorithms to generate intelligence?

IoT and ML

- Automation
  - Optimizing existing process
  - Main or improve quality
  - Reduce cost
- Enrichment
  - Add new feature or products
  - Customised manufacturing
  - New customers
- Invention
  - Create new products and categories

Machine Learning

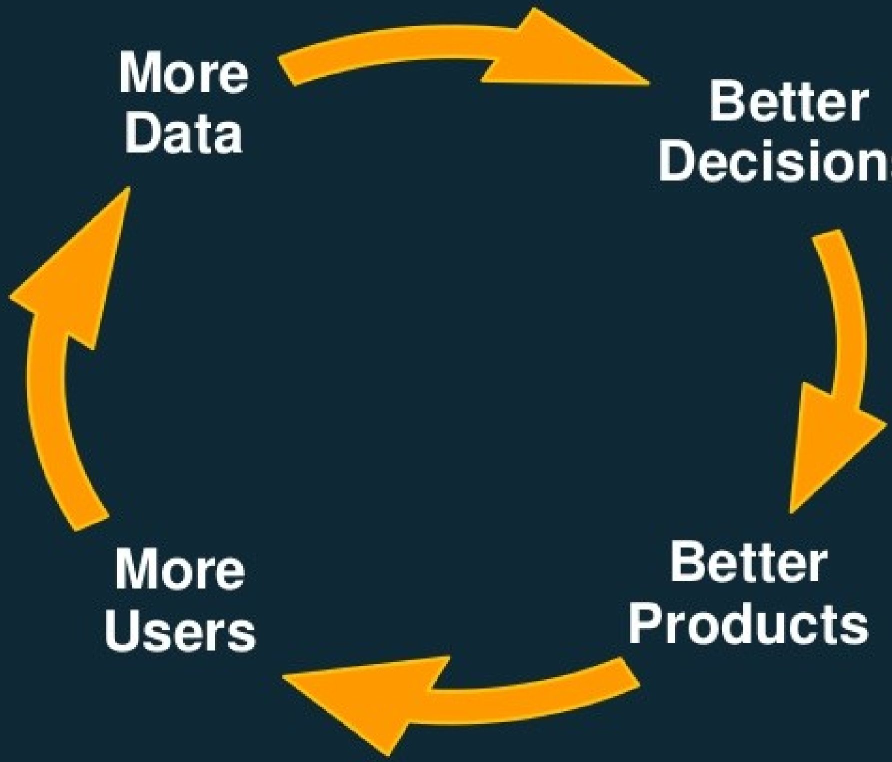
+

IoT Data

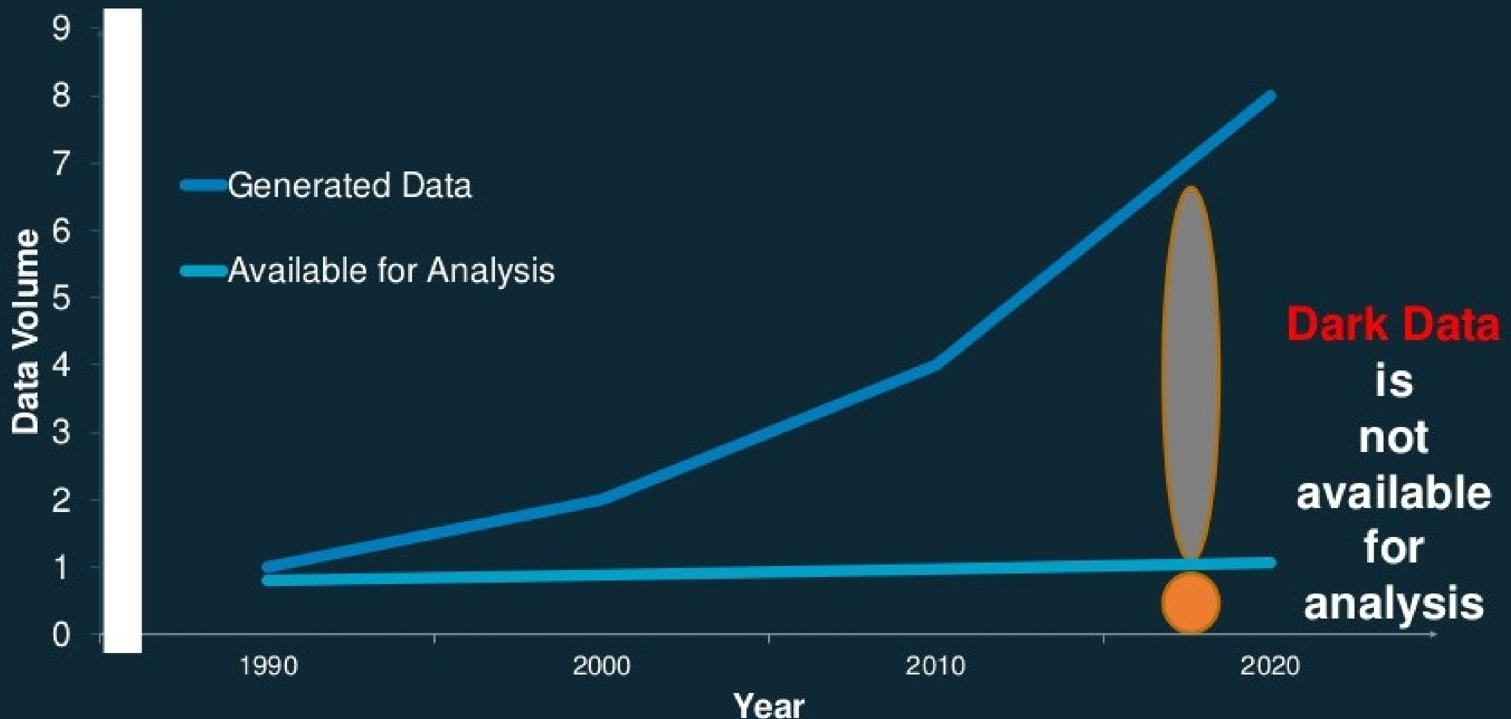
=

Smart World

A Flywheel for Data



Most generated data is unavailable for analysis



Sources:
Gartner: User Survey Analysis: Key Trends Shaping the Future of Data Center Infrastructure Through 2011
IDC: Worldwide Business Analytics Software 2012–2016 Forecast and 2011 Vendor Shares

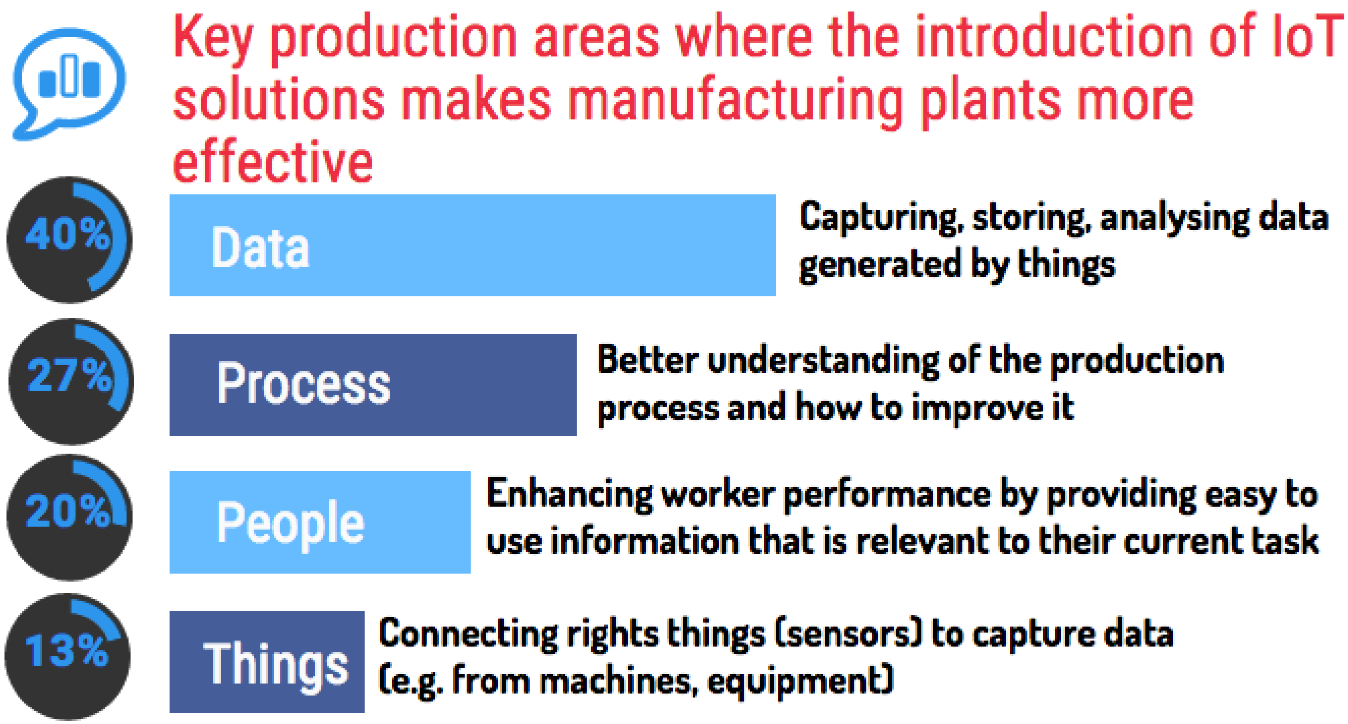
Dark data is **data** which is acquired through various computer network operations but not used in any manner to derive insights or for decision making.

AI/ML Adoption Challenges

1. While the power of ML is unrivaled, “data scientists spend around 80% of their time on preparing and managing data for analysis” … hence only 20% of their time is used to derive insights

2. While “60% of companies agree that big data will help improve their decision making and competitiveness … only 28% indicate that they are currently generating strategic value from their data”

3. The value of data science relies upon operationalizing models within business applications and processes, yet “50% of the predictive models [built] don’t get implemented”



CISCO, 2016, http://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/iot-data-analytics-white-paper.PDF (2016)

Predictive Machine Maintenance and Prevention of Stoppages in Manufacturing

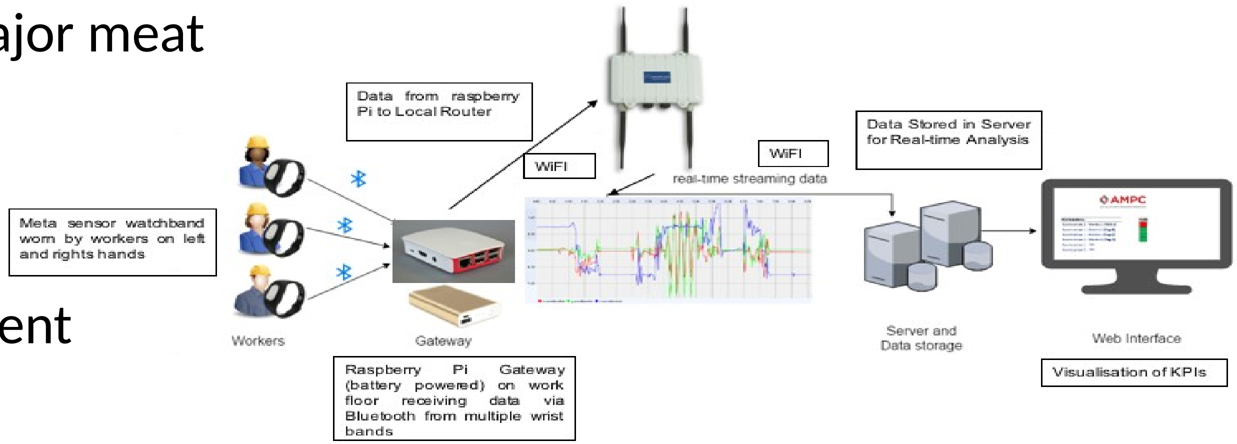
- Developing a predictive machine health monitoring system that will improve plant productivity through the reduction of unplanned downtime
- Approach involve combine manufacturing process and machine data modelling and analysis

Machine Selection and FMEA → Detailed Process Capture → Data Model Development (Process and Machine level) / Analytical Validation of Machine Selection → Data Capture → Development of Predictive Tool → Evaluation, Gap Analysis, and Iteration

- Expected Outcome: Improve plant machine efficiency (> 85%) – currently 65%

Measuring and Improving Worker Productivity

- Conducted trials in a major meat processing plant
- IoT solution payback in 0.4 years
- IoT solution will lead to significant improvement in plant productivity



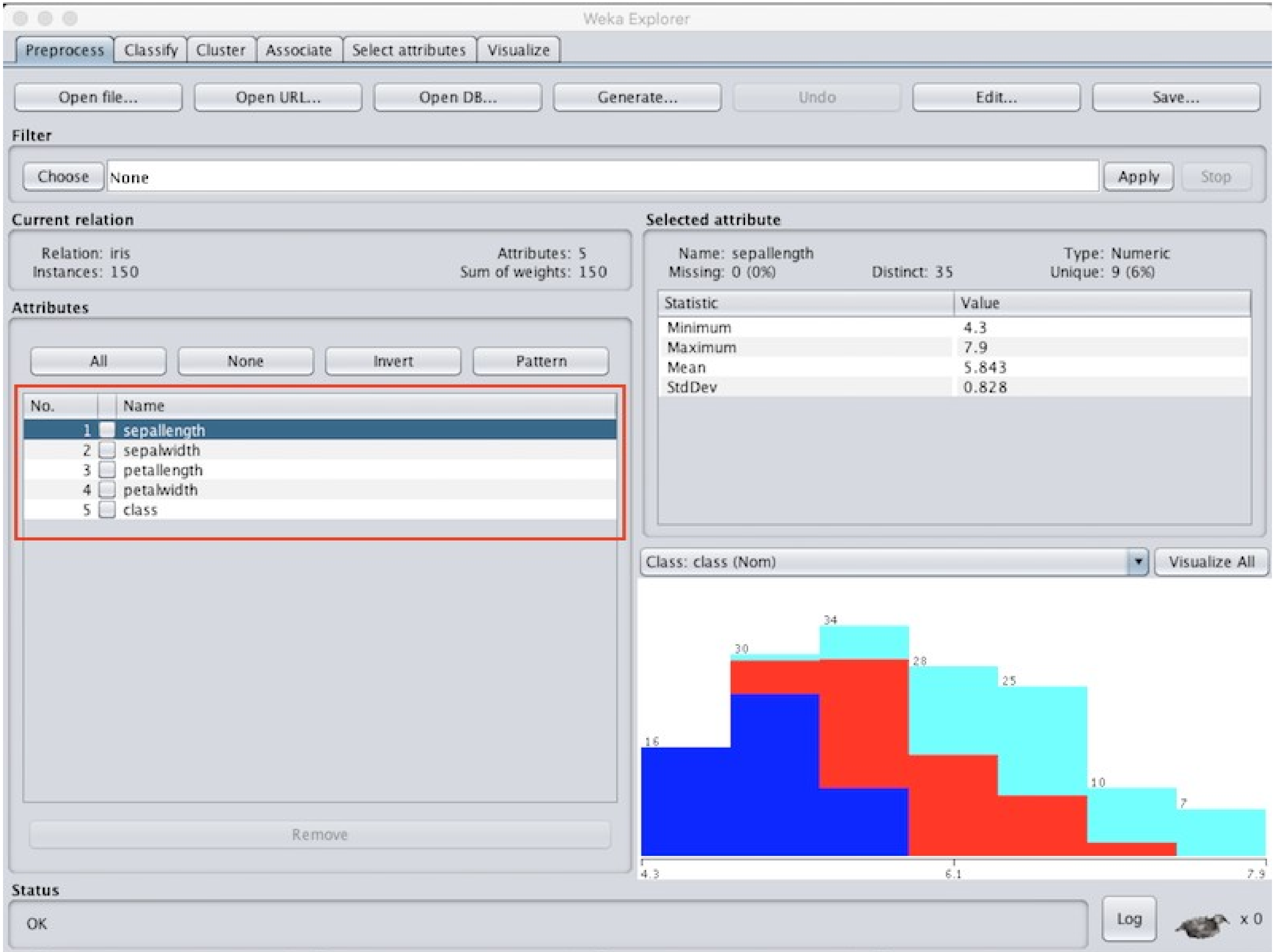
	Worker 1 (Experienced)	Worker 2 (Inexperienced)
Idle Time	36%	8%
Productive Time	63%	91%
Alignment Time	1%	1%
Active States	49	30

WEKA

- Weka is a collection of machine learning algorithms for data mining tasks
- It contains tools for data preparation, classification, regression, clustering, association rules mining, and visualization
- Found only on the islands of New Zealand, the Weka is a flightless bird with an inquisitive nature.

Weka requires Java.

http://www.cs.waikato.ac.nz/ml/weka/downloading.html



https://code.likeagirl.io/beginning-to-weka-step-by-step-93f6564d9f2

Kaggle: Machine learning and data science community

- Over 19,000 public datasets.
- 200,000 public notebooks.

https://www.kaggle.com/datasets

You may learn from COVID-19 dataset uploaded quite recently.

Questions?