A Gentle Introduction to Artificial Intelligence and Machine Learning

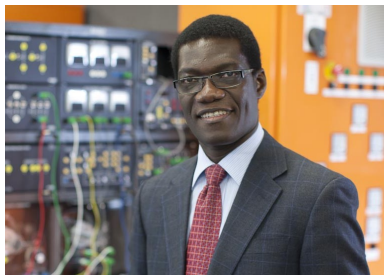
Centre for AI – Faculty of Engineering and Information Sciences

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Content

1. Centre for Artificial Intelligence
2. Introduction
   - Artificial Intelligence
3. Machine Learning
   - AI enabled by learning
   - Data
4. Learning Paradigms
   - Supervised learning
   - Unsupervised learning
   - Semi-supervised learning
   - Transfer learning
   - Reinforcement learning
   - Generative learning
   - < 15 minute break >
5. Artificial neural network
6. Some Applications
7. Grand Challenges
8. Q & A

Presenters



Professor Philip O. Ogunbona, School of Computing and Information Technology

PhD in Electrical engineering, Imperial College, London

Research and development: Image & Signal Processing; Machine Learning; Computer Vision;

Associate Professor Markus Hagenbuchner, School of Computing and Information Technology

PhD in Computer Science in 2002

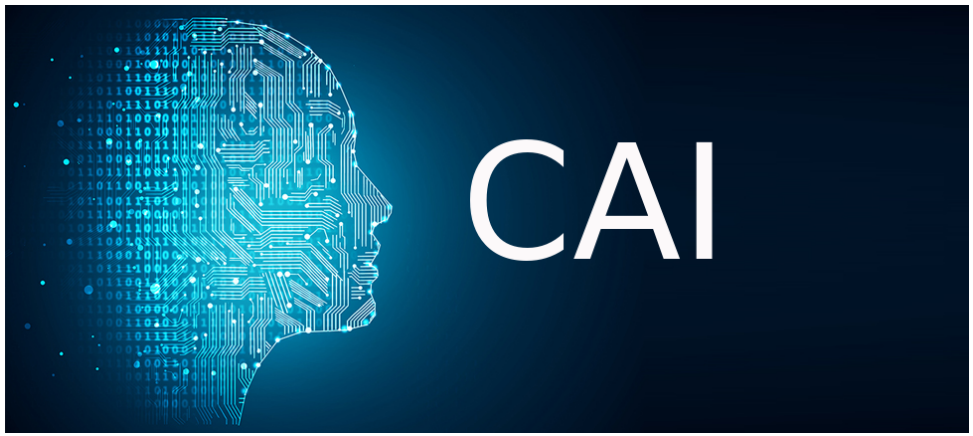
Working in AI for over 25 years: Development of machine learning algorithms and neural networks, with applications to Data Mining and Big Data.

Associate Professor Wanqing Li, School of Computing and Information Technology

PhD in Computer Science; Uni Western Australia

Research & Development of machine learning algorithms and neural networks, with applications to Computer Vision and Image Processing.

Centre for AI - Who are we?



We focus on pioneering research to develop innovative theories and techniques of AI, and transferring the knowledge and technologies to industry, community and society. CAI offers expertise in:

- Machine Learning
- Computer Vision
- Multi-agent Systems
- Big data and Applications
- Multimedia Signal Analysis
- Probabilistic Model Checking
- Natural Language Processing
- Distributed Computing

and their applications to a wide range of real-world problems.

https://uow.info/cai

Introduction

What is Artificial Intelligence (AI)?

Artificial intelligence is a constellation of many different technologies working together to enable machines to sense, comprehend, act, and learn with human-like levels of intelligence. AI isn't just one thing. (`https://accenture.com`)

Introduction

What is Artificial Intelligence (AI)?

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information or rules for using the information), reasoning (using the rules to reach approximate or definite conclusions), and self-correction. (`https://exploreai.org/p/ai-definition`)

What is Artificial Intelligence (AI)?

Human intelligence? - Take One!

Human intelligence is a mental quality that consists of the abilities to learn from experience, adapt to new situations, understand and handle abstract concepts, and use knowledge to manipulate one's environment (`https://www.britannica.com`).

Human intelligence? - Take two!

Through intelligence, humans possess the cognitive abilities to learn, form concepts, understand, apply logic and reason, including the capacities to recognize patterns, plan, innovate, solve problems, make decisions, retain information, and to communicate information (`https://www.wikipedia.com`).

A bit of history

Darthmouth meeting - 1956

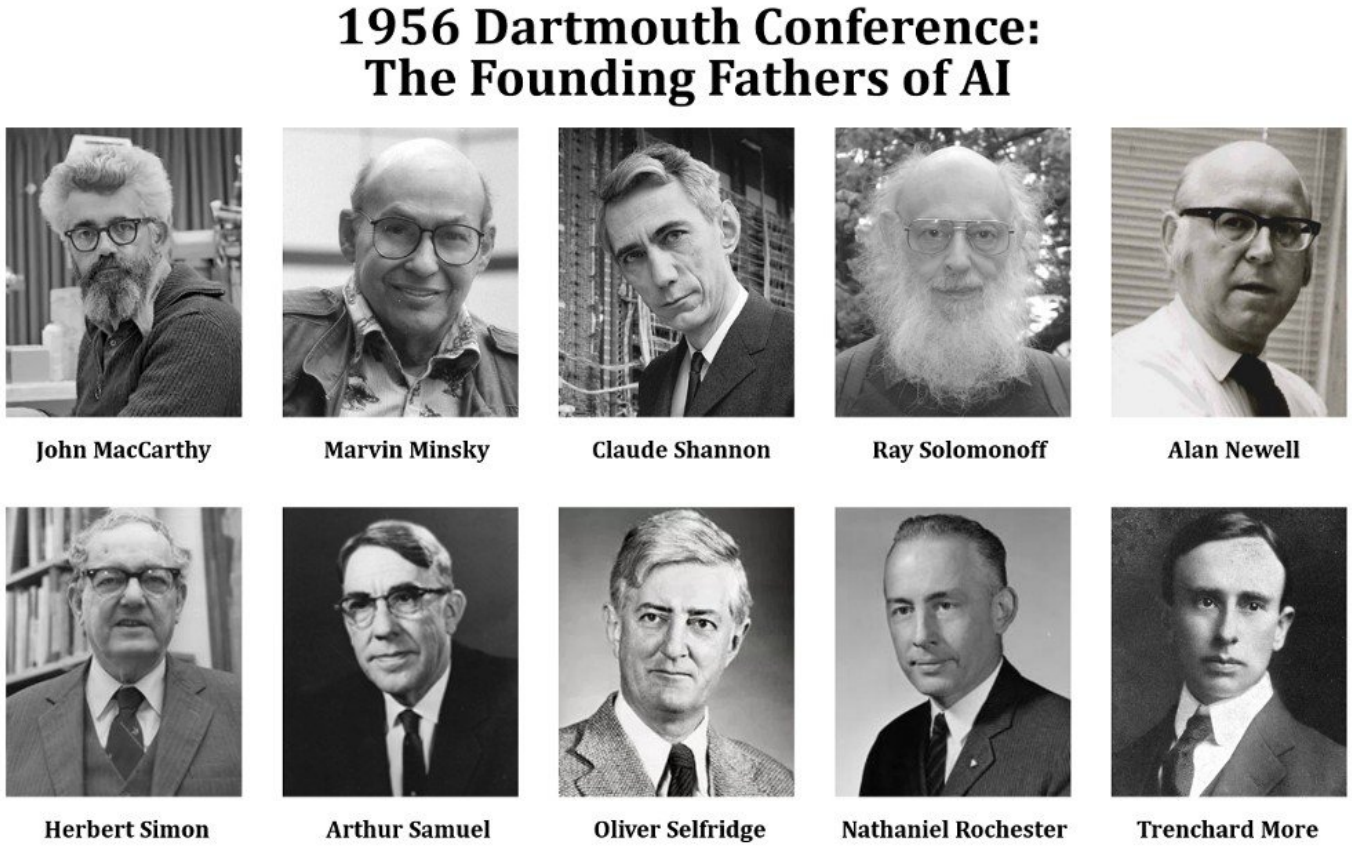


Figure: Founding Fathers of Artificial Intelligence

A bit of history

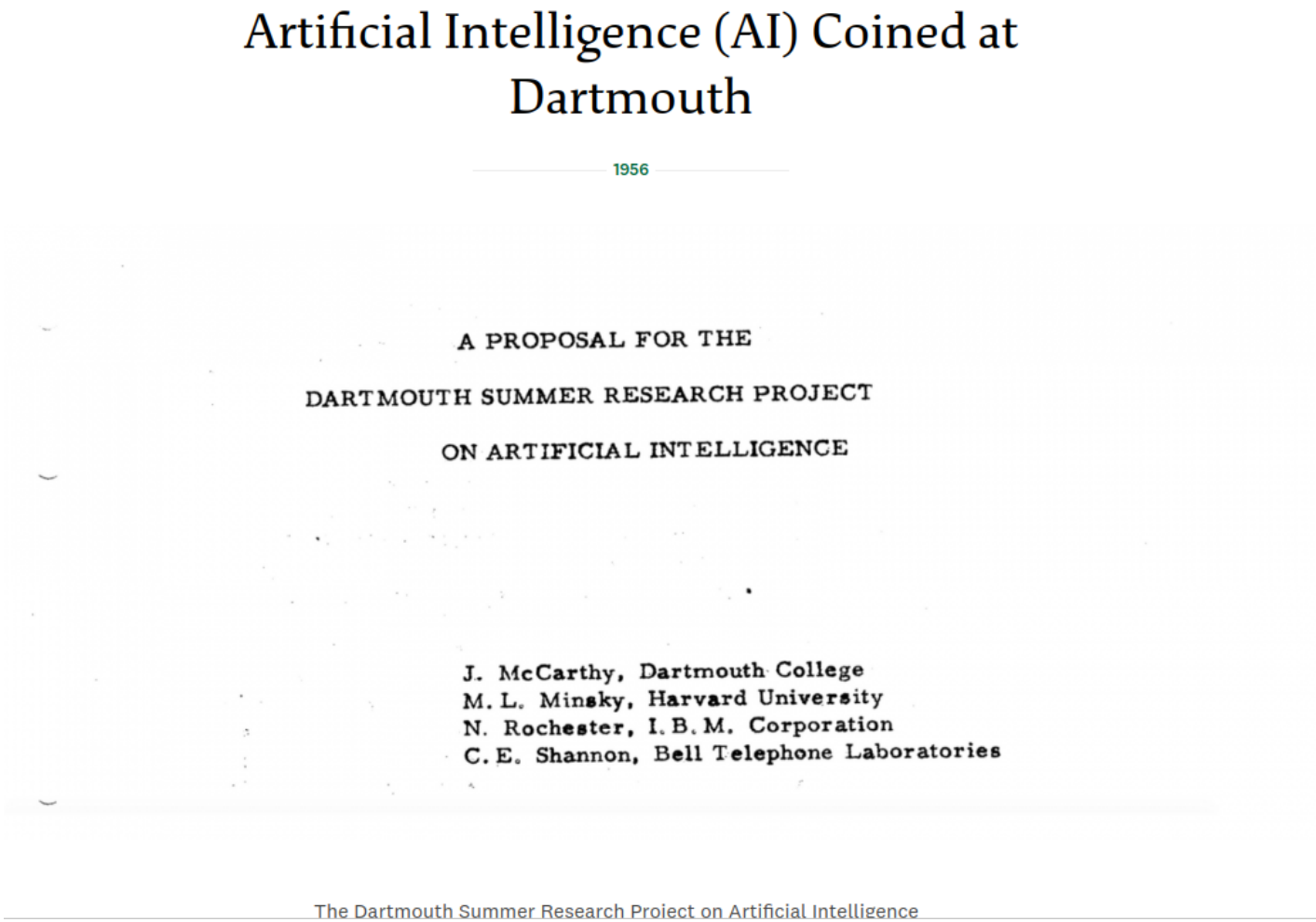


Figure: Paper proposing the AI project

Strong vs Weak AI

Strong AI

Strong artificial intelligence (AI), also known as artificial general intelligence (AGI) or general AI, is a theoretical form of AI used to describe a certain mindset of AI development.

Strong AI aims to create intelligent machines that are indistinguishable from the human mind.
(`https://www.ibm.com/cloud/learn/strong-ai`)

Weak AI

Weak AI, also known as narrow AI, focuses on performing a specific task, such as answering questions based on user input or playing chess.

It can perform one type of task, but not both, whereas Strong AI can perform a variety of functions, eventually teaching itself to solve for new problems.

Weak AI relies on human interference to define the parameters of its learning algorithms and to provide the relevant training data to ensure accuracy. (`https://www.ibm.com/cloud/learn/strong-ai`)

Strong vs Weak AI

Strong AI

Strong artificial intelligence (AI), also known as artificial general intelligence (AGI) or general AI, is a theoretical form of AI used to describe a certain mindset of AI development.

Strong AI aims to create intelligent machines that are indistinguishable from the human mind.
(https://www.ibm.com/cloud/learn/strong-ai)

Weak AI

Weak AI, also known as narrow AI, focuses on performing a specific task, such as answering questions based on user input or playing chess.

It can perform one type of task, but not both, whereas Strong AI can perform a variety of functions, eventually teaching itself to solve for new problems.

Weak AI relies on human interference to define the parameters of its learning algorithms and to provide the relevant training data to ensure accuracy. (https://www.ibm.com/cloud/learn/strong-ai)

Conceptualisation of contemporary AI

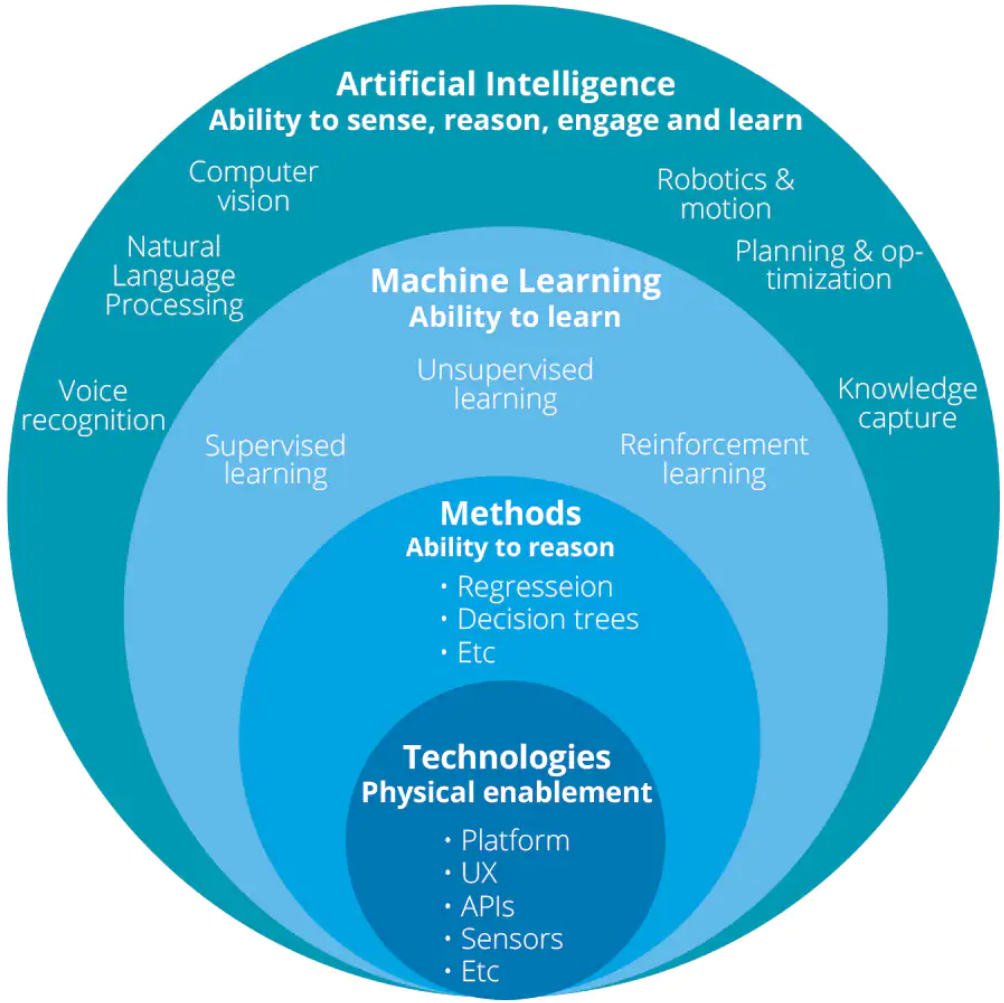


Figure: A view of AI and its components (`https://www2.deloitte.com/`)

Conceptualisation of contemporary AI

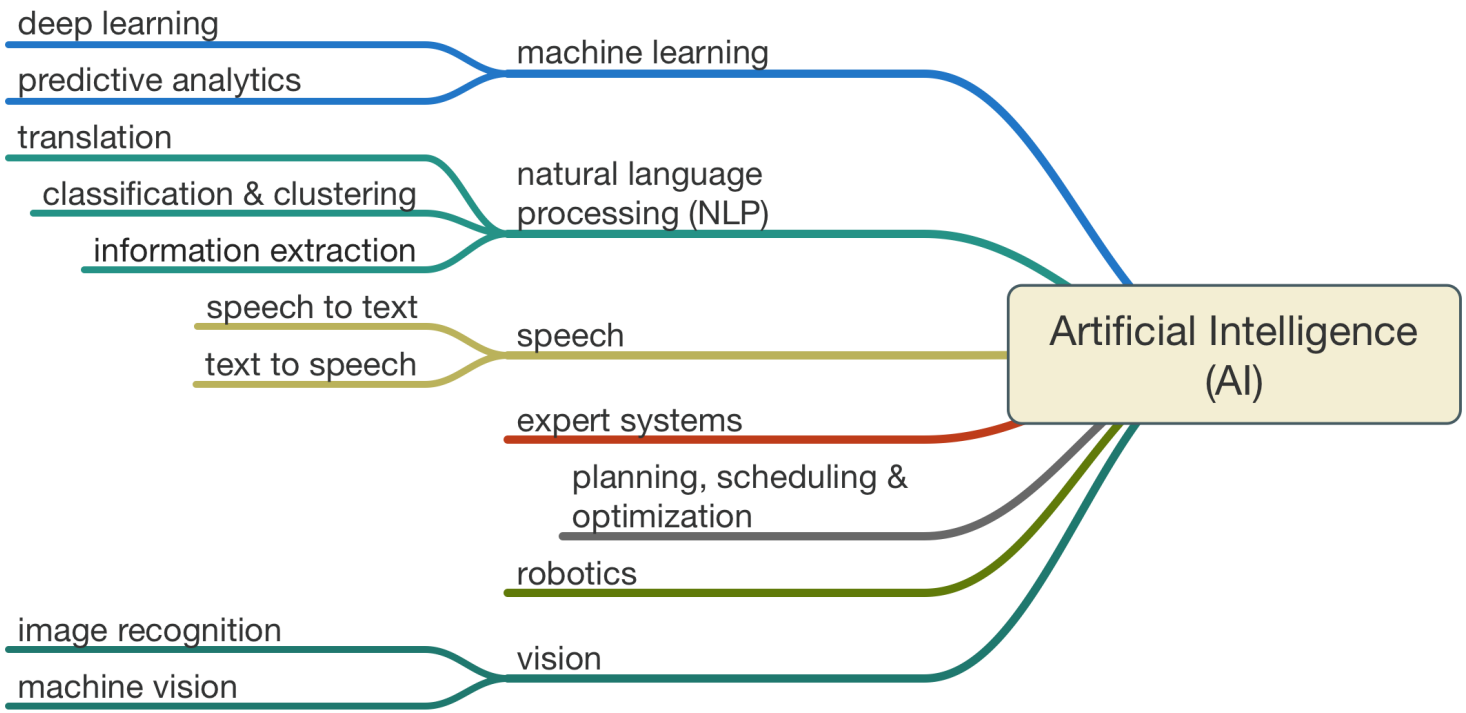


Figure: A view of technologies comprising AI (`https://wiredelta.com/`)

AI enabled by learning

AI and Machine Learning

One of the key components of Artificial Intelligence is learning.

Machine learning allows the concept of learning from data.

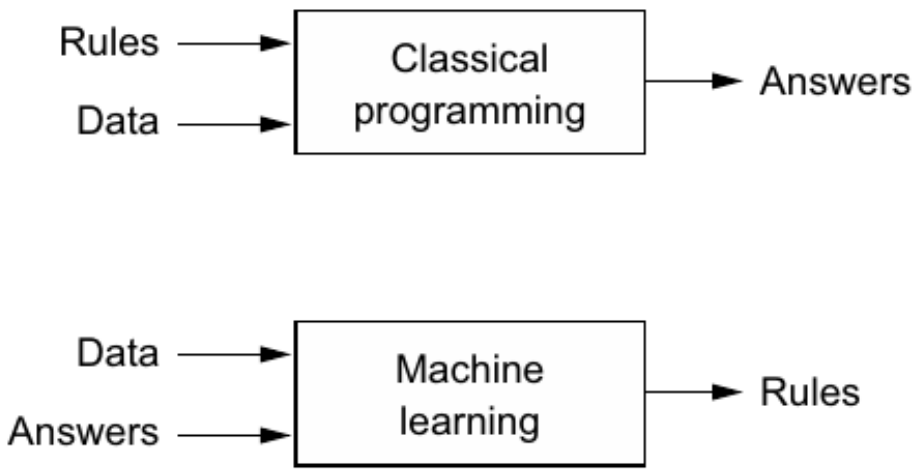


Figure: Machine learning: a new programming paradigm

Life cycle of AI

Six phases

1. Discovery, Data collection and annotation/labeling
2. Data preparation (preprocessing)
3. Model planning (methods, techniques, workflow, variables, relationships, models)
4. Model building (training and test datasets, software and hardware, analysis of results)
5. Communication of results (identify key findings)
6. Operationalize (delivery, pilot project)

Data [is king]

In general data drives model building and needs to be treated carefully.

Data quality issues:

Noisy data

Data that contains a large amount of conflicting or misleading information.

Data [is king]

In general data drives model building and needs to be treated carefully.

Data quality issues:

Dirty data

Data that contains missing values, categorical and character features with many levels, and inconsistent and erroneous values

Data [is king]

In general data drives model building and needs to be treated carefully.

Data quality issues:

Sparse data

Data that contains very few actual values, and is instead composed of mostly zeros or missing values

Inadequate data

Data that is either incomplete or biased

Data [is king]

Feature Engineering

1. Feature selection : selecting the most useful features to train on among existing features
2. Feature extraction: combining existing features to produce a more useful one
3. Creating new features by gathering new data

Supervised learning

What is it?

- In this learning scenario, there are examples of appropriate output data for specific input data.
- The problem is how to learn the best model of the relationship between the input and output from the example data.

Supervised learning

How to do it?

- Collect an appropriate amount of data.
- Separate data into three distinct subsets.
  - training, test, and validation dataset.
- There is an implied model whose parameters can be learned.
- Train the model using an appropriate methodology.
- There is a method of judging goodness of the model.
  - Optimize model by using the validation data.
- Test the model on the test data (not part of training set).

Supervised learning

Simple general science experiment - linear regression

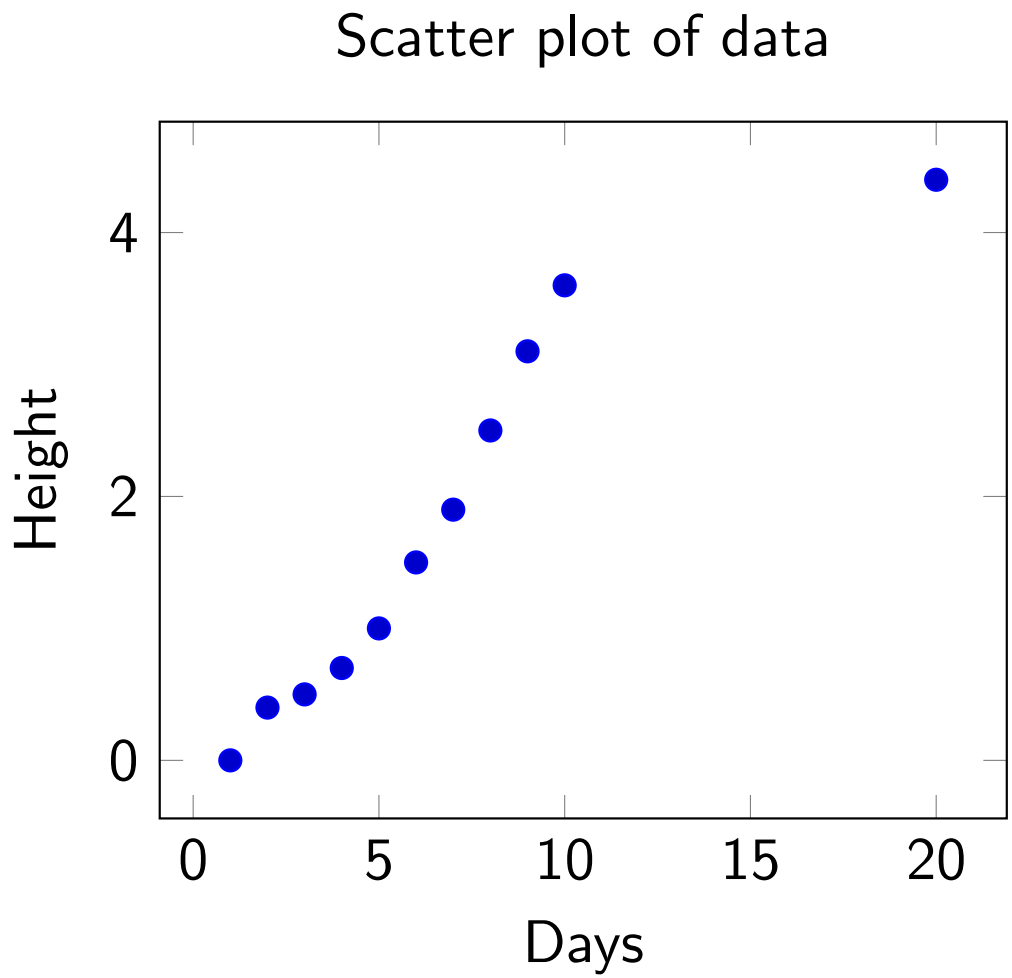
- A science teacher wanted to explore the relationship between time and the height a bean can grown from time of planting.
- Students brought clear jam jars and were asked to place bean seeds in wet cotton wool set in the jar.
- Jars were placed on the window sill of the class and left to grow over several weeks; plant was watered appropriately.
- Data collected; measure height of plant over time.

(a) Day 1 (b) Day 3 (c) Day 7 (d) Day 12 (e) Day 20

Figure: Bean seed sprouting over days

Data Collected

Day	Height (cm)
1	0
2	0.4
3	0.5
4	0.7
5	1.0
6	1.5
7	1.9
8	2.5
9	3.1
10	3.6
..	..
..	..
20	4.4



When we draw the line of best fit to the data, we are building a model that explains the height of beans sprouting over time. The line can be used to predict height at future times. This is regression problem.

Supervised learning

Distinguish between Salmon and Bass fishes - classification

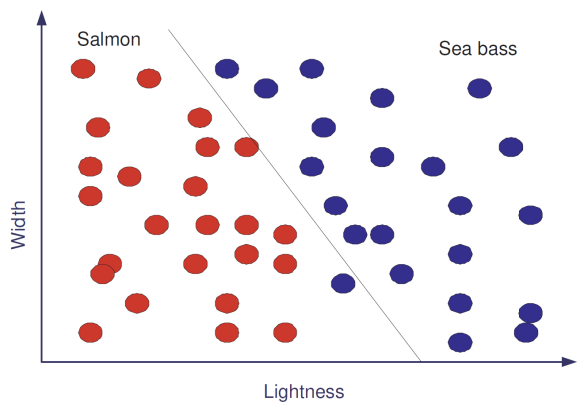
- A commercial fishing company goes to sea to catch fish.
- They catch both salmon and bass.
- During packaging they need to distinguish salmon and bass, otherwise they lose money.
- Design a AI system that will automatically sort and package the fish with minimum error.

Salmon

Bass

Data collection (using appropriate sensors)

- Lightness
- Width



Unsupervised learning

What is it?

- In this learning scenario, there are example data, but they are not identified with a specific label beforehand.
- The problem is how to learn the underlying model that can best separate the data into meaningful clusters of categories. We want to uncover the pattern inherent in the data.
- We may not know how many clusters of categories are appropriate or makes sense.

Unsupervised learning

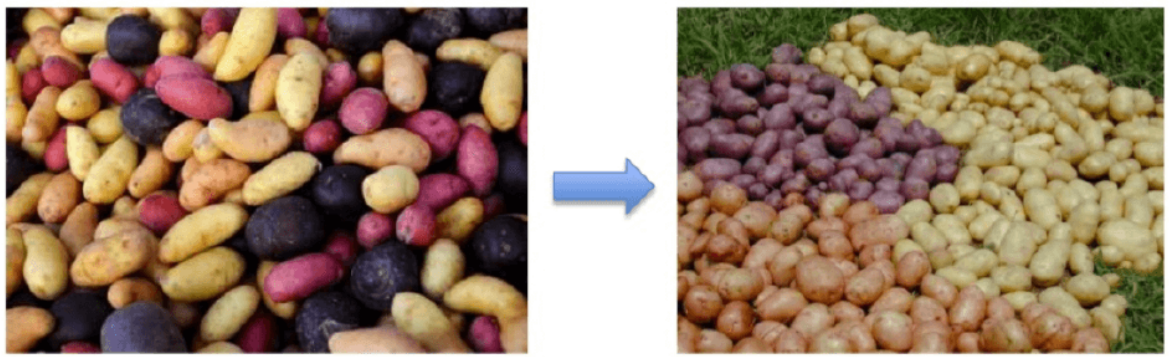


Figure: The samples are unlabeled and the model is trained to identify underlying pattern. Different properties (features) of the data are used. The neural network may discover the appropriate features.

Unsupervised learning

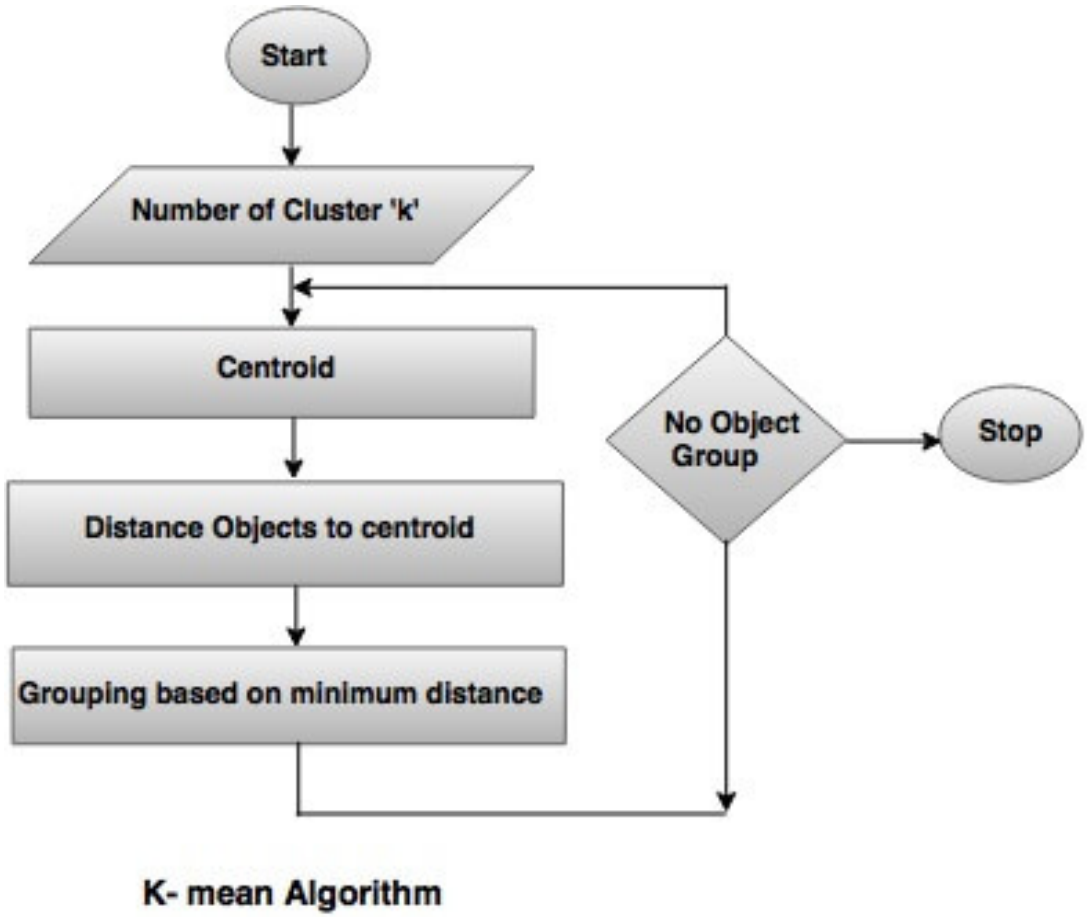


Figure: A popular method of achieving data clustering employs the k-means algorithm.

Unsupervised learning

Market basket analysis

- Market Basket Analysis is one of the key techniques used by large retailers to uncover associations between items.
- It works by looking for combinations of items that occur together frequently in transactions.
- In general, the machine learning tool used is called association rules mining and it is an unsupervised learning method.

Unsupervised learning

Rule	Support	Confidence	Lift
$A \Rightarrow D$	2/5	2/3	10/9
$C \Rightarrow A$	2/5	2/4	5/6
$A \Rightarrow C$	2/5	2/3	5/6
$B \& C \Rightarrow D$	1/5	1/3	5/9

Figure: Market basket (Credits - www.towardscience.com)

Association rules mining

1. Analyse data for patterns or co-occurrence
2. Rule has two parts: an antecedent (if) and a consequent (then)
3. Search data for frequent if-then patterns
4. Use the criteria support and confidence to identify the most important relationships.
5. Support is an indication of how frequently the items appear in the data.
6. Confidence indicates the number of times the if-then statements are found true.
7. A third metric lift, can be used to compare confidence with expected confidence

Semi-supervised learning

What is it?

- In this learning scenario, there are a few labeled example data and many unlabelled data.
- labeling data is expensive; there is abundance of unlabelled data.
- The problem is how to learn the underlying model that can best classify the whole data into given categories despite lack of many labeled data.

Semi-supervised learning

How to do it?

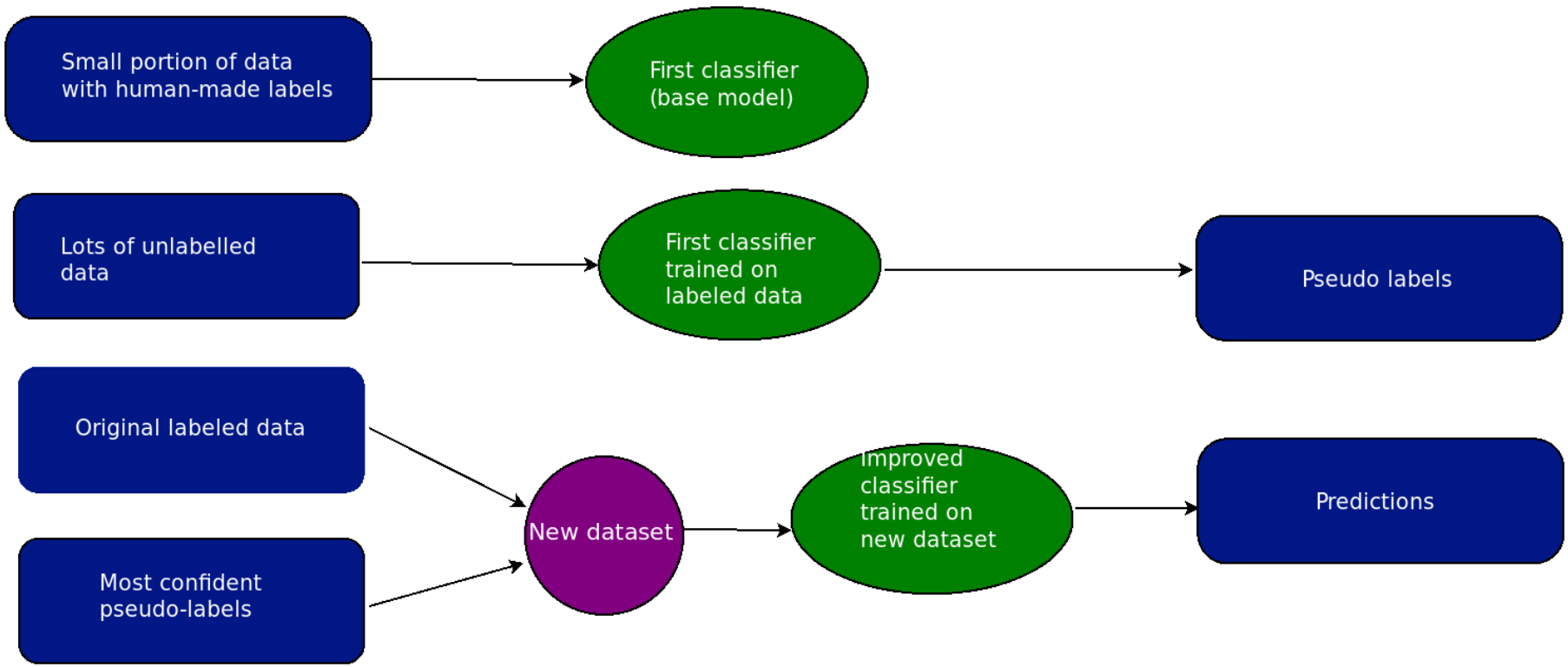


Figure: Semi-supervised self training (Credits- `altexsoft.com`)

Transfer learning

What is it?

- In this learning scenario, the focus is on storing knowledge gained while solving one problem and applying it to a different but related problem.
- Transfer learning method uses the similarity of data, tasks, models in different domains and applies the knowledge from one domain to another domain.

Transfer learning

How to do it?

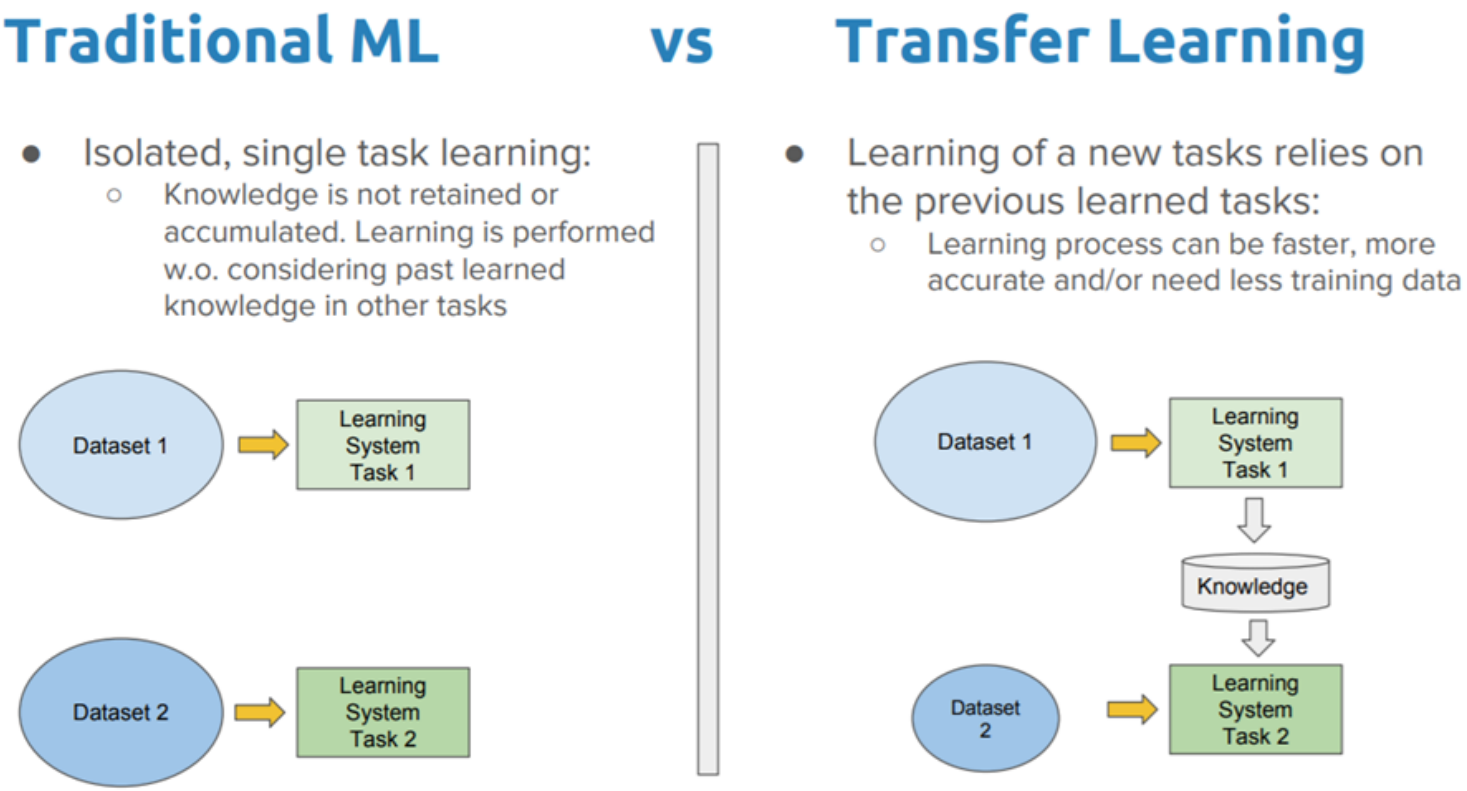


Figure: Transfer learning compared to conventional learning (Credits - `https://towardsdatascience.com/`)

Reinforcement learning

What is it?

- In this scenario the problem is modeled as the interaction between an agent and an environment.
- The idea is to come up with a set of actions that maximise a cumulative reward.
- An agent is connected to its environment via perception and action
- On each step of interaction the agent receives as input, some indication of the current state of the environment
- The agent then chooses an action as its output
- The action changes the state of the environment, and the value of this state transition is communicated to the agent as some reinforcement signal
- The output in this learning paradigm is the action performed by the agent (i.e. the AI)

Reinforcement learning

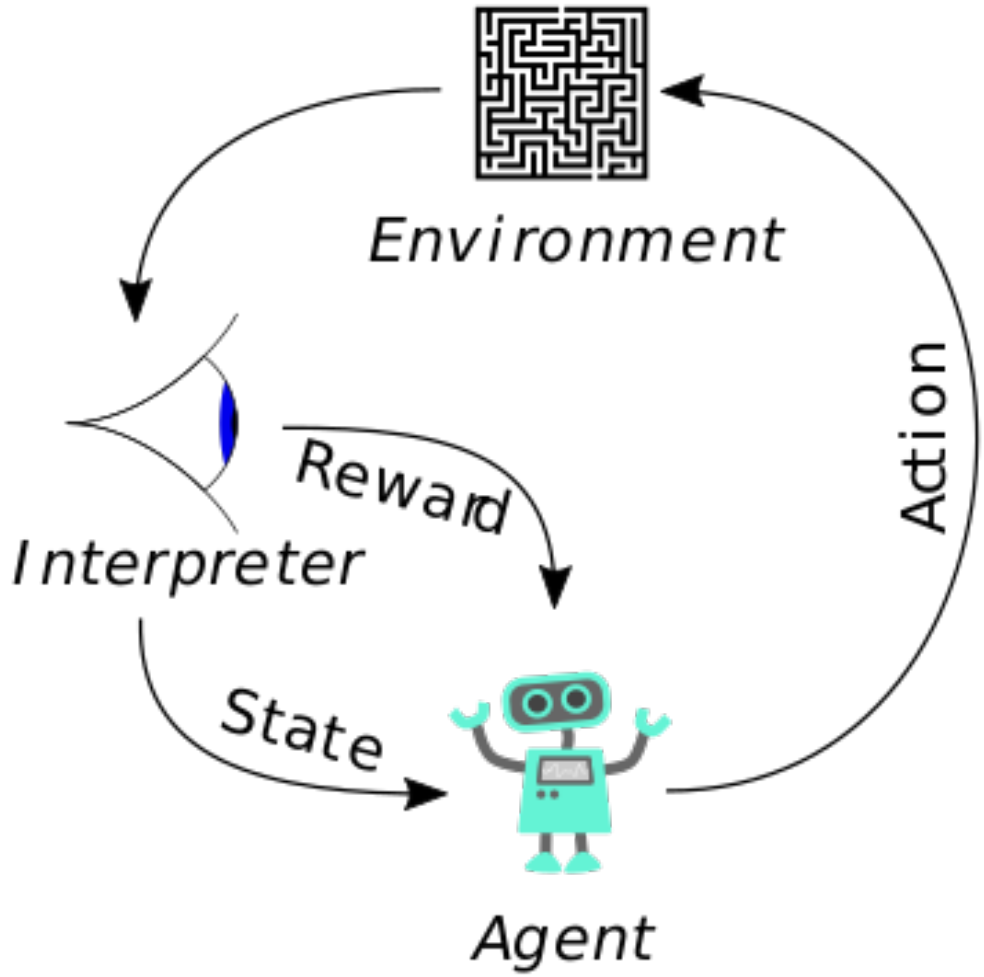


Figure: Reinforcement learning (Credits - Wikipedia)

Reinforcement learning

Conceptual dialogue between the agent and the environment[1]

Environment:	You are in state 65. You have 4 possible actions
Agent:	I'll take action 2.
Environment:	You received a reinforcement of 7 units. You are now in state 15. You have 2 possible actions.
Agent:	I'll take action 1.
Environment:	You received a reinforcement of -4 units. You are now in state 65. You have 4 possible actions.
Agent:	Agent: I'll take action 2.
Environment:	You received a reinforcement of 5 units. You are now in state 44. You have 5 possible actions.
Agent:	...
Environment:	...

[1] Journal of Artificial Intelligence Research 4 (1996) 237-285

Generative learning

What is it?

- A generative model describes how a dataset is generated, in terms of a probabilistic model.
- By sampling from this model, we are able to generate new data.

Generative learning

Scenario

Suppose we have a dataset containing images of horses. We may wish to build a model that can generate a new image of a horse that has never existed but still looks real because the model has learned the general rules that govern the appearance of a horse. This is the kind of problem that can be solved using generative modeling.

Generative learning

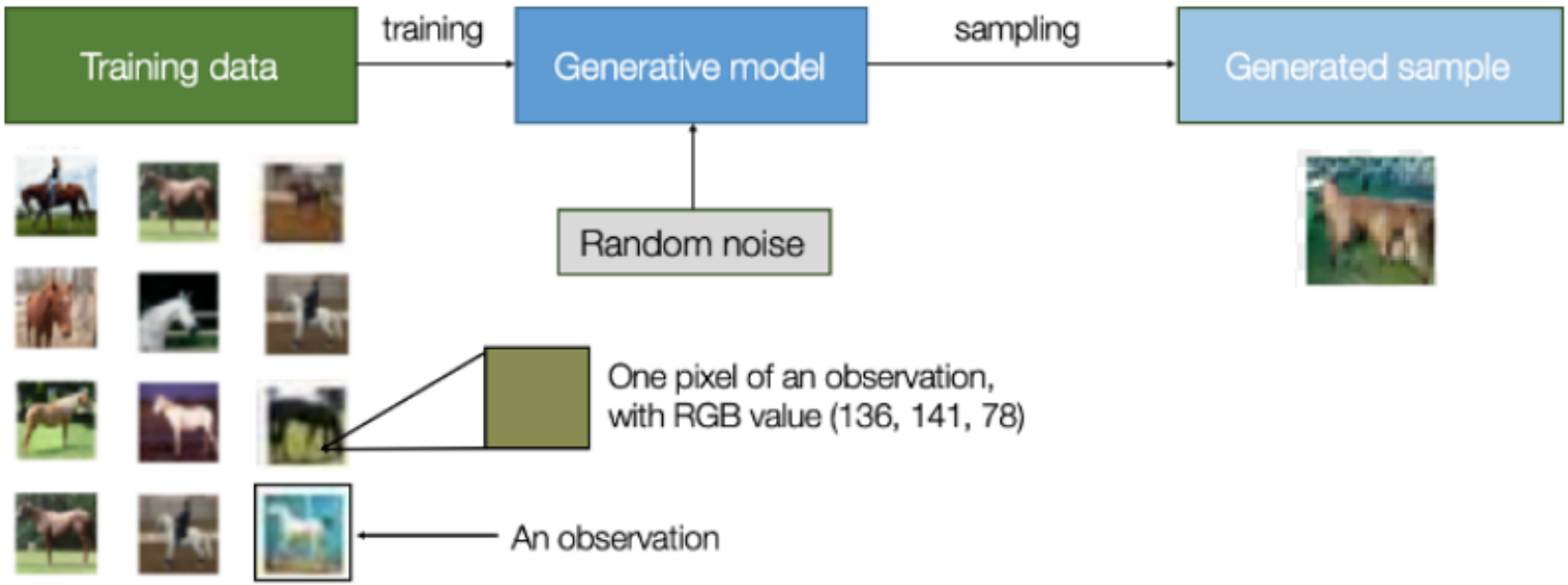


Figure: Generative modeling that can create new samples from examples.

Artificial neural network

Artificial neural networks

- Artificial Neural Networks are by far the most successful ML technique.
- Inspired the way a biological brain works.
- Brain "learns" by adjusting the connecting links (Axons and synaptic weights)

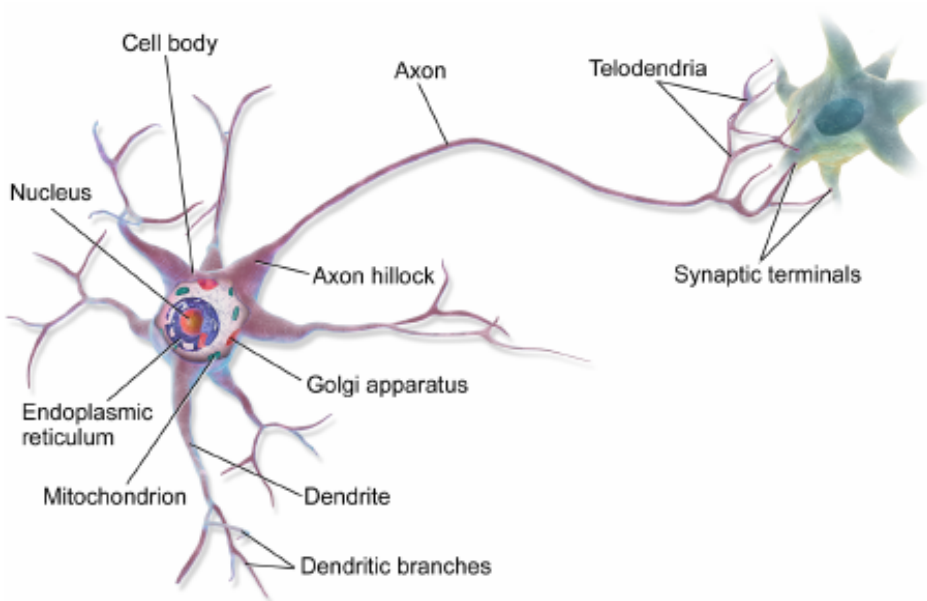


Figure: Biological neuron of mammalian brain (Géron, 2019).

Artificial neural network

- For simplicity, neurons are organized in one or more layers - Input, hidden, and output layers
- Layers are fully connected by weighted links
- Input signal travels from input to output neurons

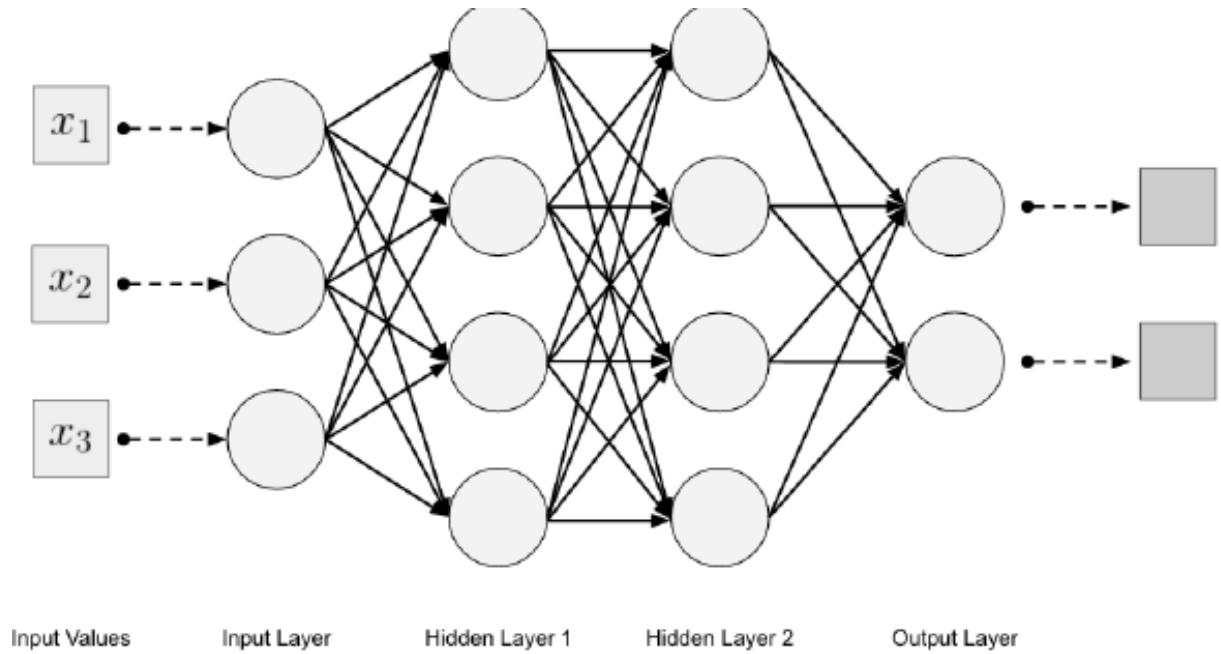


Figure: Artificial neural network (Géron, 2019)

Deep Learning

- All Deep Learning methods are Neural Networks
- Networks with many layers are said to be “Deep”
  - Can have hundreds of layers
- The corresponding training algorithms are called “deep learning”.

Artificial neural network

- The success of artificial neural networks lies in their ability to approximate the functional relationship the input data and the desired output.
- The approximation is encoded in the weights of the neural network.
- The process of training consists of using the given data to successively modify the value of the weights.

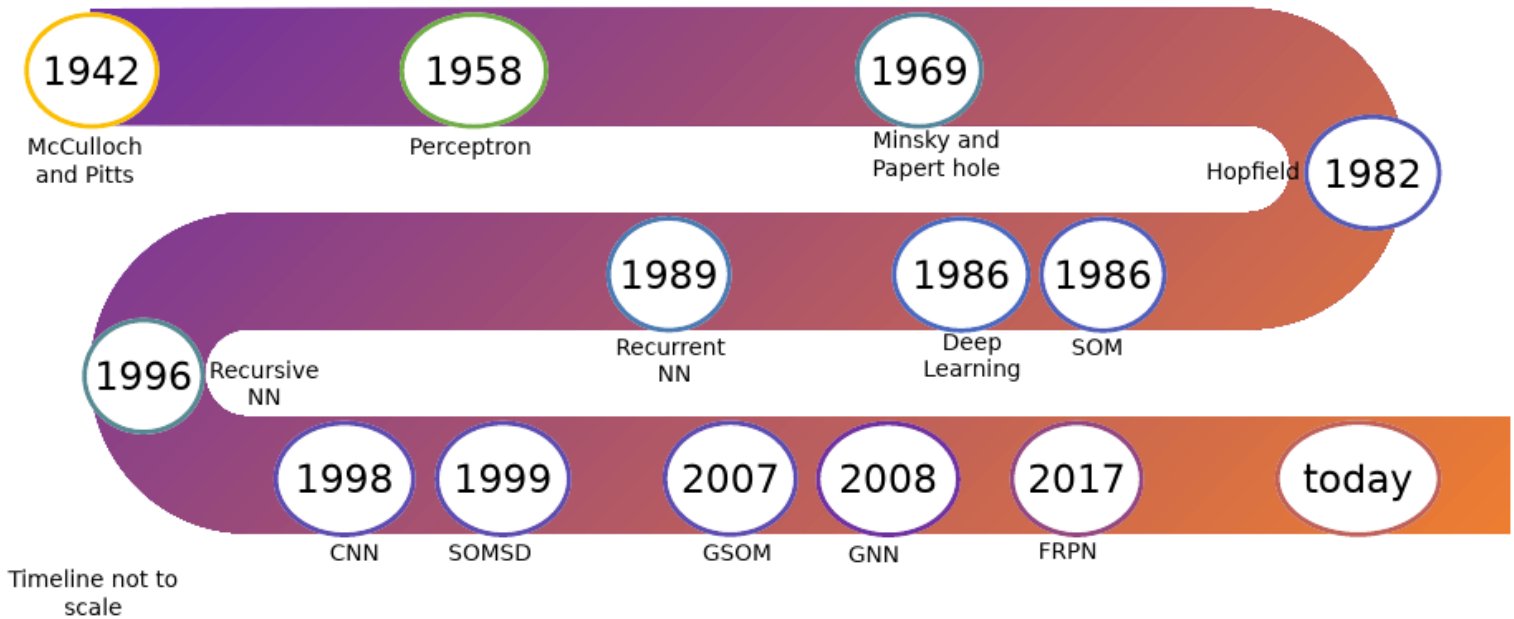


Figure: Historical timeline of development of ANN

Some Applications

Market basket analysis

How to increase sales by better understanding customer purchasing patterns. Use large data sets (e.g. purchase history) to discover product groupings, & co-purchase pattern

Machine Learning Tool

Model problem as a basket-item link prediction task and employ graph convolutional neural network for analysis (Liu, Wan, Guo, Achan, & Yu, 2020)

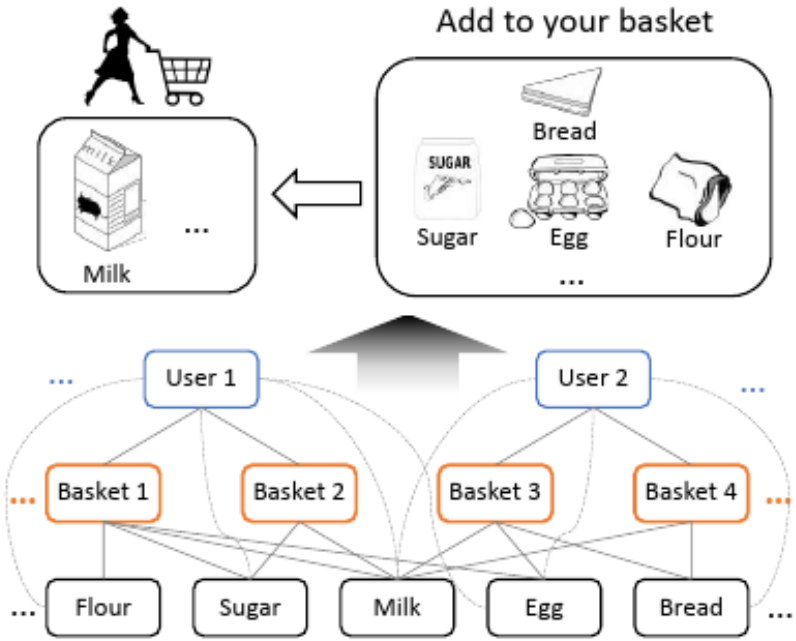


Figure: Basket recommendation; User-basket-item graph; user-basket, basket-item, and user-item interactions.(Liu et al., 2020)

Some applications

Prison term prediction

Judgment prediction aims at automatically predicting the judgment result given a textual description of a legal case.

Machine Learning Tool

Based on the observation that fine-grained feature selection is the key to achieving good performance, a Deep Gating Network (DGN) is proposed for charge-specific feature selection and aggregation. (Chen, Cai, Dai, Dai, & Ding, 2019)

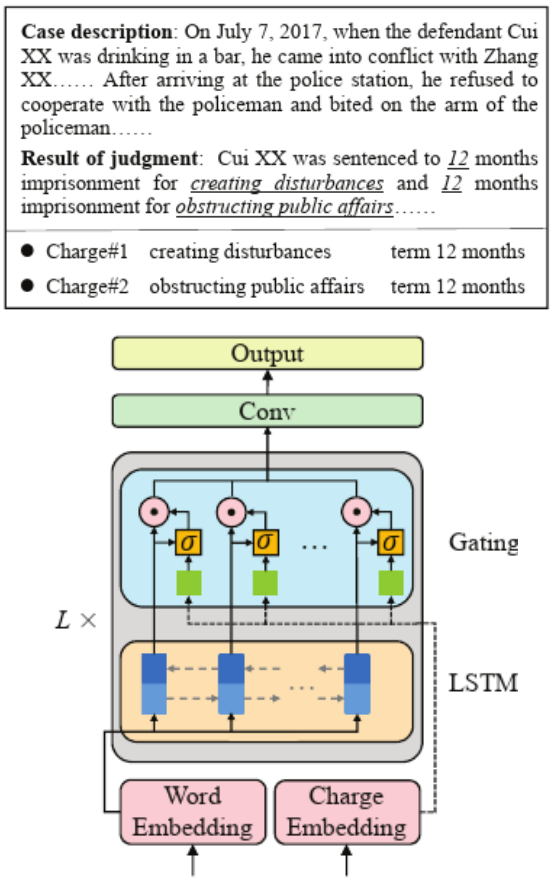


Figure: (Chen et al., 2019).

Some Applications

Automated news generation

1. How can we use data and algorithms to lower the costs of discovering stories?
2. How do we tell stories in more personalized and engaging ways?

Machine Learning Tools

1. Knowledge representation
2. Document classification
3. Natural language generation
4. Machine translation

"Democracy's Detectives"

Investigative journalism is underprovided in the market, but new combinations of data and algorithms may make it easier for journalists to discover and tell the stories that hold institutions accountable - James T. Hamilton

Stanford Computational Journalism Lab

Figure: `http://cjlab.stanford.edu/`.

Some Applications

Sentiment analysis

1. How do we computationally identify and categorize opinions expressed in a piece of text?
2. How to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral

Machine Learning Tools

1. Word representation / embedding
2. Self-attention network
3. Fully connected neural network

Aspect-based sentiment analysis

The aspect-based sentiment analysis (ABSA) consists of two conceptual tasks, namely an aspect extraction and an aspect sentiment classification (Phan & Ogunbona, 2020)

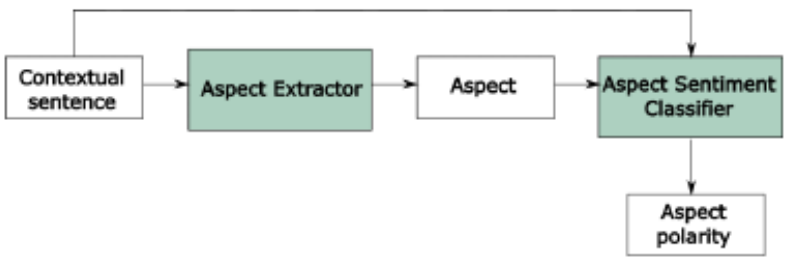


Figure: Aspect-based sentiment analysis (Phan & Ogunbona, 2020).

Some Applications

Network Intrusion Detection

1. Identify attacks before damage.
2. Identify effectiveness of a security policy.
3. Identify effectiveness of security mechanisms.
4. Identify zero-day attacks.

Machine Learning Tools

1. Hybrid ensemble systems.
2. Nonlinear classification methods.
3. Visual analytics.

Ensemble intrusion detection system

- Combine SOM (for clustering, visualization and data augmentation) and DL (for classification).

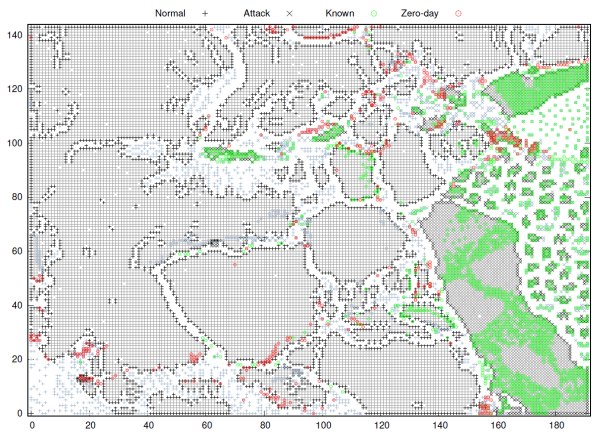


Figure: Intrusion Detection System by A.Saraswati, M.Hagenbuchner, and A.C.Tsoi (submitted 2022).

Some Applications

Big Data

1. Working with data streams.
2. Working with unreliable data.
3. Working with changing data.
4. Working with different forms of data.

Machine Learning Tools

1. Scalable, just-in-time learning and processing methods.
2. Massive parallel methods.
3. Adaptive systems.
4. Simple methods.

Location Analytics

- Find spatio-temporal patterns indicative of physical activities or social behavior.
- Information enrichment: add maps, imagery, demographics, consumer and lifestyle data, environment and weather, etc.

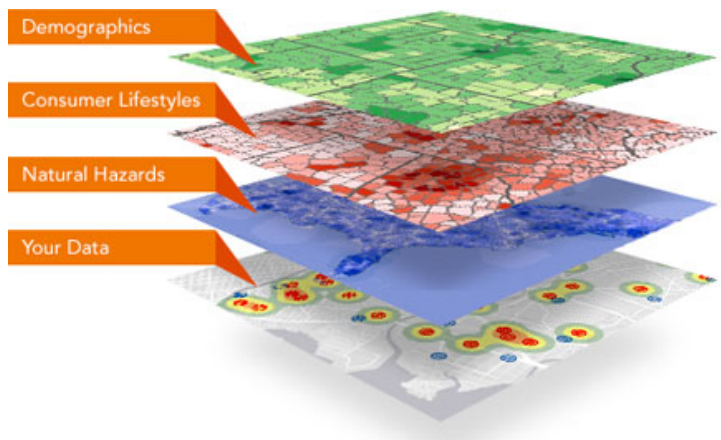


Figure: Location Analytics (Source: `http://www.esri.com/software/location-analytics`.

Grand challenges

Some Grand Challenges

1. Modeling Our Planet's Systems: Assessing global warming and determining mitigating actions
2. Confronting Existential Risk: What is the impact of a dangerous genetically modified pathogen
3. Exploring Transhumanism: What is the impact of embedded nanotechnology, genetic therapy, and "smart" prosthetics?
4. The Singularity? What happens when systems approach the level of human intelligence? Emotional intelligence?
5. Dealing Effectively with Globalism: Modeling the interconnected of human societies/organizations
6. How do we keep AI accountable? Issues of ethics and bias.

Q & A

QUESTIONS & ANSWERS

Subsequent lectures

Module 2: A gentle introduction to Deep Learning
(Thursday 1/December/2022)

Module 3: A gentle introduction to tools and systems of AI, ML and DL
(14/February/2022)

Module 4: A gentle introduction to Natural Language Processing
(16/February/2022)

Consultation

- We are available to offer advise and consultation services.
- Explore opportunities to solve problems in your area of research/work by using AI.
- The power of joined expertise.
- Contact us. Visit https://uow.info/cai

Bibliography I

Chen, H., Cai, D., Dai, W., Dai, Z., & Ding, Y. (2019, November 3 - 7). Charge-based prison term prediction with deep gating network. In *Proceedings, 2019 ACL conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing* (p. 6362 - 6367). Hong Kong, China.

Géron, A. (2019). *Hands-on machine learning with scikit-learn, keras & tensorflow concepts, tools, and techniques to build intelligent systems*. Boston, Beijing, Tokyo: O'Reilly.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Boston Massachusettes, London England: The MIT Press.

Liu, Z., Wan, M., Guo, S., Achan, K., & Yu, P. S. (2020). BasConv: Aggregating heterogeneous interactions for basket recommendation with graph convolutional neural network. *CoRR, abs/2001.09900*. Retrieved from `https://arxiv.org/abs/2001.09900`

Bibliography II

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. Boston Massachusettes, London England: The MIT Press.

Phan, M. H., & Ogunbona, P. O. (2020, July 5 - 10). Modelling context and syntactical features for aspect-based sentiment analysis. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 3211–3220). Seatle, Washington, USA.

Thank you

UNIVERSITY
OF WOLLONGONG
AUSTRALIA