Introduction to Neural Networks

Alexander Trofimov
PhD, professor, NRNU MEPHI

lab@neuroinfo.ru
http://datalearning.ru

Course “Neural Networks”

March 2020

Artificial Neural Network

Neural networks are mathematical models that use learning algorithms inspired by the brain to process information

Since neural networks are used in computers, they are collectively called as "artificial neural networks" (ANNs)

Keywords:

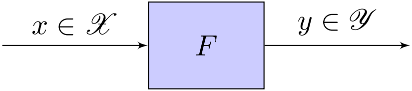
- Mathematical models
  ANNs are described with mathematical symbols and formulas
- Learning algorithms
  ANNs automatically learn to recognize complex mappings or patterns and make intelligent decisions based on data
- Inspired by the brain
  Similar to the brain, neural networks are built up of many neurons with many connections between them
- To process information
  ANNs receive input data, process them and output the results

ANN as a Computational Network

The ANNs are computational networks, they are used to process information

Usually ANN has inputs and outputs and can be viewed as a mapping F from input domain $\mathscr{X}$ to output domain $\mathscr{Y}$:

$$F : \mathscr{X} \to \mathscr{Y}$$



Input domain $\mathscr{X}$ and output domain $\mathscr{Y}$ are usually some multidimensional real coordinate spaces ($\mathbb{R}^M$ and $\mathbb{R}^K$ respectively)

The neural networks theory is based on mathematical analysis, linear algebra, probability theory and mathematical statistics

ANNs are Inspired by the Brain

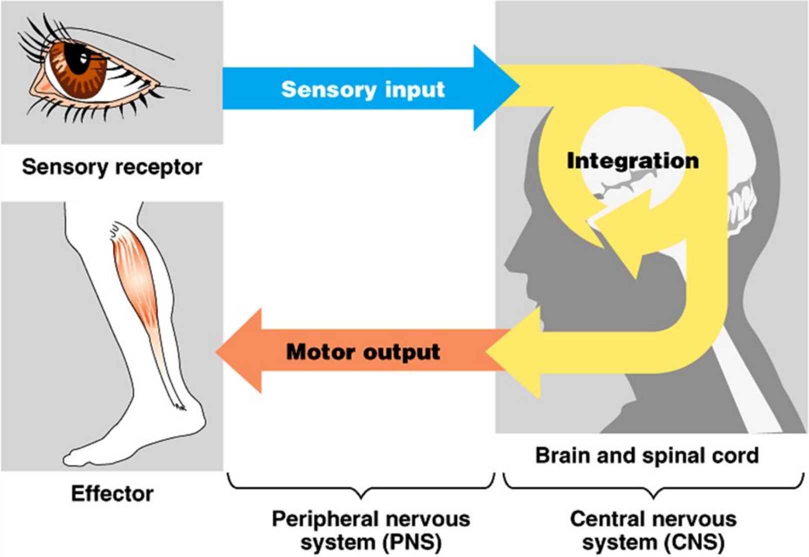
The entire information processing system of the vertebrate is the nervous system that consists of the central nervous system (CNS) and the peripheral nervous system (PNS)

PNS comprises the nerves that are situated outside of the brain or the spinal cord. These nerves form a branched and very dense network throughout the whole body

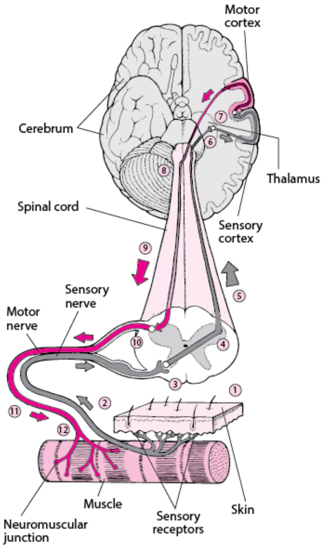
PNS consists of the sensory division that transmits information from both inside and outside of the body to the CNS y the motor division that receives signals from the CNS that cause an action to occur

CNS is the place where information received by the sense organs are stored and managed. Furthermore, it controls the inner processes in the body and coordinates the motor functions of the organism. The CNS consists of the brain and the spinal cord

Vertebrate Nervous System. Illustration



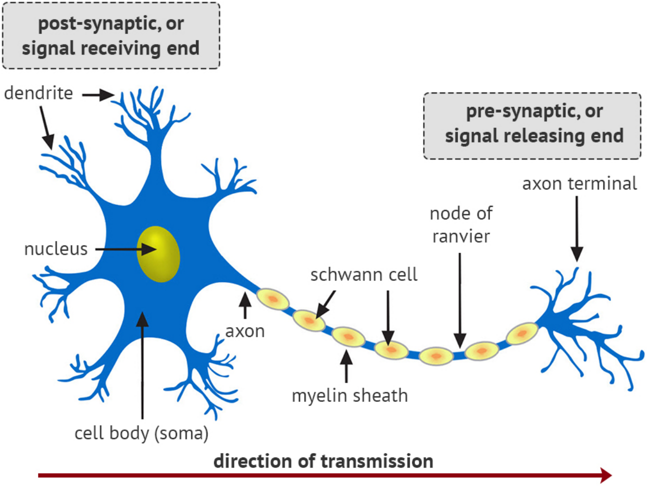
Using the Brain to Move a Muscle



1. If sensory receptors in the skin detect pain or a change in temperature, they transmit an impulse (signal), which ultimately reaches the brain.
2. The impulse travels along a sensory nerve to the spinal cord.
3. The impulse crosses a synapse (the junction between two nerve cells) between the sensory nerve and a nerve cell in the spinal cord.
4. The impulse crosses from the nerve cell in the spinal cord to the opposite side of the spinal cord.
5. The impulse is sent up the spinal cord and through the brain stem to the thalamus, which is a processing center for sensory information, located deep in the brain.
6. The impulse crosses a synapse in the thalamus to nerve fibers that carry the impulse to the sensory cortex of the cerebrum (the area that receives and interprets information from sensory receptors).
7. The sensory cortex perceives the impulse. A person may then decide to initiate movement, which triggers the motor cortex to generate an impulse.
8. The nerve carrying the impulse crosses to the opposite side at the base of the brain.
9. The impulse is sent down the spinal cord.
10. The impulse crosses a synapse between the nerve fibers in the spinal cord and a motor nerve, which is located in the spinal cord.
11. The impulse travels out of the spinal cord along the length of the motor nerve.
12. At the neuromuscular junction (where nerves connect to muscles), the impulse crosses from the motor nerve to receptors on the motor end plate of the muscle, where the impulse stimulates the muscle to move.

The Structure of a Neuron

The nervous system contains billions of cells referred to as neurons



Synapse

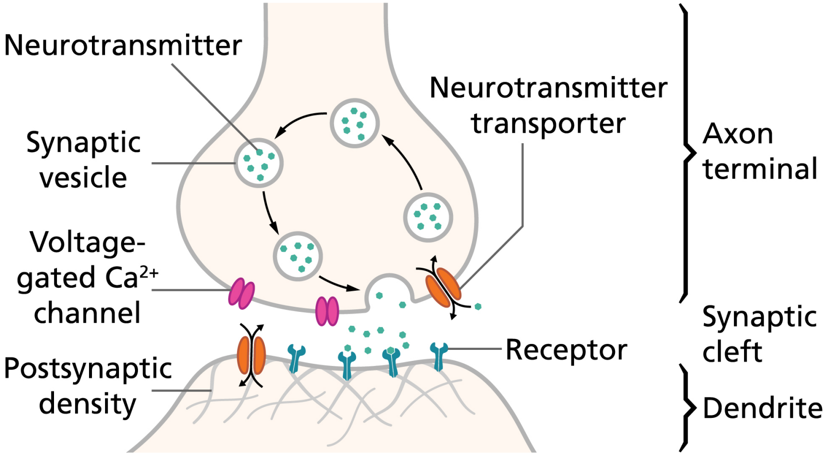
Neurons communicate with each other across unique biological structures called synapses

Synapse formed by:

- Presynaptic nerve terminal
  It releases chemical neurotransmitter in response to the arrival of an electrical action potential from the axon
- Postsynaptic cell
  It's in which neurotransmitter receptors, signaling molecules, and ion channels are concentrated
- Glial cells
  They participate in the metabolism of neurotransmitters and maintain ion homeostasis

Information is transmitted through the synapse in one direction only, from axon terminal to dendrite receptors

Synapse. Illustration



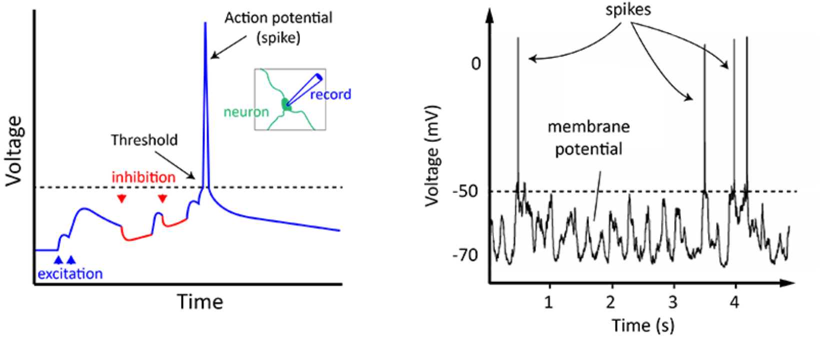
Action Potential

Normally, the inside of the neuron is more negative electrically (approx. -70 mV) than the outside (because of more positive ions are outside), neuroscientists say that the neuron's resting membrane potential is -70 mV

The membrane potential constantly going up and down, depending mostly on the inputs coming from the axons of other neurons. Some inputs make the neuron's membrane potential become more positive (or less negative, e.g. from -70 mV to -65 mV), and others do the opposite. It depends on the type of neurotransmitter used

When the total sum of all of the excitatory and inhibitory inputs makes the neuron's membrane potential reach a threshold referred to as the action potential threshold (approx. -50 mV), it generates the action potential (up to +40 mV) also called a spike

Action Potential. Illustration



A neuron spikes when a combination of all the excitation and inhibition it receives makes it reach action potential threshold

Spike is a brief (approx. 1 ms) electrical event followed by a refractory period

Nervous Transmission

Synaptic transmission:
Synapses can be thought of as converting an electrical signal (the action potential) into a chemical signal in the form of neurotransmitter release, and then, upon binding of the transmitter to the postsynaptic receptor, switching the signal back again into an electrical form, as charged ions flow into or out of the postsynaptic neuron

The width of synaptic cleft is 20-40 nm

Neural transmission:
The generation and propagation of action potential across the neuron is a complex electrochemical process including openings and closings of membrane ion channels, exchanging the Na^+, Ca^+, K^+ ions with the outside and saltatory conduction through Schwann cells

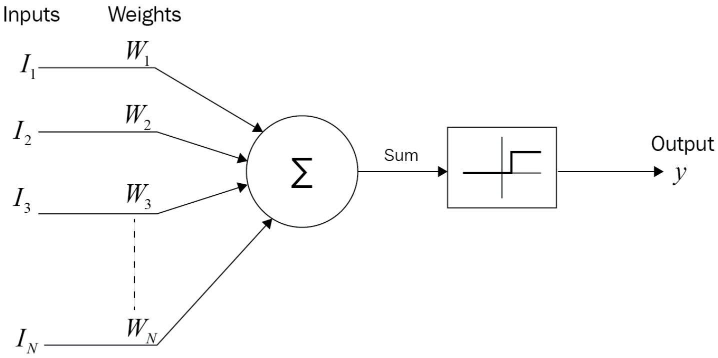
Video on neural transmission on YouTube

Technical View on Nervous Transmission

- The neuron can have thousands of dendrites and synapses, they receive electrical signals from other neurons and transfer them into the soma of the cell
  Technical view: the neuron receives multiple inputs
- The soma accumulates the signals that can be activating or inhibiting signals
  Technical view: the soma weights the neuron inputs
- As soon as the accumulated signal exceeds the action potential threshold a spike is generated
  Technical view: if the weighted sum exceeds a threshold the neuron's output should be active (e.g. equal to 1)
- As soon as spike achieves axon terminal, it can influence dendrites of adjacent neurons
  Technical view: neuron's output can be the input for other neurons in neural network

McCulloch-Pitts Neuron

The artificial neuron was proposed by Warren McCulloch and Walter Pitts in 1943*. The model was specifically targeted as a computational model of the "nerve net" in the brain



*W.McCulloch, W.Pitts. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics. 1943, V.5(4), Pp.115–133.

History of Artificial Neural Networks: Golden Age

1943. W.McCulloch and W.Pitts introduced artificial neural networks as the models of natural nerve nets

1947. W.McCulloch and W.Pitts indicated a practical field of ANNs applications namely the recognition of patterns

1949. Donald Hebb formulated the rule which represents in its more generalized form the basis of nearly all neural learning procedures named Hebbian rule

1957–1958. Frank Rosenblatt, Charles Wightman and their co-workers at the MIT developed the first successful neurocomputer, the Mark 1 perceptron, which was capable to recognize simple numerics by means of a 20*20 pixels image sensor and electromechanically worked with 512 motor driven potentiometers. Each potentiometer representing one variable weight

History of Artificial Neural Networks: Golden Age

1960. Bernard Widrow and Ted Hoff introduced the ADALINE (ADAptive LInear NEuron), a fast and precise adaptive learning model being the first widely commercially used neural network. It could be found in nearly every analog telephone for real-time adaptive echo filtering and was trained by menas of the Widrow-Hoff rule or delta rule

1969. Marvin Minsky and Seymour Papert published a precise mathematical analysis of the Rosenblatt's perceptron to show that the perceptron model was not capable to learn an XOR function (and other linearly non-separable functions), so called XOR-problem. Often-miscited Minsky/Papert text put an end to overestimation and popularity of perceptron and caused a significant decline in interest and funding of neural network research

*M.Minsky, S.Papert. Perceptrons: An Introduction to Computational Geometry. 1969. MIT Press.

History of Artificial Neural Networks: Long Silence

1974. Paul Werbos in his dissertation in Harvard developed a learning procedure called backpropagation of error

1976–1980. Stephen Grossberg presented many papers in which numerous neural models are analyzed mathematically. Furthermore, he dedicated himself to the problem of keeping a neural network capable of learning without destroying already learned associations. Under cooperation of Gail Carpenter this led to models of adaptive resonance theory (ART)

1982. Teuvo Kohonen described the self-organizing feature maps (SOM) also known as Kohonen maps. He was looking for the mechanisms involving self-organization in the brain

1983. Fukushima, Miyake and Ito introduced the neural model of the Neocognitron which could recognize handwritten characters

History of Artificial Neural Networks: Renaissance

1985. John Hopfield published an article describing a way of finding acceptable solutions for the Travelling salesman problem by using Hopfield nets

1986. D.Rumelhart, G.Hinton and R.Williams rediscovered the backpropagation algorithm and finally succeeded in making it widely known. Non-linearly separable problems could be solved by multilayer perceptrons, and negative evaluations around Marvin Minsky's text were discredited at a single blow

1989. Yann LeCun et al. used back-propagation to learn the convolution kernel coefficients from images of hand-written digits

1997. J.Schmidhuber and S.Hochreiter proposed a sophisticated type of recurrent neural networks named Long short-term memory (LSTM) network that overcame the vanishing gradient problem

1998. Yann LeCun et al. introduced a pioneering convolutional network (CNN) LeNet-5 to recognize hand-written digits

History of Artificial Neural Networks: Going Deeper

2006. Geoffrey Hinton presented Deep Belief Networks and one of the first effective deep learning algorithms based on Restricted Bolzman machines and autoencoders

2009. As deep learning models require a tremendous amount of labelled data to train themselves in supervised learning, Fei-Fei Li launched a large database of labelled images ImageNet

2012. Alex Krizhevsky proposed Convolutional neural network named Alexnet in ILSVRC ImageNet competition that demonstrated best results

2014–2015. Convolutional networks GoogLeNet (Google), VGGNet (Visual Geometry Group, Oxford), ResNet (Microsoft) were developed and have shown the best results in ILSVRC ImageNet competitions

2015–. Development of deep neural networks in wide range of practical applications

Why Neural Networks?

The ANNs are computational (data flow) networks, they are used to process data

- **Capability to learn**
  Neural networks are universal function approximators. If a problem can be described by a mathematical function (no matter how complex), we can use a neural network to represent it
- **Capability to generalize data**
  After successful training a neural network can find reasonable solutions for similar problems of the same class that were not explicitly trained
- **Robustness**
  Fault tolerance against noisy input data and network's parameters

Approaches to Modelling

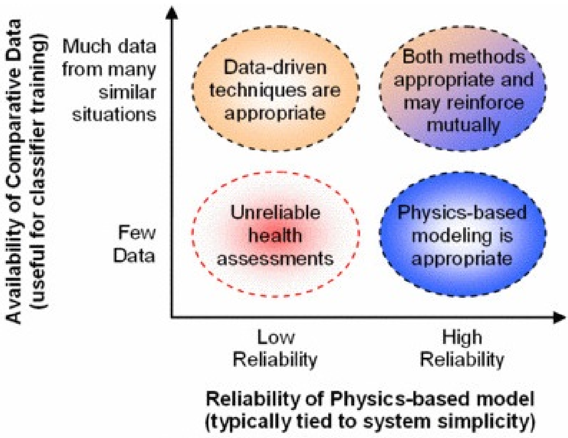
Approaches to modelling:

- Model-based
  Use known physical, economical, biological, etc. laws and expert rules to construct the model
- Data-driven
  Use data to construct the model

Key features of data-driven approach:

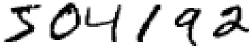
- The precise mathematical model is absent or unacceptably complex
- There are statistical "input-output" data about the system under modelling
- The data processing algorithm is unknown a priori, it is a result of learning procedure

Data-driven vs Model-based Approaches



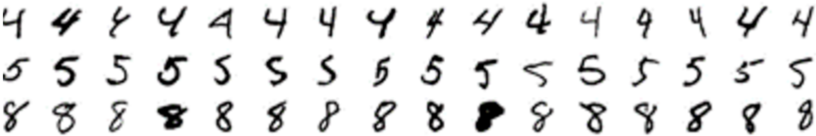
The model-based approach is generally more robust in the sense that it can deal more easily with new or unforeseen situations

Example: Handwritten Digits Recognition

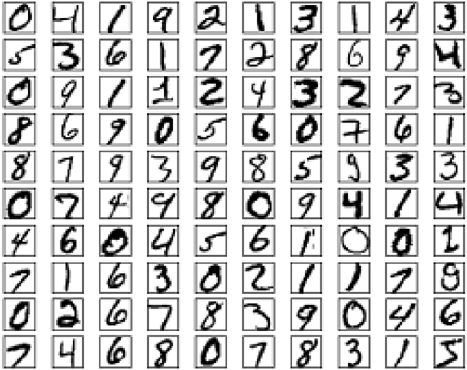


The human brain easily recognize the digits. The difficulty of visual pattern recognition becomes apparent if you attempt to write a computer program to recognize them

Simple intuitions about how we recognize shapes – "a 9 has a loop at the top, and a vertical stroke in the bottom right" – turn out to be not so simple to express algorithmically. When you try to make such rules precise, you quickly get lost in a morass of exceptions and special cases. It seems hopeless



Neural Network Approach to Handwritten Digits Recognition



If we take a large number of handwritten digits (training examples), we develop a system which can learn from those training examples

The neural network uses the examples to automatically infer rules for recognizing handwritten digits

Knowledge-Based and Machine Learning Approaches in AI

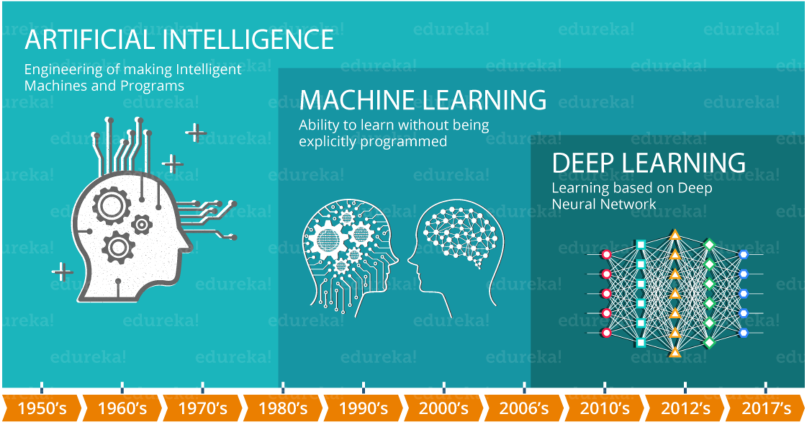
In the early days, the artificial intelligence (AI) tackled and solved problems that are intellectually difficult for humans but relatively straightforward for computers – problems that can be described by a list of formal, mathematical rules

But the true challenge to AI resulted to be solving the tasks that are easy for people to perform but hard for people to describe formally – problems that we solve intuitively, like recognizing spoken words or faces in images

Up to 2000s the knowledge-based approach to AI was popular. It based on hard-coding the knowledge about the world in formal languages and computer can reason automatically using logical inference rules

In 2000s took place the machine learning approach that avoids the need for human operators to formally specify all the knowledge that the computer needs

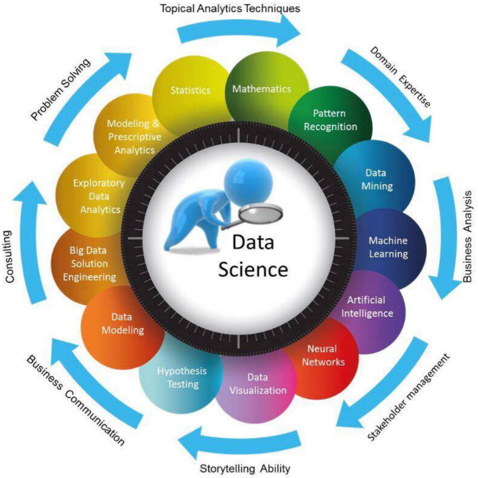
AI, ML and ANN



For now artificial intelligence is mostly related to machine learning and neural networks

Neural Networks and Data Science

Neural networks deal with the data and therefore the theory of neural networks is highly related to data science



Neural Networks and Machine Learning

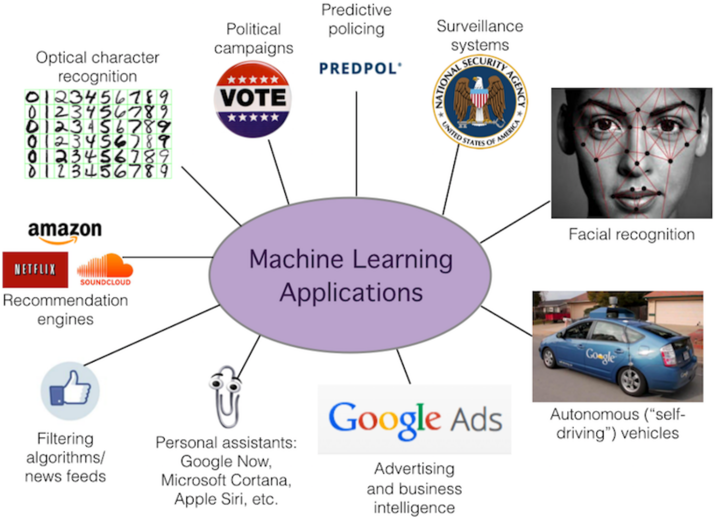
Neural networks are used in almost every branch of machine learning:

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Ensemble learning
- Reinforcement learning
- Deep learning
- Transfer learning
- ...

Types of Problems Solved by Neural Networks

- **Supervised learning**
  - Regression
  - Classification
- **Unsupervised learning**
  - Clustering
  - Anomaly detection and novelty detection
  - Visualization and dimensionality reduction
- **Reinforcement learning**
  - Approximation of state/action values
- **Optimization**
  - Approximate solution of some combinatorial problems (e.g. salesman problem)

Neural Networks Applications

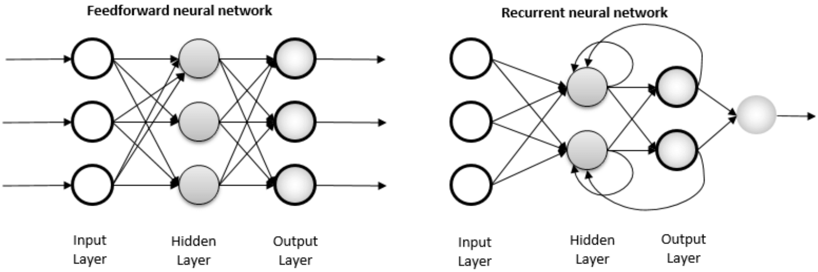


Neural Networks Milestones

1997 IBM's Deep Blue beats the world champion at chess
2005 Autonomous ground vehicles: DARPA Grand Challenge
2006 Google Translate
2011 DARPA CALO project, Apple Siri
2011 IBM's Watson beats two human champions in a Jeopardy! competition
2012 The Google Brain team create a neural network that learns to recognise cats by watching unlabeled images taken from frames of YouTube videos
2014 Facebook DeepFace identifies faces with 97% accuracy
2015 OpenAI by Elon Musk and Sam Altman - $1 bln.
2016 OpenAI, Google's DeepMind: Atari games
2016 Google's AlphaGo beats the world champion at Go
2018 Tesla launches self-driving vehicle

Neural Network Architectures

- Feed-forward neural networks
  The data flow propagates in one direction, from inputs to outputs
- Dynamic (recurrent) neural networks
  The network is considered as a dynamic system, the concept of time is needed



Types of Neural Networks Architectures

By depth:

- Shallow neural networks
  Multilayer perceptron, RBF network, Kohonen network, etc.
- Deep neural networks
  Convolutional neural networks, Deep belief networks, LSTM networks, etc.

By training manner:

- Trained in supervised manner
  Multilayer perceptron, convolutional neural networks, etc.
- Trained in unsupervised manner
  Self-organizing maps, autoencoders, etc.

By stochasticity:

- Determined
  Multilayer perceptron, convolutional neural networks, etc.
- Stochastic
  Stochastic Hopfield network, Restricted Boltzmann machine, etc.

Shallow Feed-Forward Neural Networks

- Multilayer neural network (multilayer perceptron)
- Radial basis function neural network (RBF network)
- Probabilistic neural network (PNN)
- Generalized regression neural network (GRNN)
- Learning Vector Quantization (LVQ) Neural Networks
- Linear Neural Networks
- Time-delay neural networks (TDNN)
- Kohonen network
- Counter-propagation network (CPN)
- ...

Shallow Recurrent Neural Networks

- Fully connected recurrent neural network (RNN)
- Elman network
- Jordan network
- Runge-Kutta neural network
- Recurrent multilayer perceptron
- Non-linear autoregressive network with exogenous inputs (NARX) neural network
- Hopfield network
- Boltzmann machine
- Continuous-time recurrent neural network (CTRNN)
- Spiking neural network (SNN)
- ...

Deep Neural Networks

- Convolutional neural network (CNN)
- Deep Belief Network (DBN)
- Long-short term memory networks (LSTM)
- Gated recurrent neural network (GRU)
- Generative Adversarial Network (GAN)
- Deep Convolutional Generative Adversarial Networks (DCGAN)
- Deep autoencoders
- ...