Computer Science Department
Shahid Beheshti University

A Friendly Introduction to Deep Learning

Dr. Hadi Farahani

What is Artificial Intelligence?

And why we need it?

Artificial Intelligence

Artificial Intelligence is nothing but the capability of a machine to imitate intelligent human behavior. AI is achieved by mimicking a human brain, by understanding how it thinks, how it learns, decides, and work while trying to solve a problem.

Applications of Artificial Intelligence

Now think about this, instead of you doing all your work, you have a machine to finish it for you or it can do something which you thought was not possible at all. For instance:

Marketing

Recommender Systems

Health Care

Bio-markers

Predicting Heart Attacks

Ridesharing Apps

Determine the Price of Your Ride

Autonomous vehicles

Self-driving Cars

Social Media

Face Verification

Tagging People in Photos

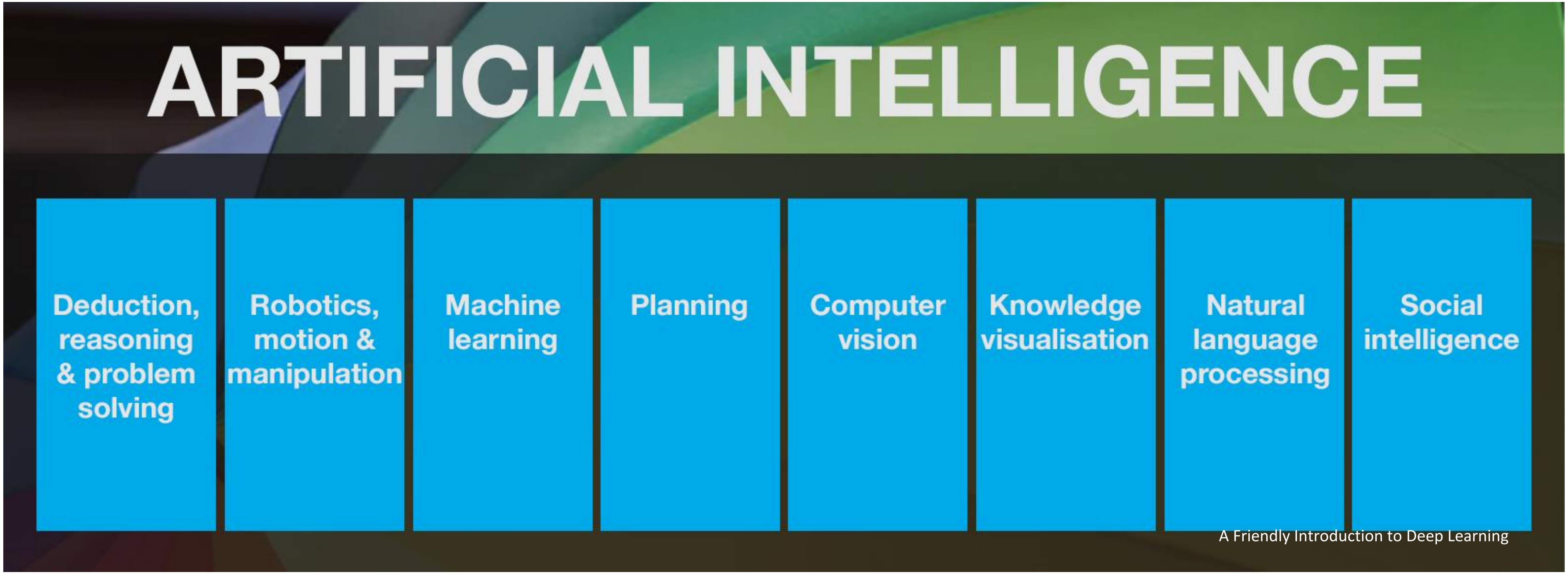
Speech Recognition

Voice-to-Text

Smart Personal Assistants

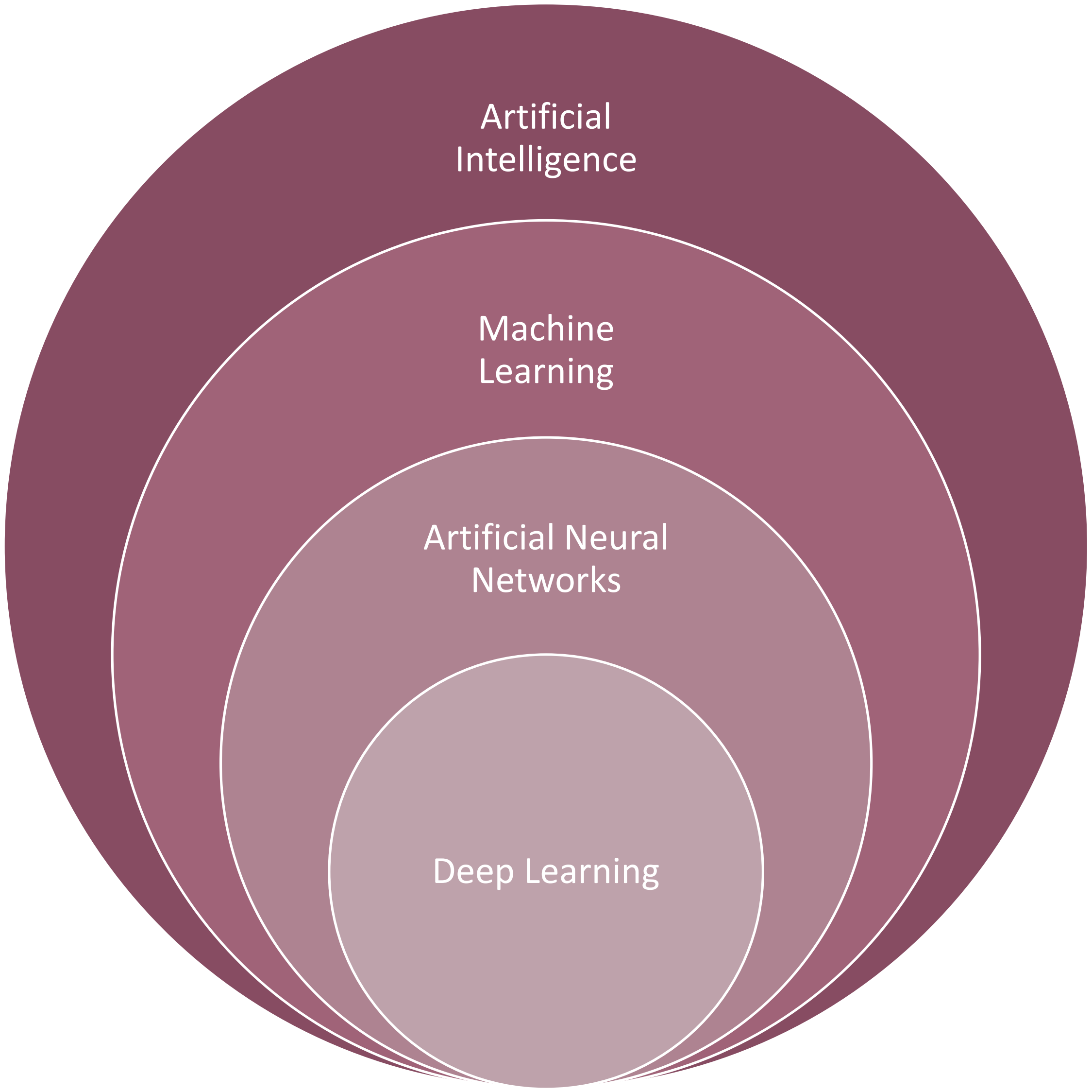
Artificial Intelligence

And its areas of application

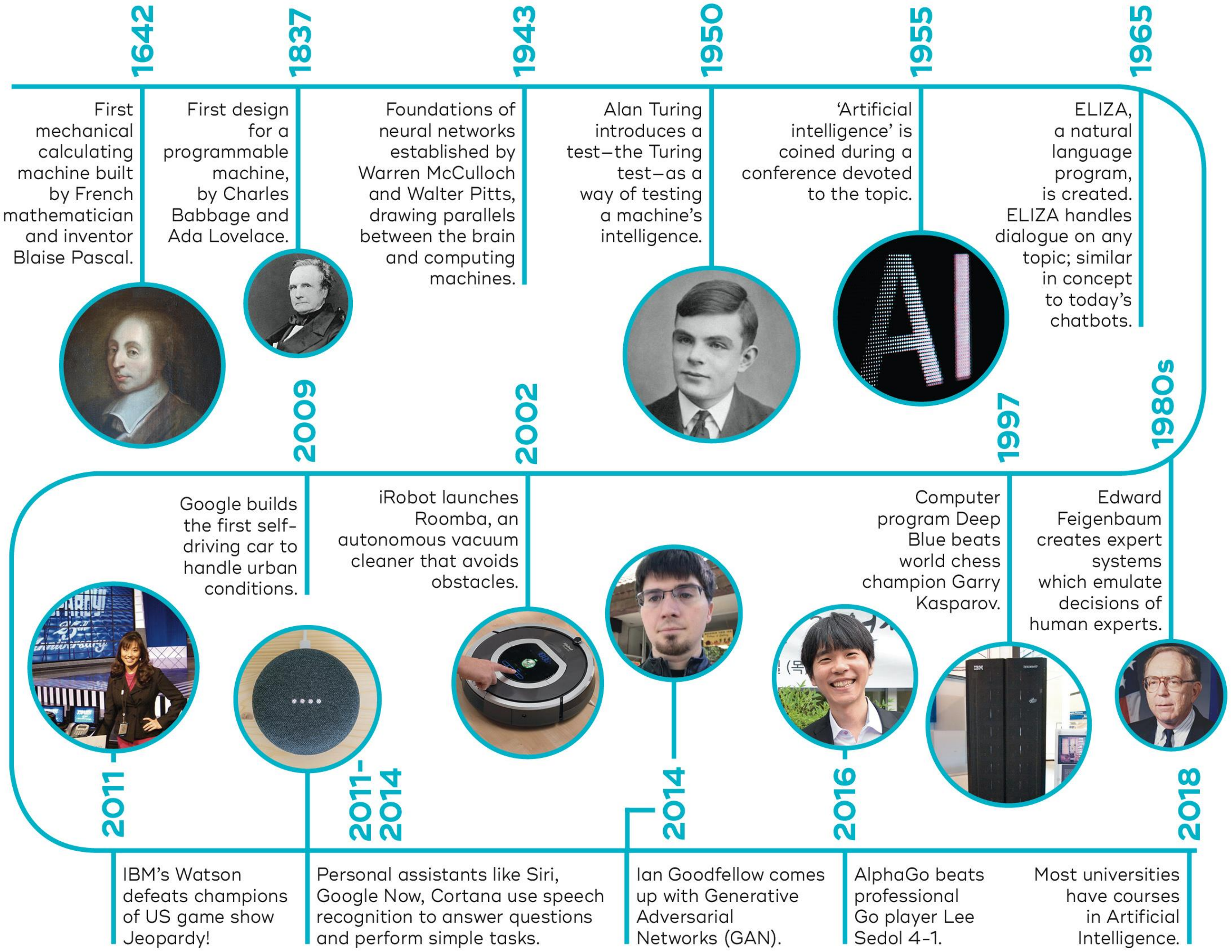


Subsets of Artificial Intelligence

Basically, Deep learning is a sub-field of Artificial Neural Networks, and Artificial Neural Networks are sub-field of Machine Learning, and Machine Learning is a sub-field of Artificial Intelligence.

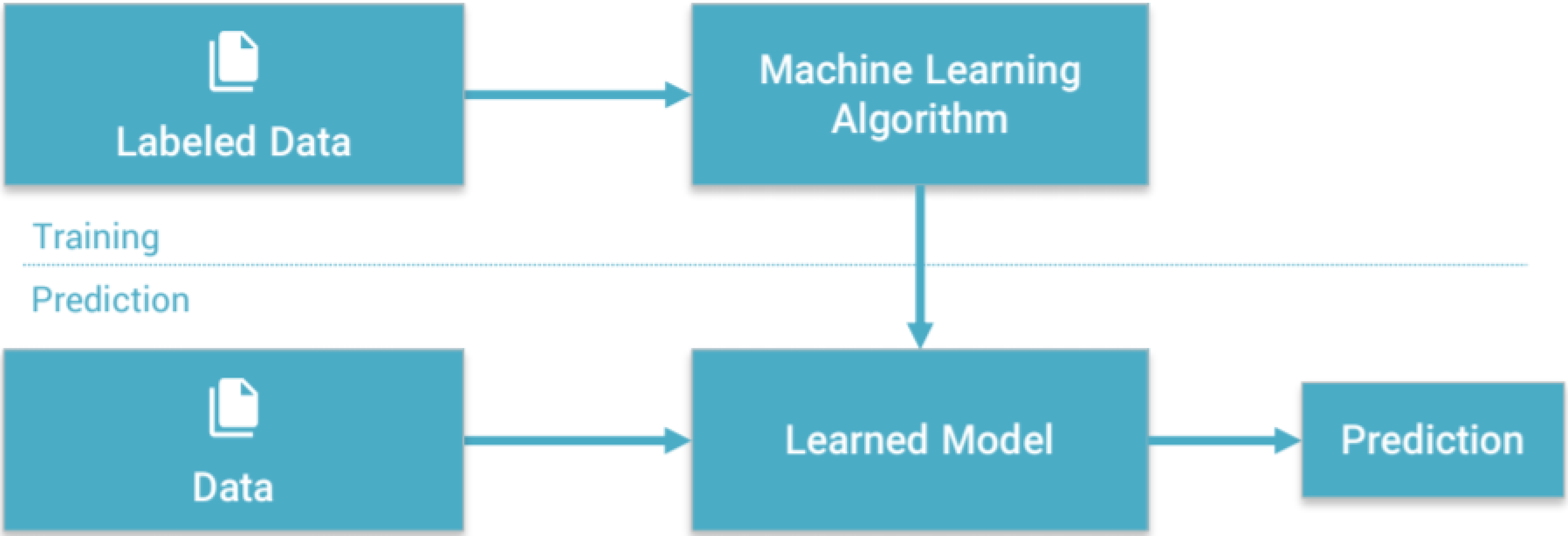


AI Brief History



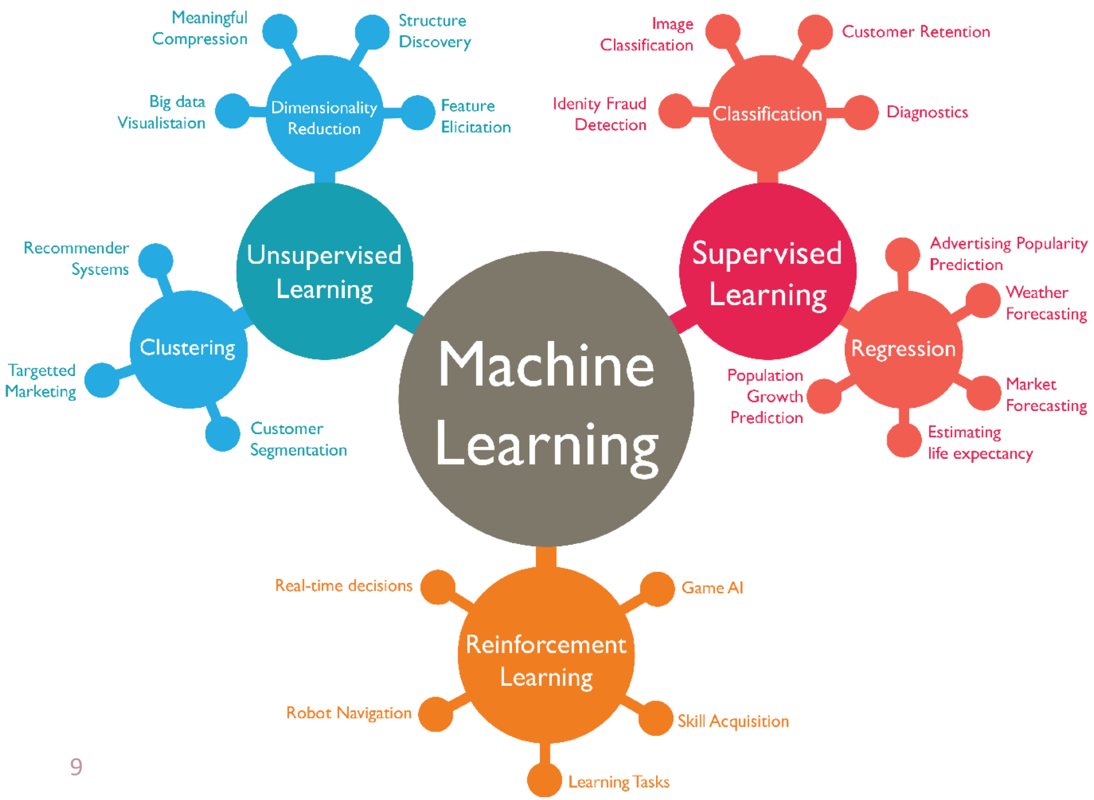
Machine Learning basics

Machine Learning is a type of Artificial Intelligence that provides computers with the ability to learn without being explicitly programmed. It provides various techniques that can learn from and make predictions on data.



Machine Learning Basics

Learning Approaches



Supervised Learning: Learning with a labeled training set

Example: email spam detector with training set of already labeled emails

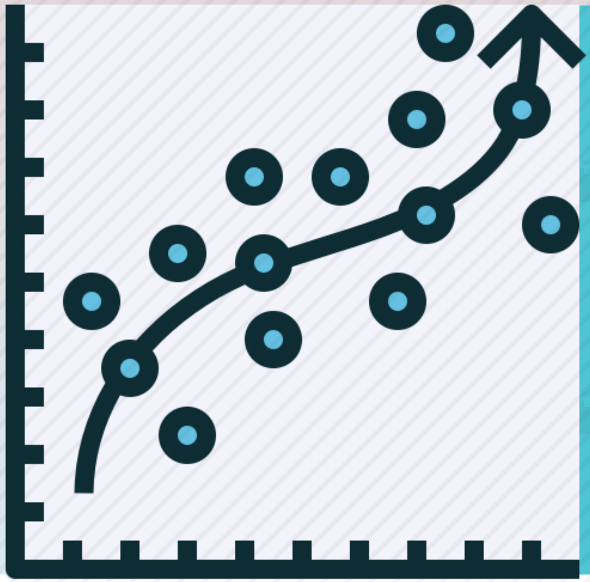
Unsupervised Learning: Discovering patterns in unlabeled data

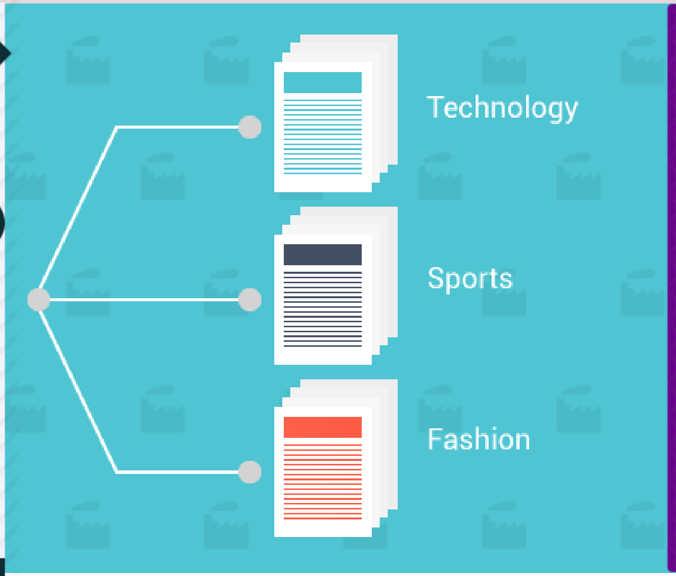
Example: cluster similar documents based on the text content

Reinforcement Learning: learning based on feedback or reward

Example: learn to play chess by winning or losing

Problem Types of Machine Learning





Regression

Supervised- Predictive

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas porttitor congue massa.

Classification

Supervised - Predictive

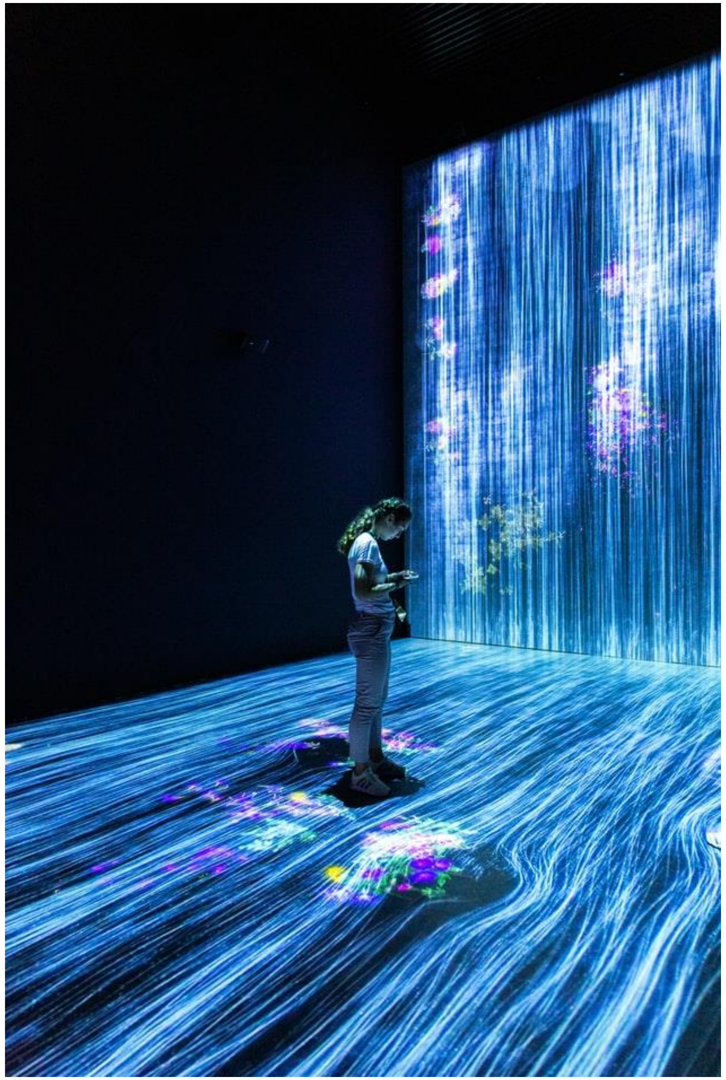
Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas porttitor congue massa.

Clustering

Unsupervised - Descriptive

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas porttitor congue massa.

What is Deep Learning?



Part of the machine learning field of learning representations of data. Exceptional effective at learning patterns.

Utilizes learning algorithms that derive meaning out of data by using a hierarchy of multiple layers that mimic the neural networks of our brain.

If you provide the system tons of information, it begins to understand it and respond in useful ways.

Inspired by Brain

- At this stage, it could be said that deep learning mimics an infant’s brain. An infant’s brain is like a sponge, and it learns through training. It takes some years for the web of neural networks in it to mature and infer or deduce multiple things through one set of training.

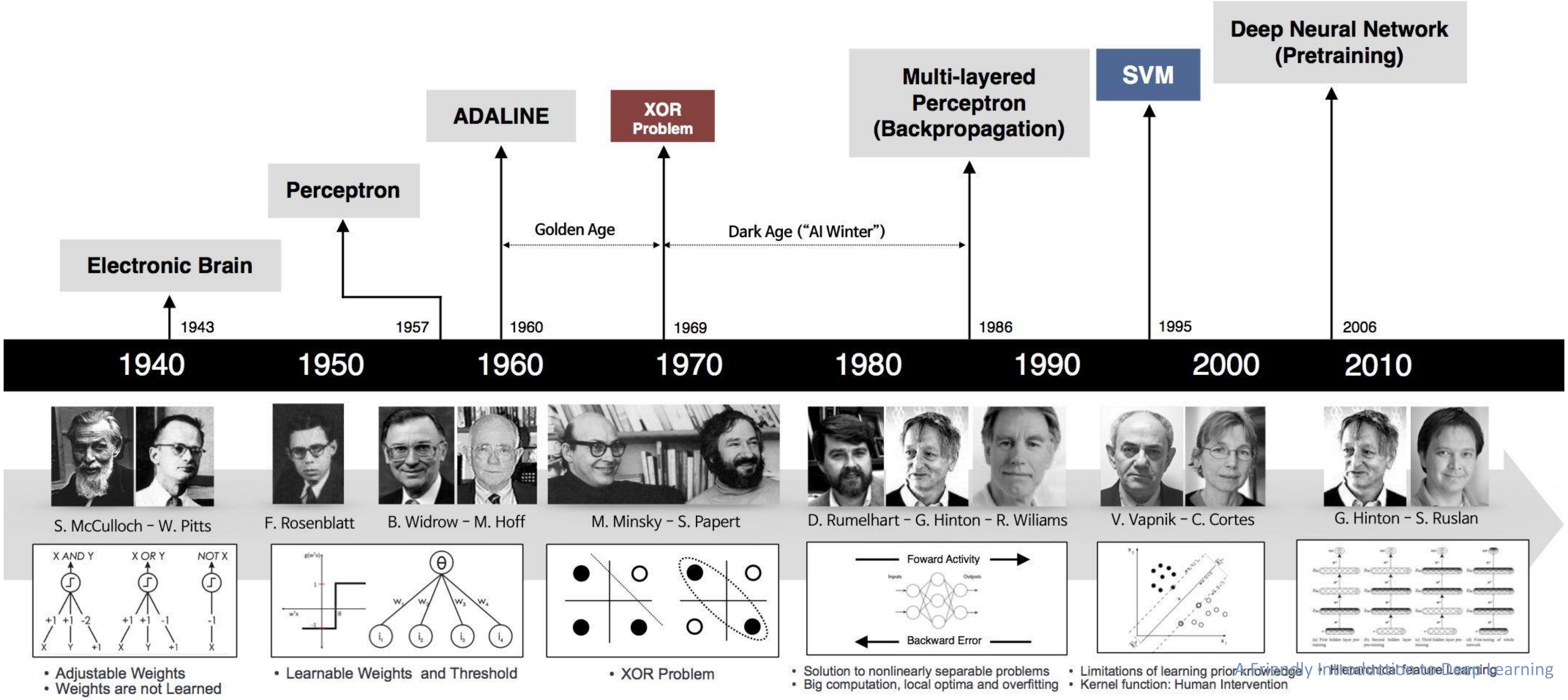
Deep Learning Applications

Speech Recognition

Natural Language Processing

Computer Vision

A Brief History of Deep Learning



Top Players

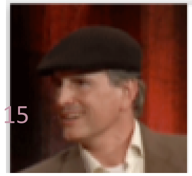
Superstar Researchers

Geoffrey Hinton: University of Toronto & Google

Yann LeCun: New York University & Facebook

Andrew Ng: Stanford & Baidu

Yoshua Bengio: University of Montreal



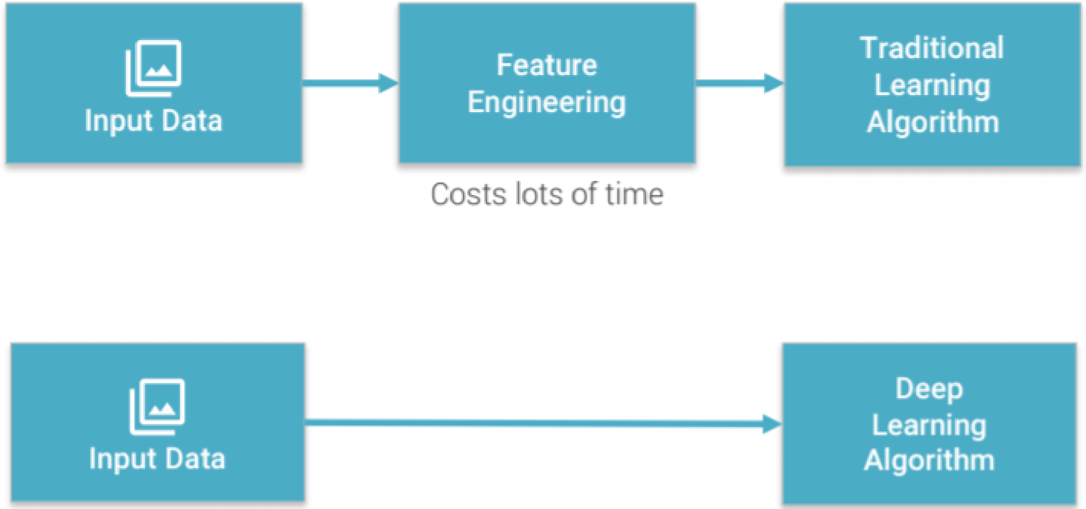
Jürgen Schmidhuber: Swiss AI Lab & NNAISENSE

The Big Players

Companies

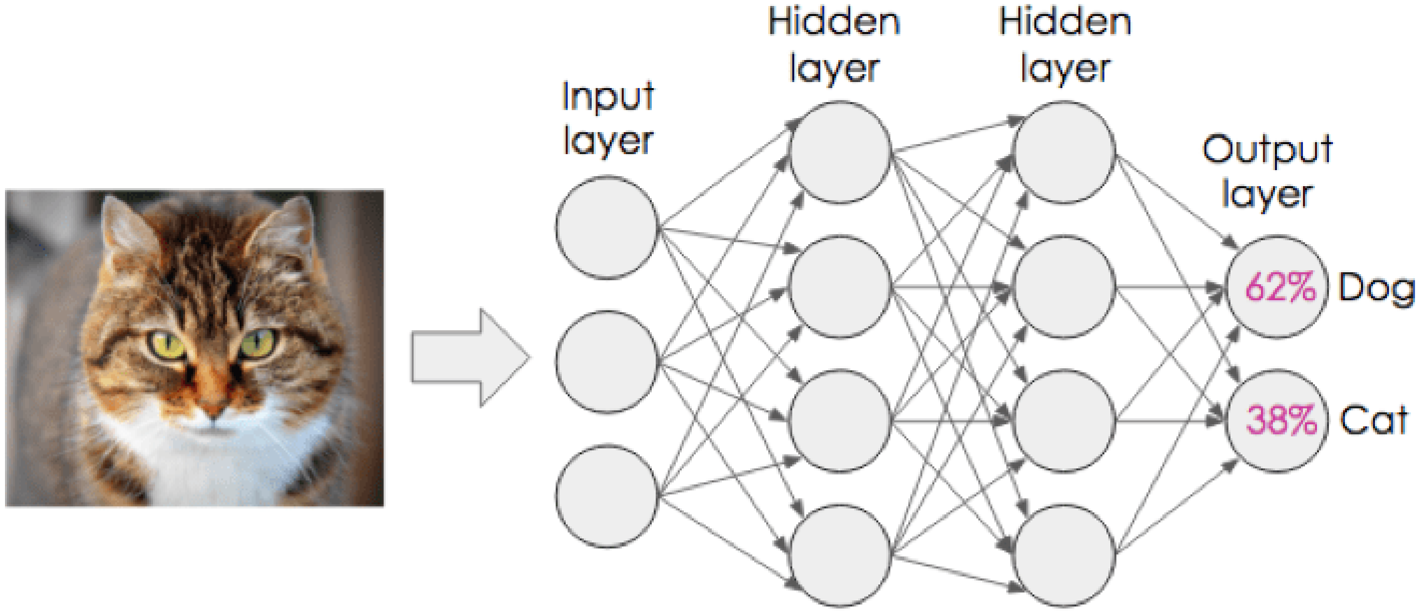
Deep Learning Basics

No more feature engineering



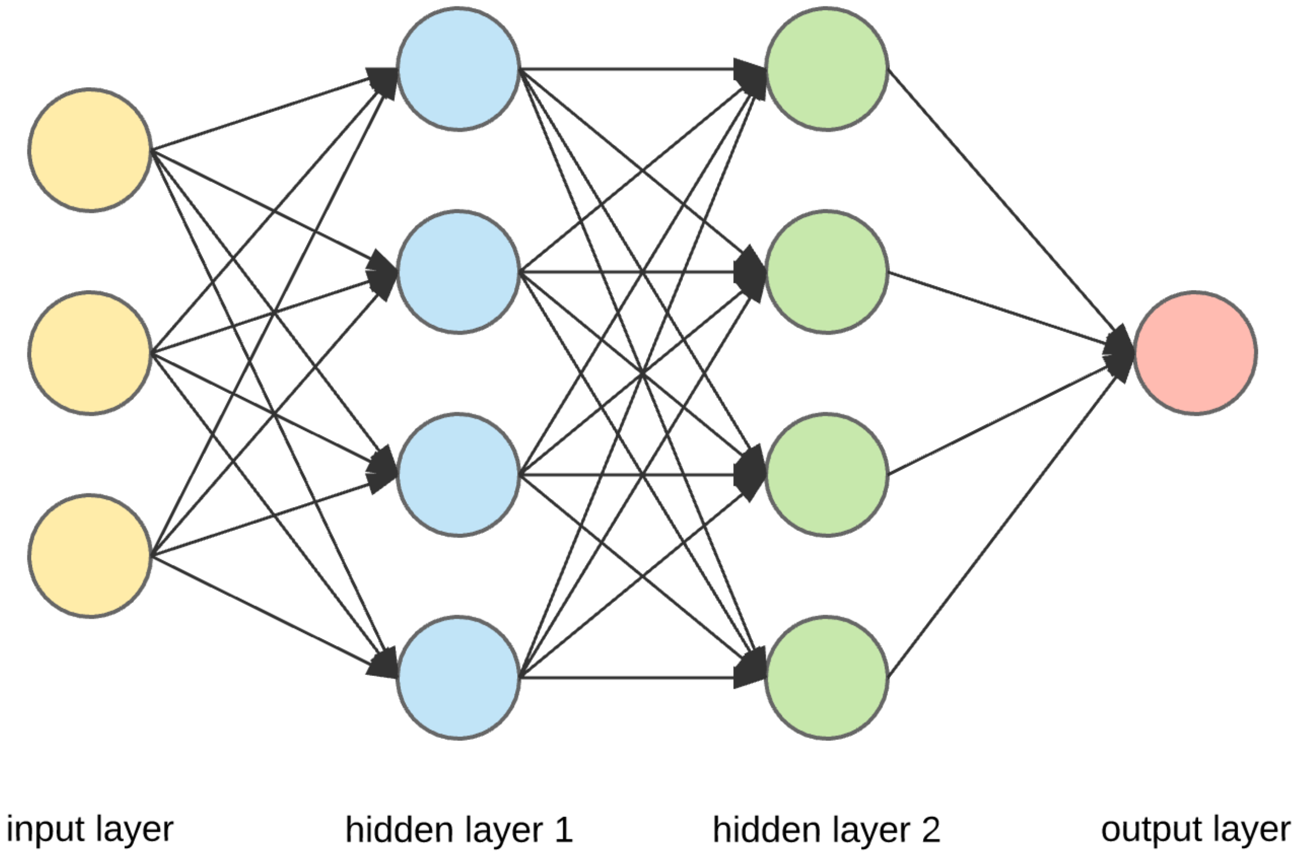
Architecture of Deep Neural Networks

A deep neural network consists of a hierarchy of layers, whereby each layer transforms the input data into more abstract representations (e.g. edge -> nose -> face). The output layer combines those features to make predictions.



Artificial Neural Networks

Consists of one input, one output and multiple fully-connected hidden layers in-between. Each layer is represented as a series of neurons and progressively extracts higher and higher-level features of the input until the final layer essentially makes a decision about what the input shows. The more layers the network has, the higher-level features it will learn.

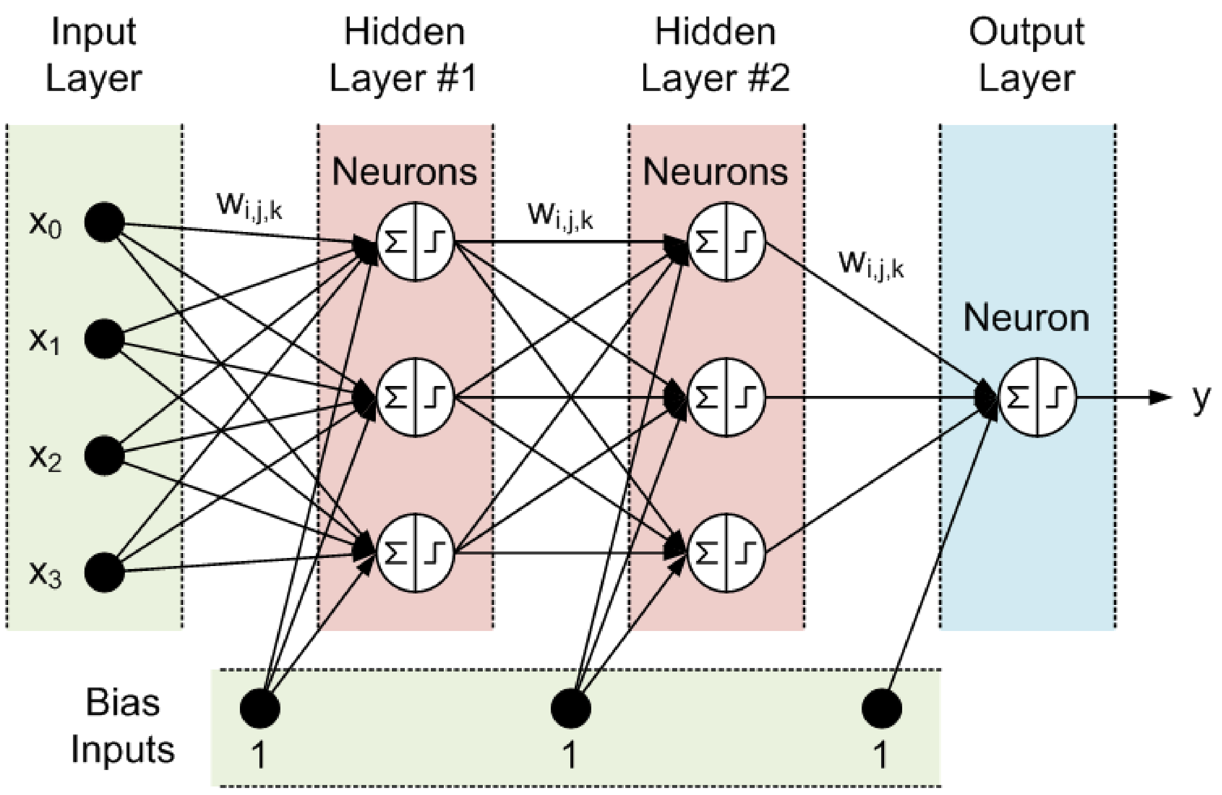


Layers of ANN

- **Input Nodes (input layer):** No computation is done here within this layer, they just pass the information to the next layer (hidden layer most of the time). A block of nodes is also called **layer**.

- **Hidden nodes (hidden layer):** In Hidden layers is where intermediate processing or computation is done, they perform computations and then transfer the weights (signals or information) from the input layer to the following layer (another hidden layer or to the output layer). It is possible to have a neural network without a hidden layer and I'll come later to explain this.

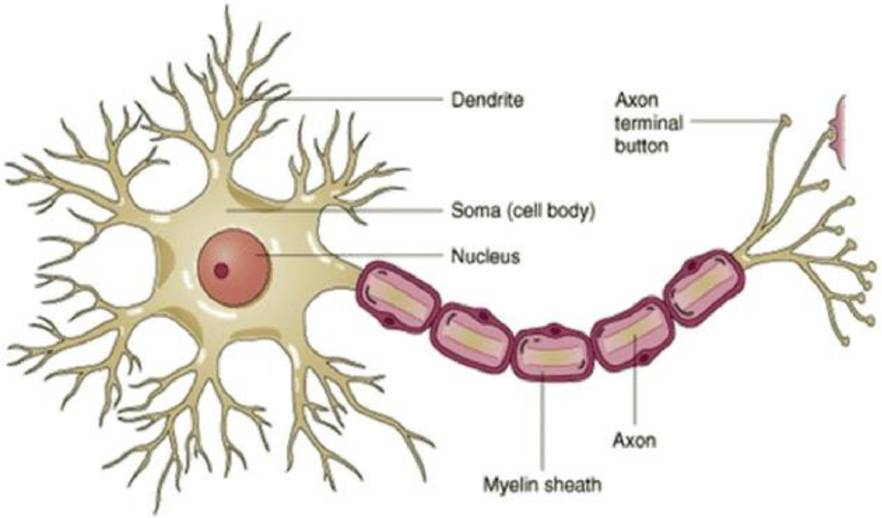
- **Output Nodes (output layer):** Here we finally use an activation function that maps to the desired output format (e.g. softmax for classification).



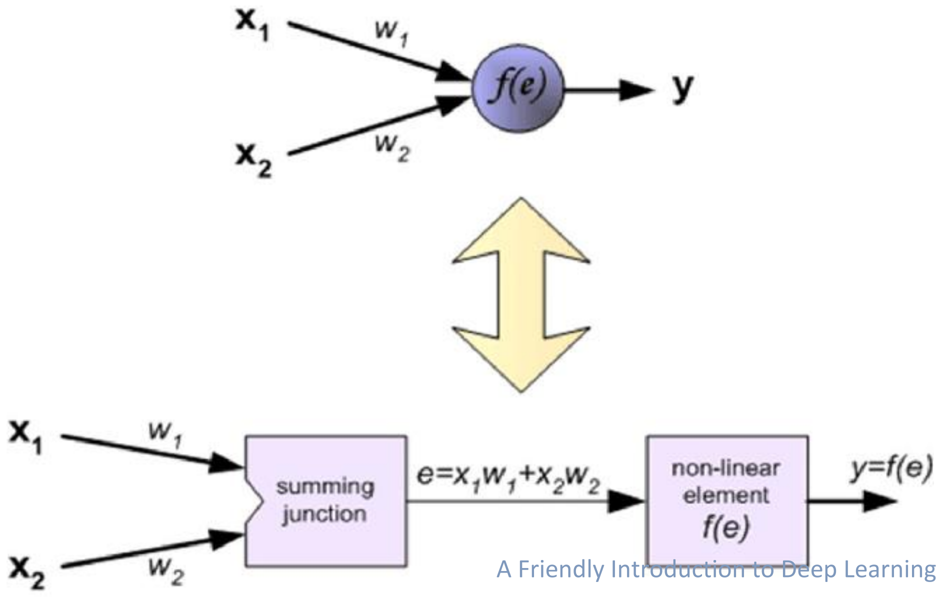
The Neuron

- The basic computational unit of the brain is a **neuron**. Approximately 86 billion neurons can be found in the human nervous system and they are connected with approximately 10^{14} — 10^{15} **synapses**.

- The basic unit of computation in a neural network is the neuron , often called a node or unit. It receives input from some other nodes, or from an external source and computes an output. Each input has an associated weight (w), which is assigned on the basis of its relative importance to other inputs. The node applies a function to the weighted sum of its inputs.



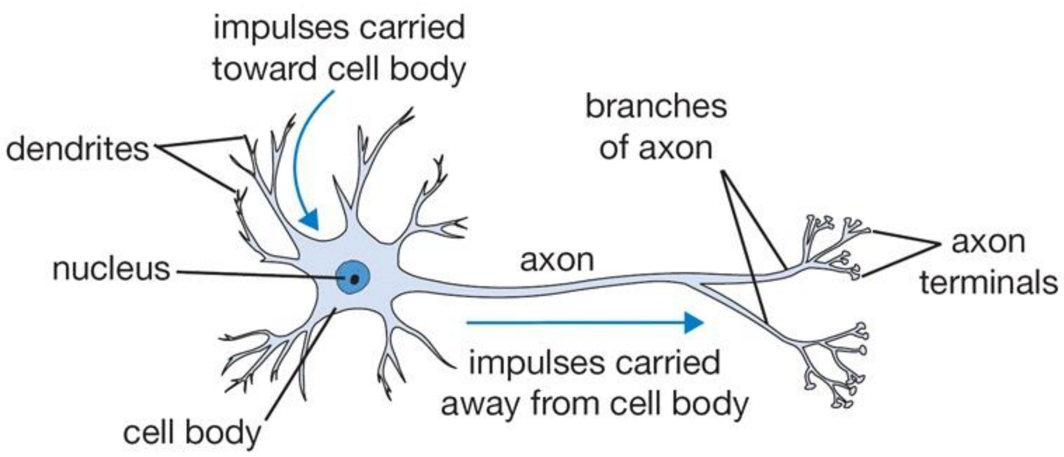
(a) Biological neuron

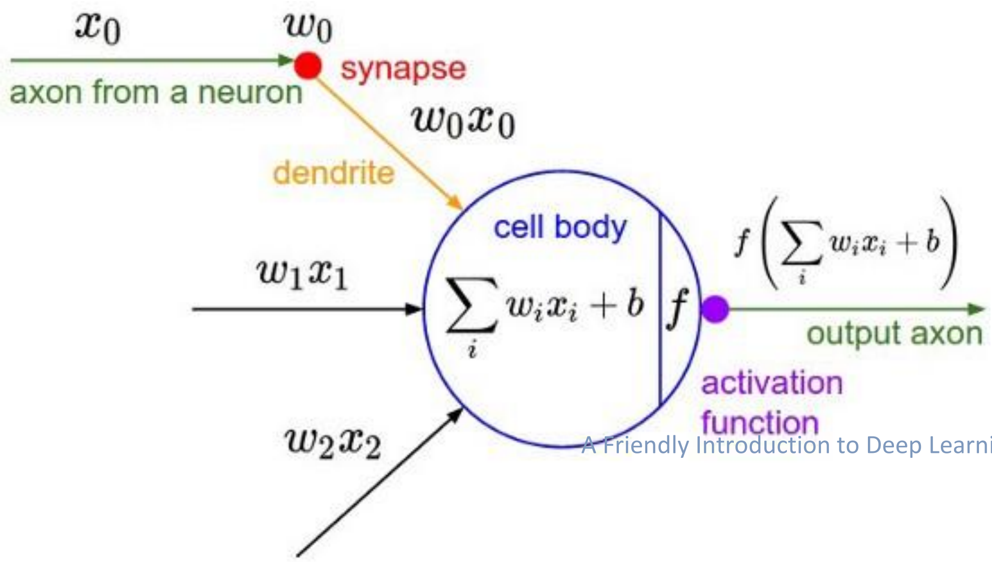


(b) Artificial neural network

Neuron and Its Mathematical Model

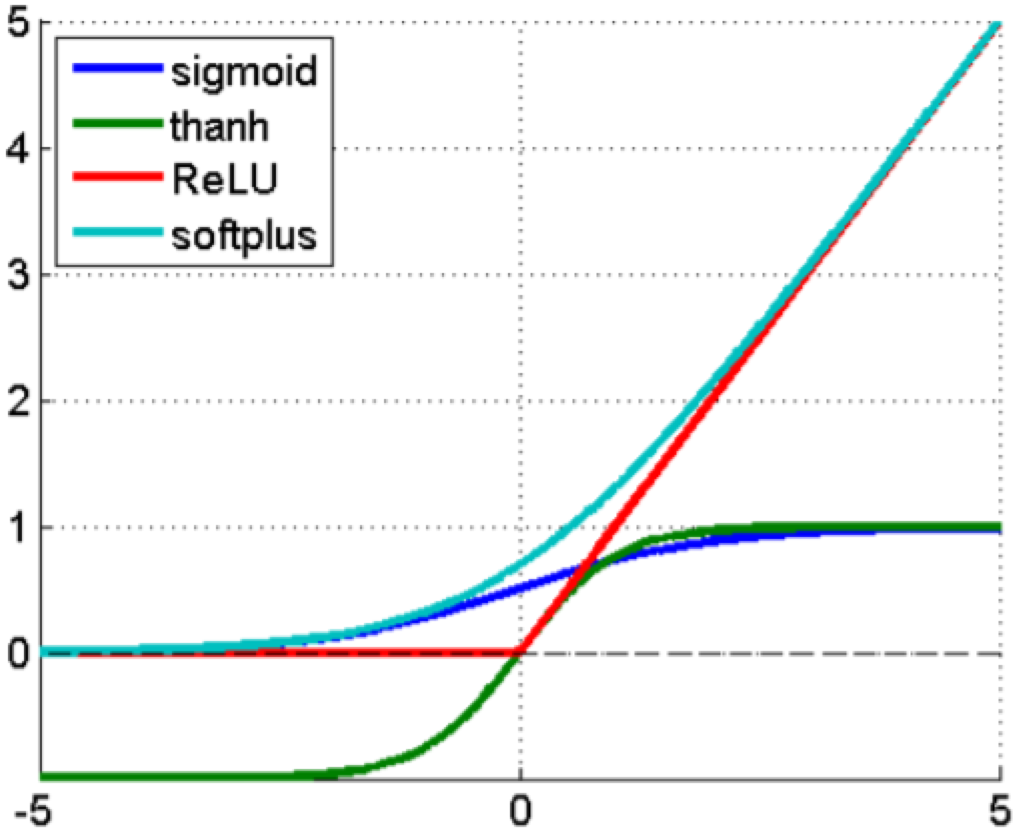
The idea is that the synaptic strengths (the weights *w*) are learnable and control the strength of influence and its direction: excitory (positive weight) or inhibitory (negative weight) of one neuron on another. In the basic model, the dendrites carry the signal to the cell body where they all get summed. If the final sum is above a certain threshold, the neuron can *fire*, sending a spike along its axon. In the computational model, we assume that the precise timings of the spikes do not matter, and that only the frequency of the firing communicates information. we model the *firing rate* of the neuron with an **activation function** *(e.x. sigmoid function)*, which represents the frequency of the spikes along the axon.





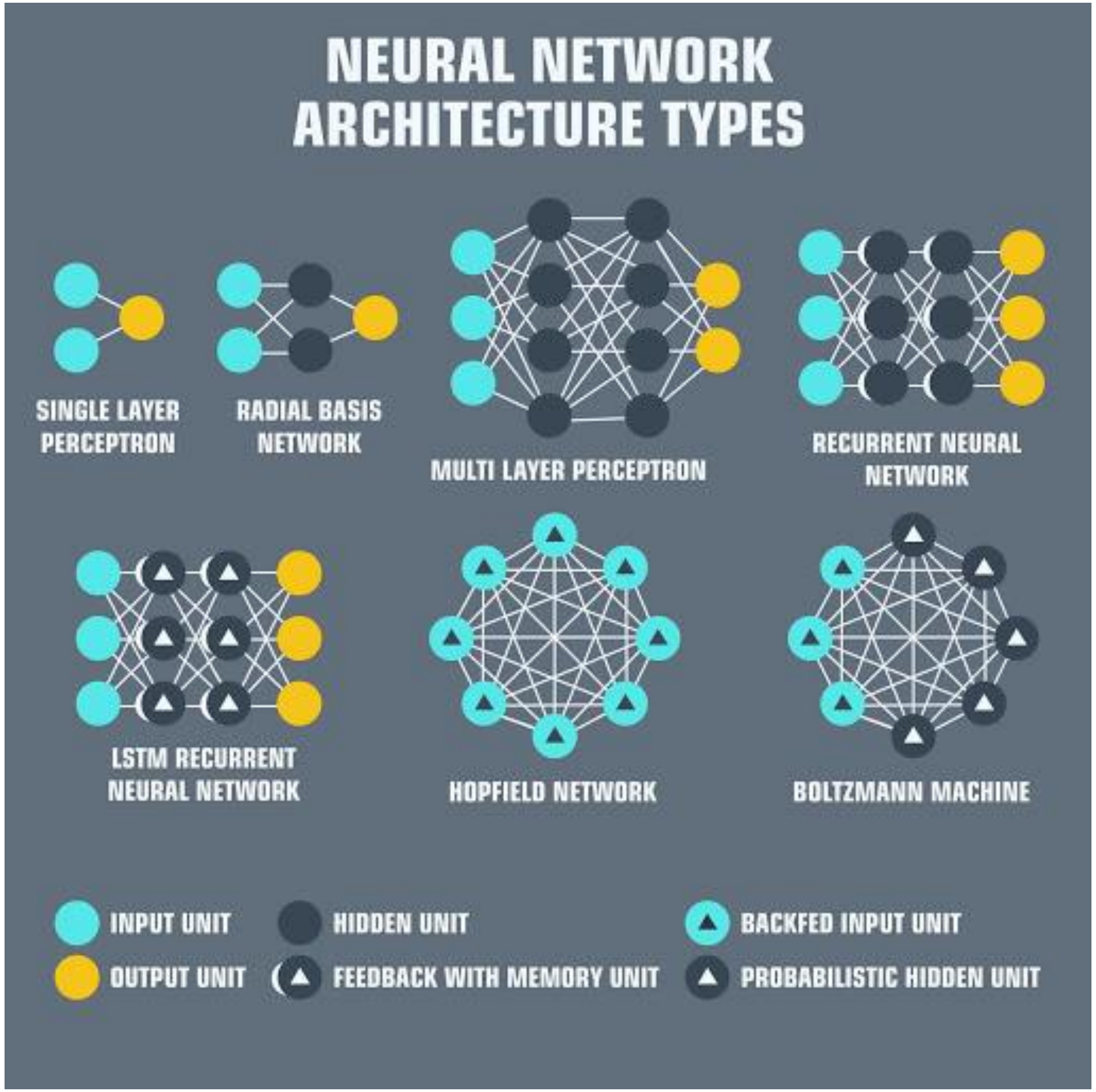
Activation Function

- The **activation function** of a node defines the output of that node given an input or set of inputs.
- In the 1940s and 1950s artificial neurons used a step activation function and were called *perceptron*. Modern neural networks may *say* they are using perceptron, but they actually have smooth activation functions, such as the logistic or sigmoid function, the hyperbolic tangent, and the Rectified Linear Unit (ReLU). ReLU is usually the best choice for fast convergence, although it has an issue of neurons "dying" during training if the learning rate is set too high.



Types of Neural Networks

There are many classes of neural networks and these classes also have sub-classes, here we will list the most used ones:



Feedforward Neural Network

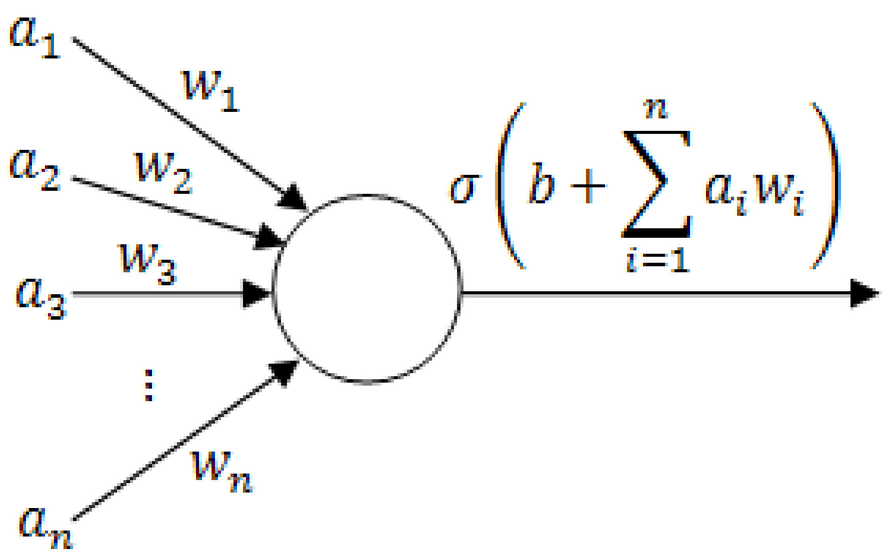
- A feedforward neural network is an artificial neural network where connections between the units do *not* form a cycle. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network.

- We can distinguish three types of feedforward neural networks:

  - ✓ Single-layer Perceptron
  - ✓ Multi-layer Perceptron
  - ✓ convolutional Neural Network (CNN)

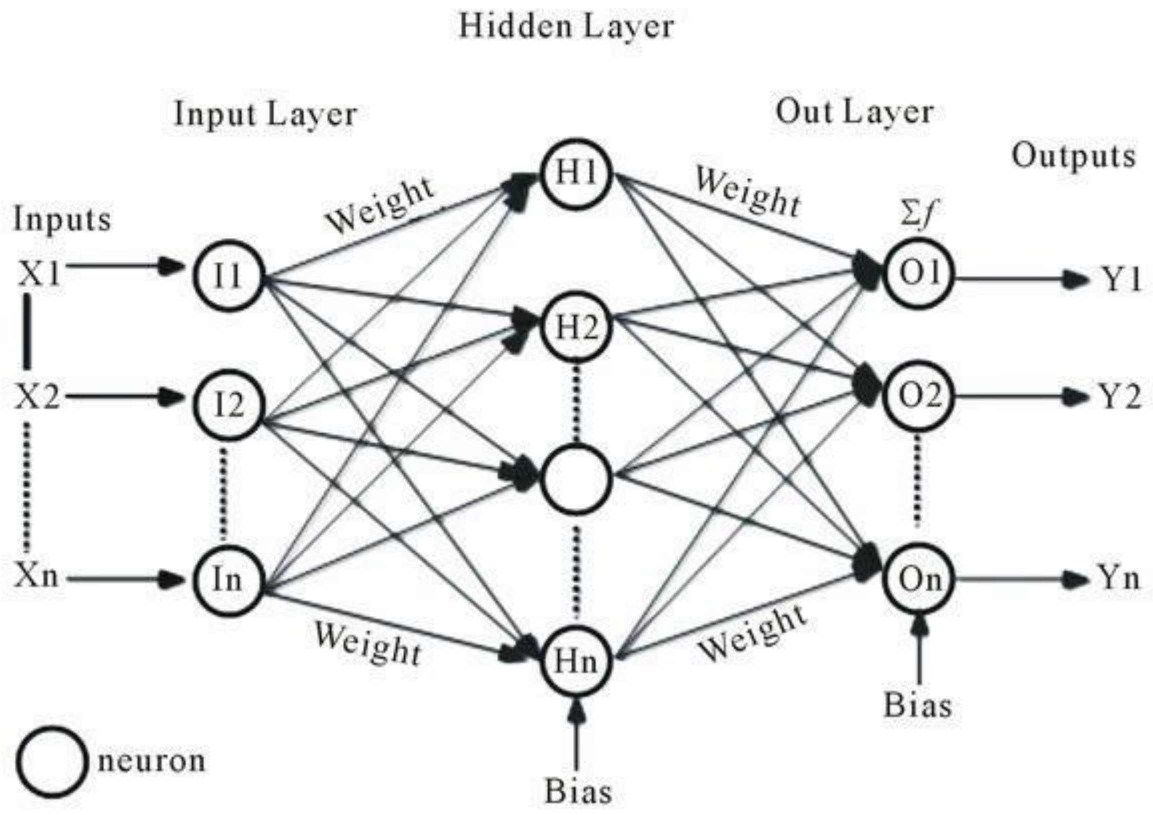
Single-layer Perceptron

- This is the simplest feedforward neural Network and does not contain any hidden layer, Which means it only consists of a single layer of output nodes. This is said to be single because when we count the layers we do not include the input layer, the reason for that is because at the input layer no computations is done, the inputs are fed directly to the outputs via a series of weights.



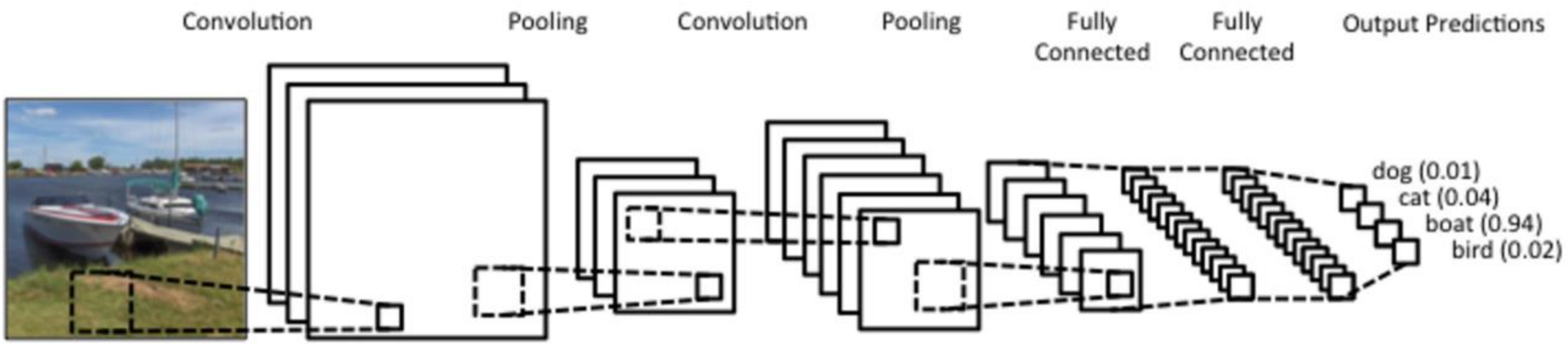
Multi-layer perceptron (MLP)

- This class of networks consists of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications the units of these networks apply a sigmoid function as an activation function. MLP are very more useful and one good reason is that, they are able to learn non-linear representations (most of the cases the data presented to us is not linearly separable).



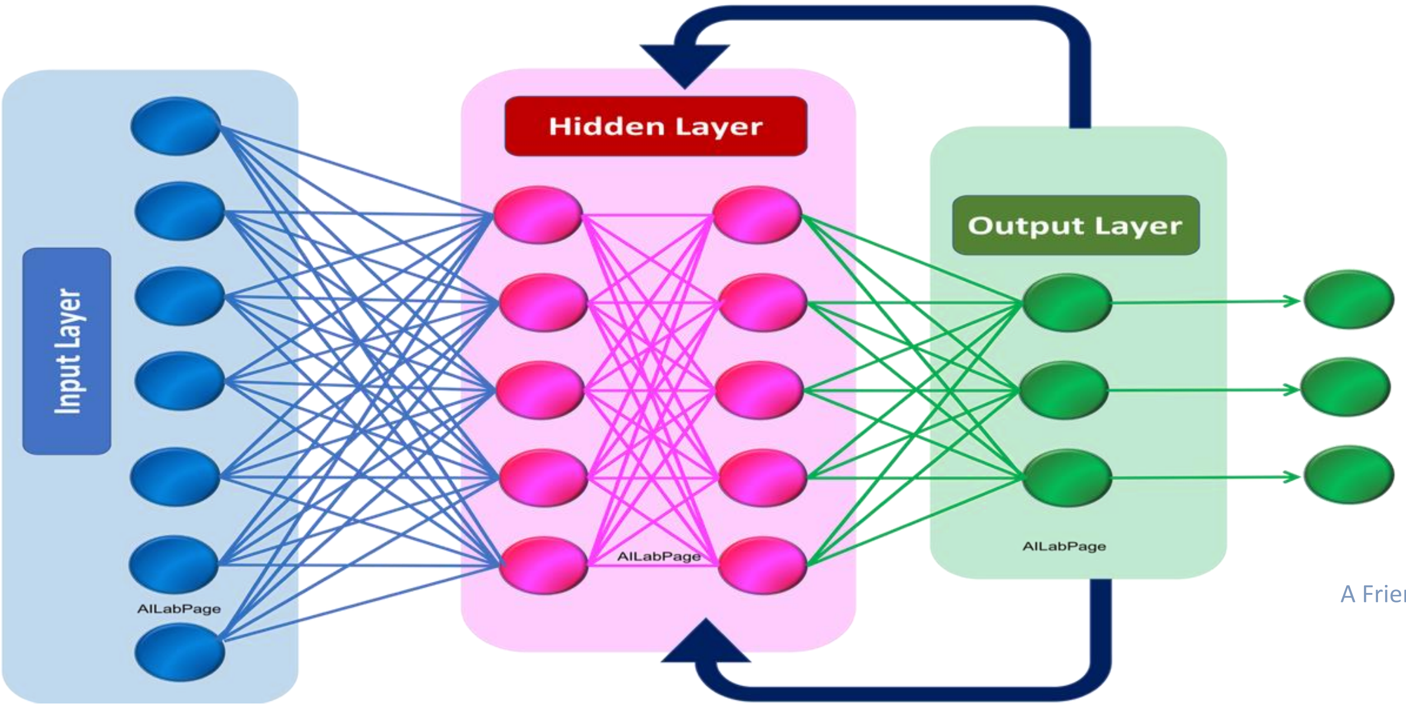
Convolutional Neural Network (CNN)

- Convolutional Neural Networks are very similar to ordinary Neural Networks, they are made up of neurons that have learnable weights and biases. In convolutional neural network (CNN, or ConvNet or shift invariant or space invariant) the unit connectivity pattern is inspired by the organization of the visual cortex, Units respond to stimuli in a restricted region of space known as the receptive field. Receptive fields partially overlap, over-covering the entire visual field. Unit response can be approximated mathematically by a convolution operation. They are variations of multilayer perceptrons that use minimal preprocessing. Their wide applications is in image and video recognition, recommender systems and natural language processing. CNNs requires large data to train on.

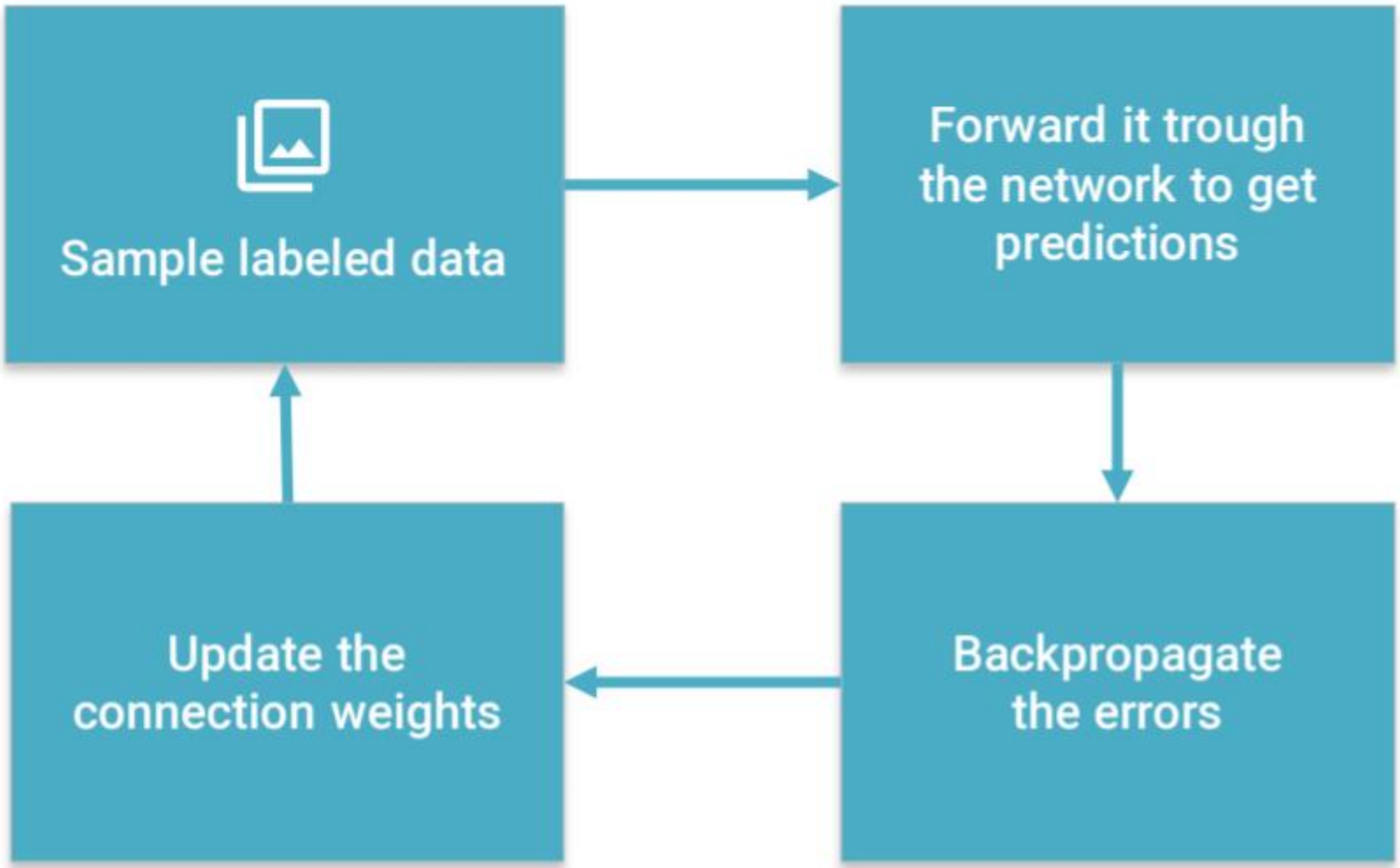


Recurrent Neural Networks

- In recurrent neural network (RNN), connections between units form a directed cycle (they propagate data forward, but also backwards, from later processing stages to earlier stages). This allows it to exhibit dynamic temporal behavior. Unlike feedforward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition, speech recognition and other general sequence processors.



The Training Process



- Learns by generating an error signal that measures the difference between the predictions of the network and the desired values and then using this error signal to change the weights (or parameters) so that predictions get more accurate.

Image Captioning – Combining CNN and RNN

Neural Image Caption Generator generates fitting natural-language captions only based on the pixels by combining a vision CNN and a language-generating RNN.

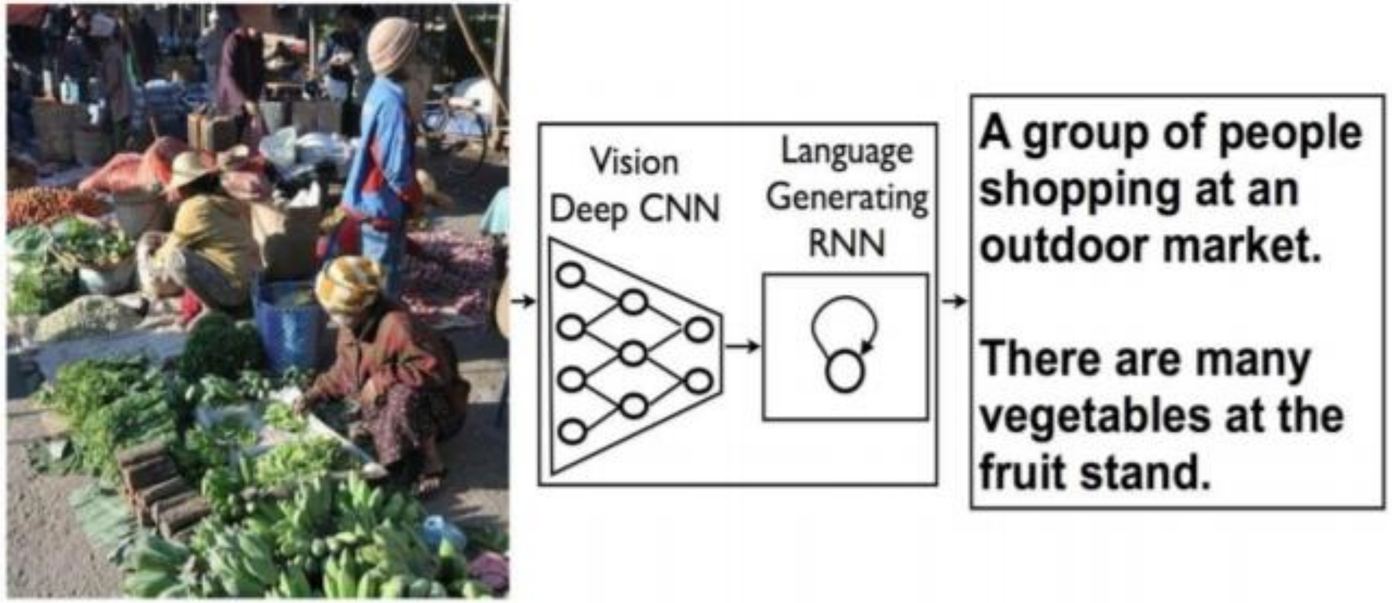


Image Captioning – Combining CNN and RNN

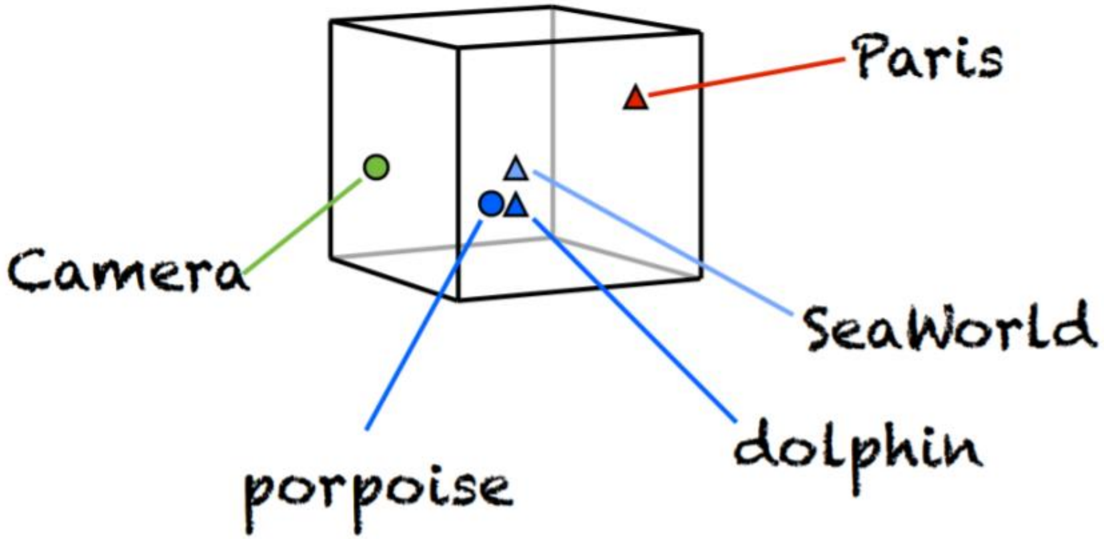
A close up of a child holding a stuffed animal

Two pizzas sitting on top of a stove top oven

A man flying through the air while riding a skateboard

Natural Language Processing – Embeddings

- Embeddings are used to turn textual data (words, sentences, paragraphs) into high-dimensional vector representations and group them together with semantically similar data in a vector-space. Thereby, computer can detect similarities mathematically.



Usage Requirements of Deep Learning

- Large data set with good quality (input-output mappings)
- Measurable and describable goals (define the cost)
- Enough computing power (AWS GPU Instance)
- Excels in tasks where the basic unit (pixel, word) has very little meaning in itself, but the combination of such units has a useful meaning

Outlook

Significant advances in deep reinforcement and unsupervised learning

Bigger and more complex architectures based on various interchangeable modules/techniques

Deeper models that can learn from much fewer training cases and Harder problems such as video understanding and natural language processing will be successfully tackled by deep learning algorithms.

Conclusion

- Deep Learning solutions are very powerful
    - State of the art in several problems
    - Still room for improvement
    - Still young solutions (hype)
- However
    - They are complex to implement
    - Free variables need to be configured with care
    - Results from paper are hard to reproduce
    - Heavy to train

Thank You!

Email:

h_Farahani@sbu.ac.ir