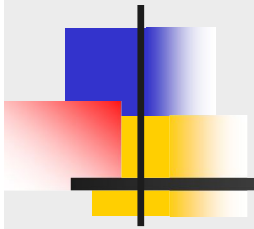


Neural Networks and Deep Learning

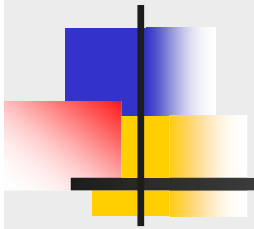
Presenter

Md. Ferdous

Lecturer

Dept. of Computer Science and Engineering (CSE)

Bangabandhu Sheikh Mujibur Rahman Science and Technology University



Slide Credits and heartful thanks to

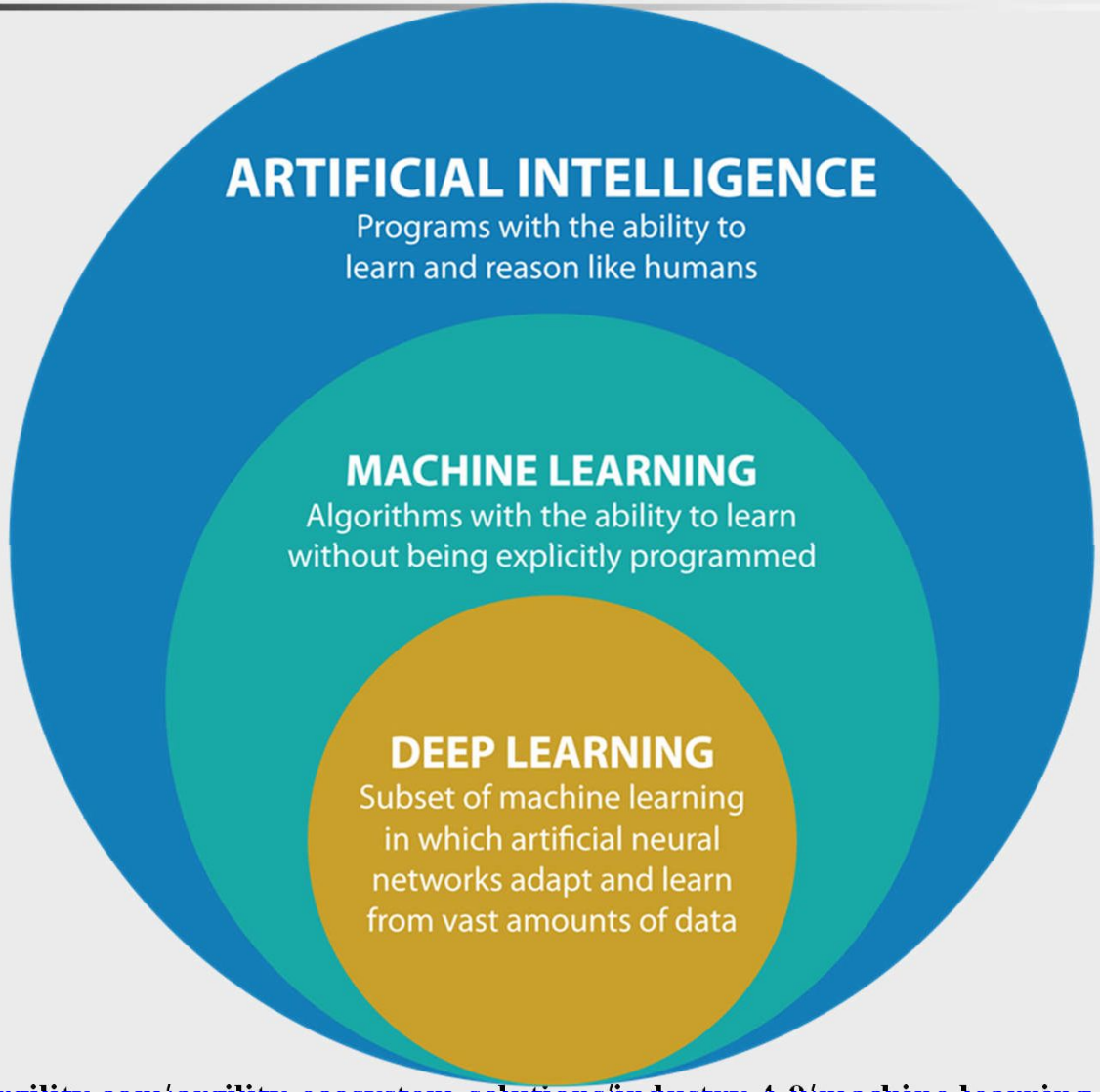
Dr. Md. Aminul Haque Akhand
Professor and Ex-head
Dept. of Computer Science and Engineering (CSE)
Khulna University of Engineering & Technology (KUET)
www.kuet.ac.bd/cse/akhand

“Yesterday I was
clever, so I wanted to change the world.

Today I am
wise,
so I am changing myself.”

— **Rumi**

AI ->Machine Learning->Deep Learning



https://www.argility.com/argility-ecosystem-solutions/industry-4-0/machine-learning-deep-learning/

AI ->Machine Learning->Deep Learning

Artificial Intelligence (AI)

AI is the broadest term, applying to any technique that enables computers to mimic human intelligence, using logic, if-then rules, decision trees, and machine learning (including deep learning).

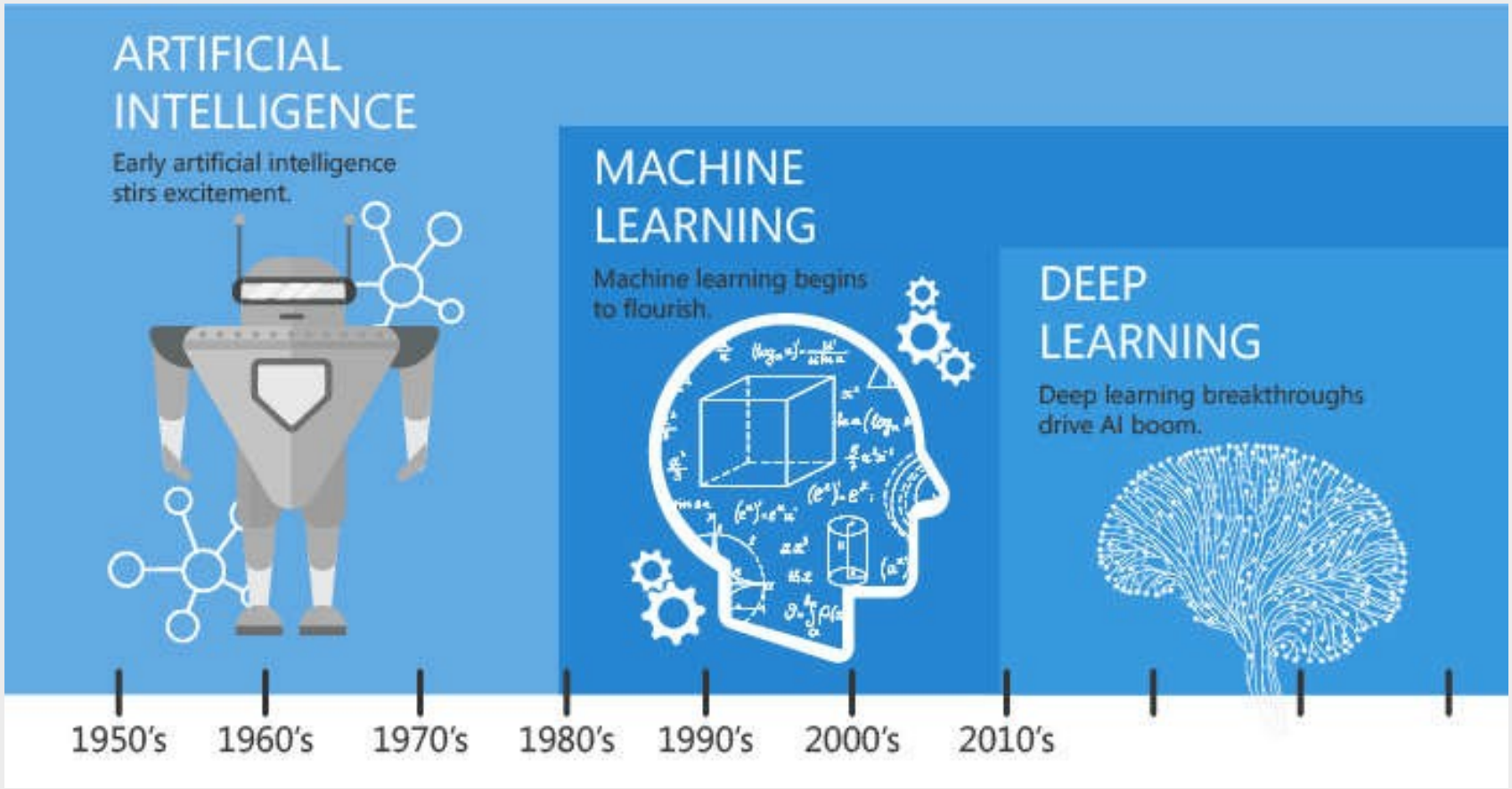
Machine Learning

The subset of AI that includes that enable machines to improve at tasks with experience. The category includes deep learning.

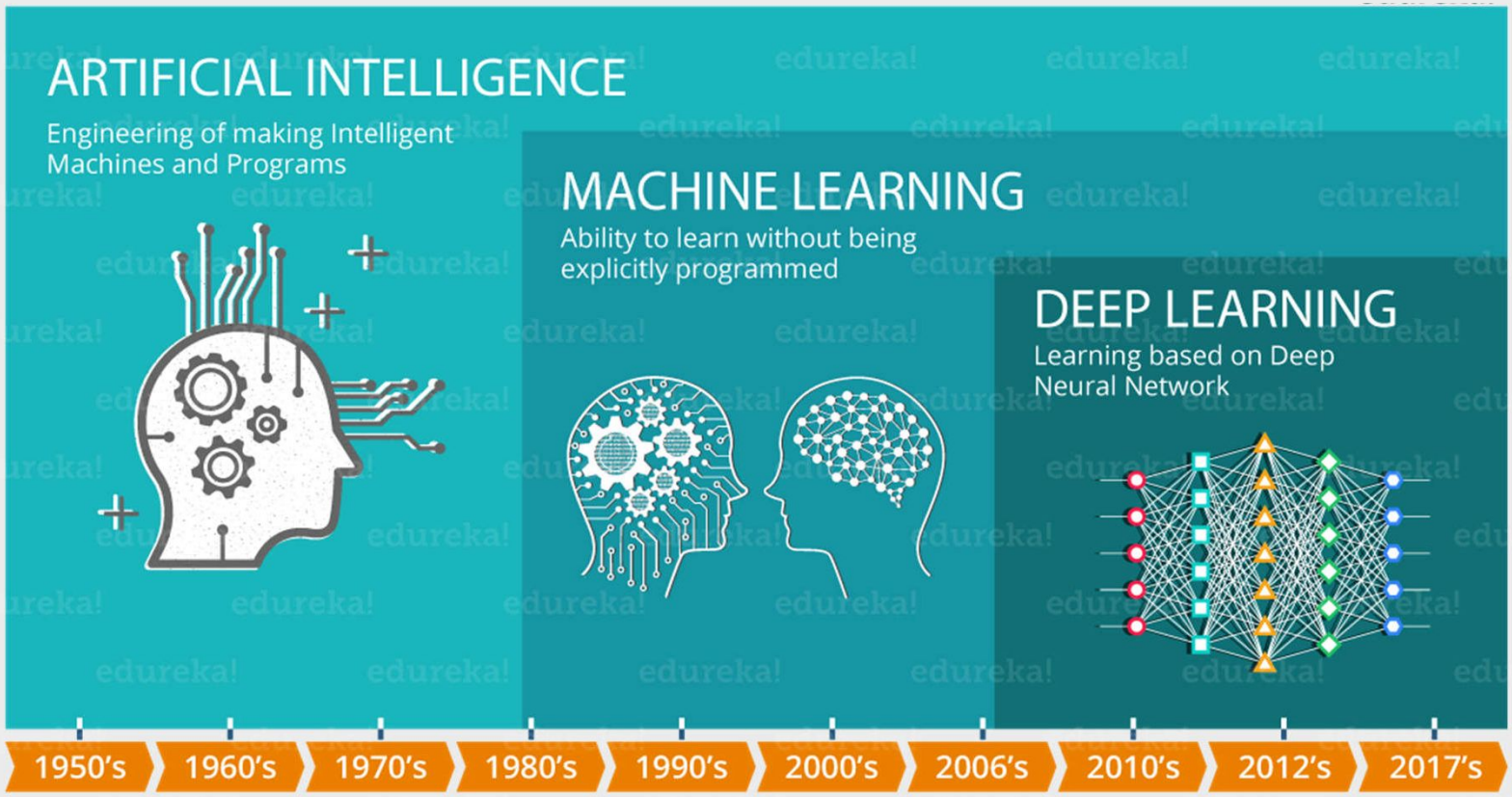
Deep Learning

The subset of machine learning composed of algorithms to train itself to perform tasks, like speech and image recognition, by exposing multilayered neural networks to vast amounts of data.

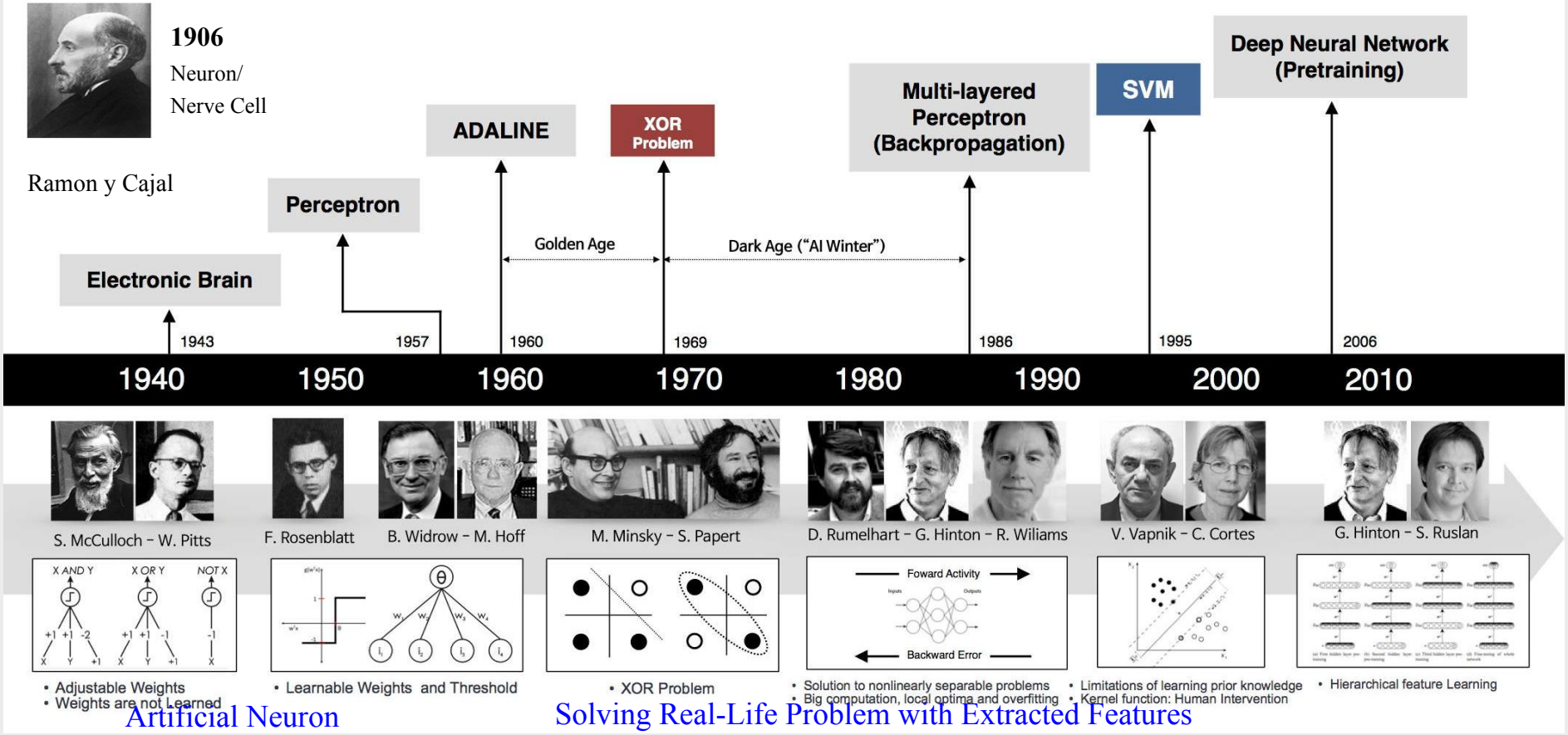
AI ->Machine Learning->Deep Learning



AI ->Machine Learning->Deep Learning



History and Background of Deep Learning



Artificial Neuron

Solving Real-Life Problem with Extracted Features

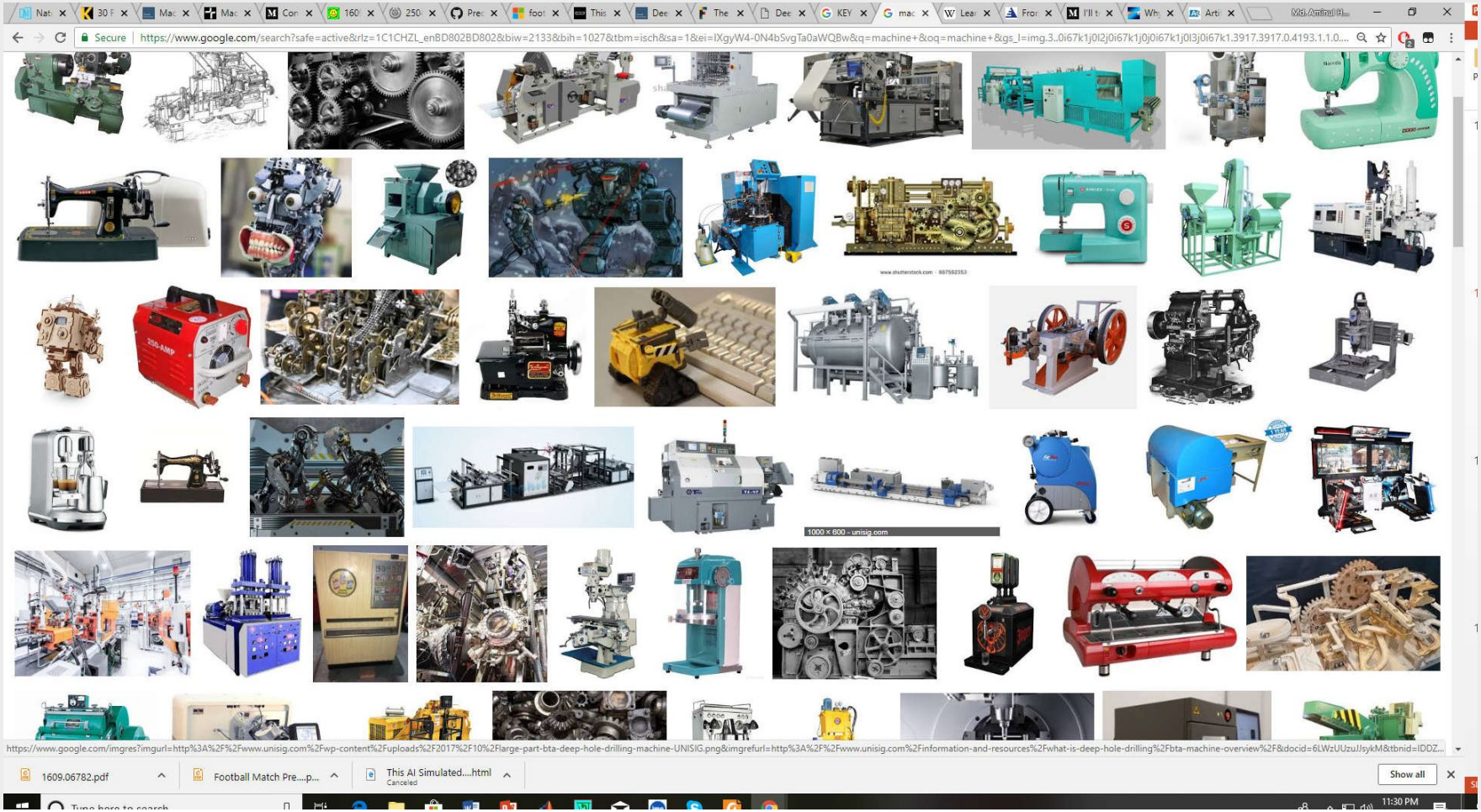
Development of Traditional Machine Learning

Deep Leaning

https://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html

Machine

A **machine** uses power to apply forces and control movement to perform an intended action. Machines include a system of mechanisms that shape the actuator input to achieve a specific application of output forces and movement.

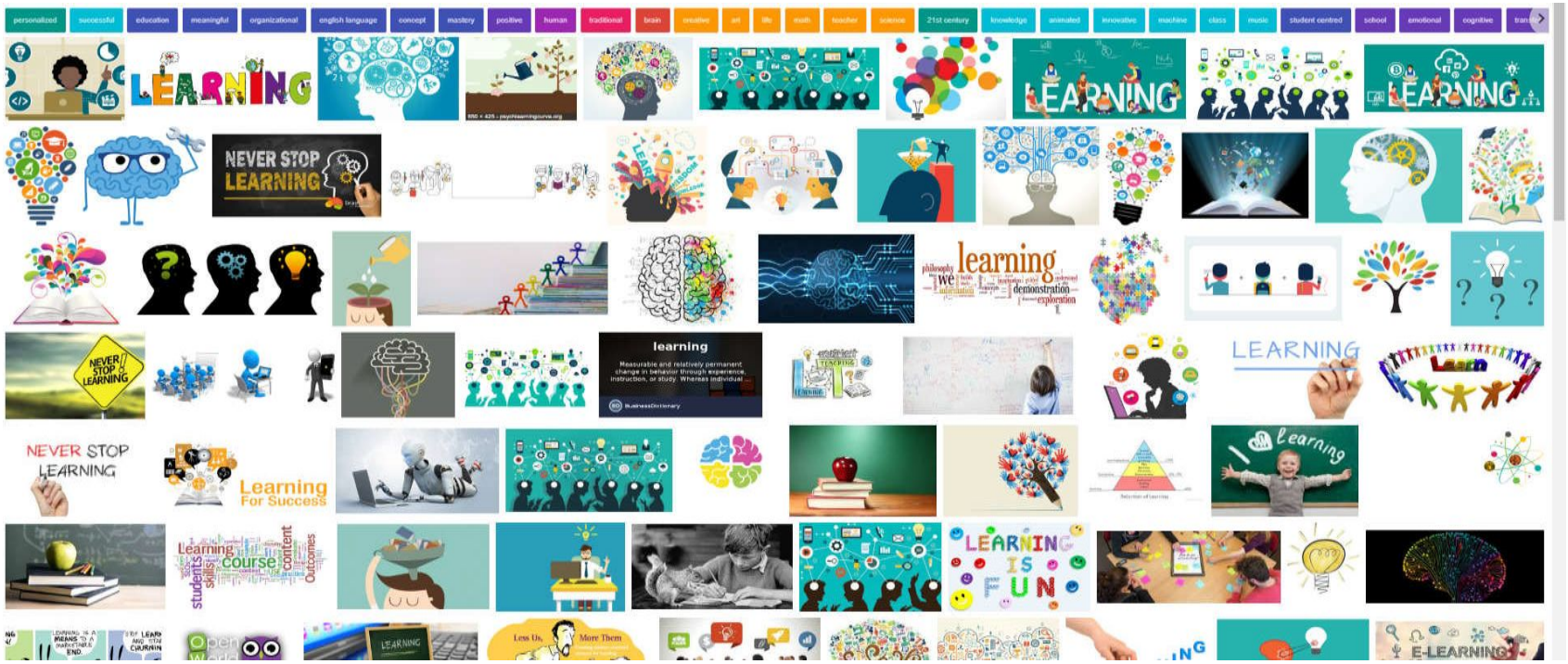


Learning (Human)

Learning is the process of acquiring new or modifying existing knowledge, behaviors, skills, values, or preferences

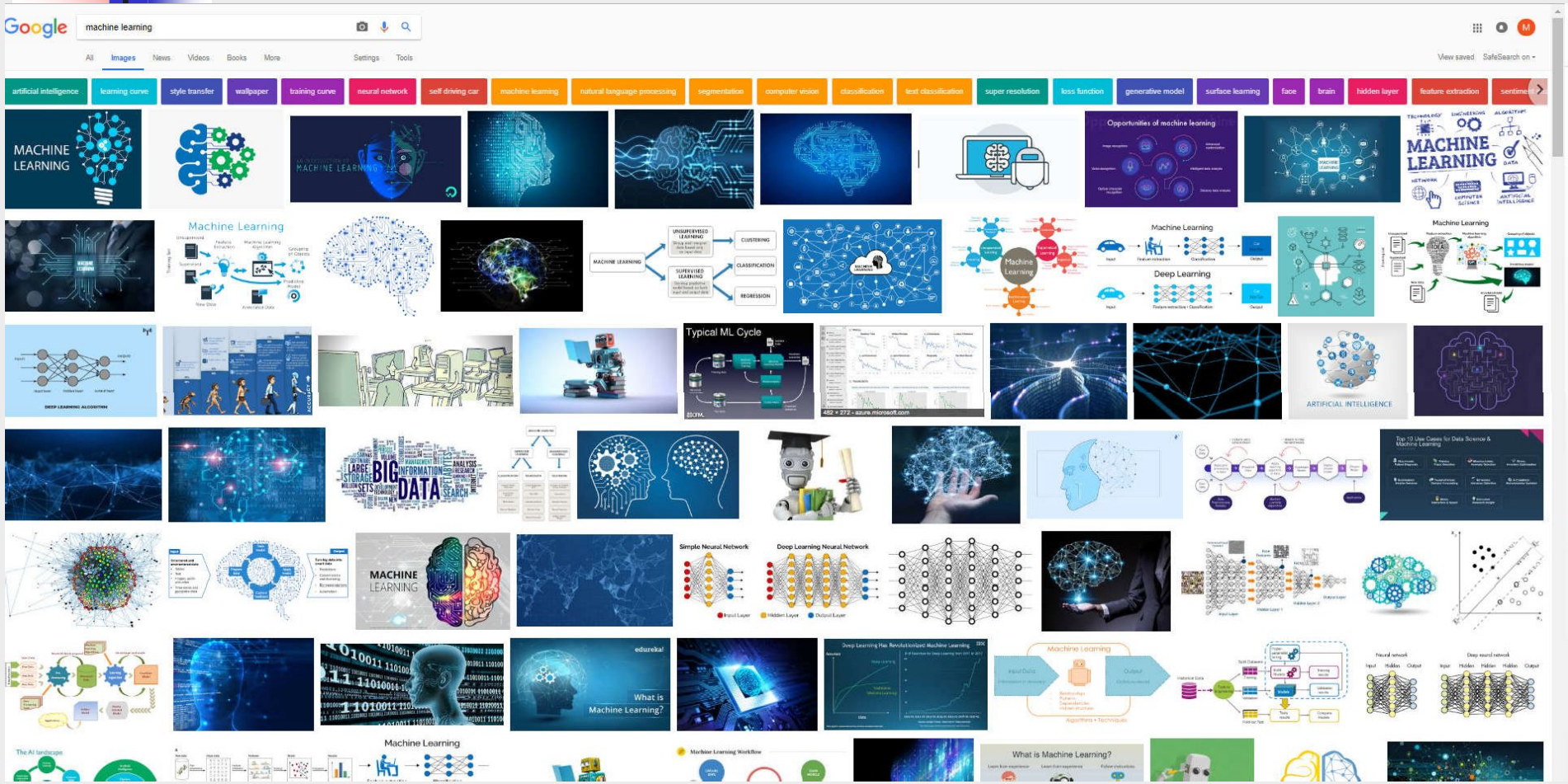
Another Definition

"A change in human disposition or capability that persists over a period of time and is not simply ascribable to processes of growth."



Human brain is the main element of learning

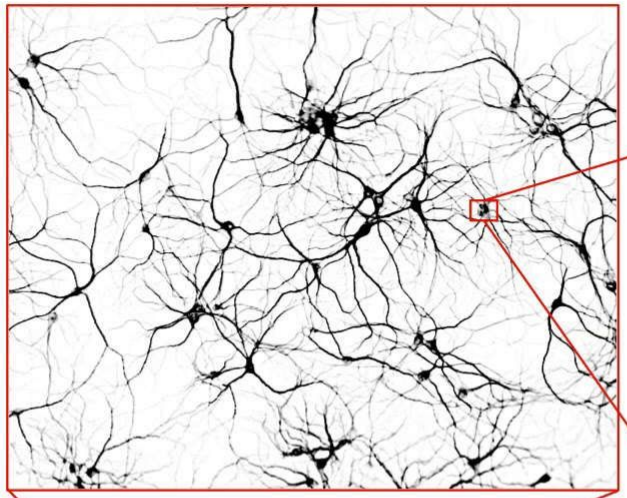
Machine Learning

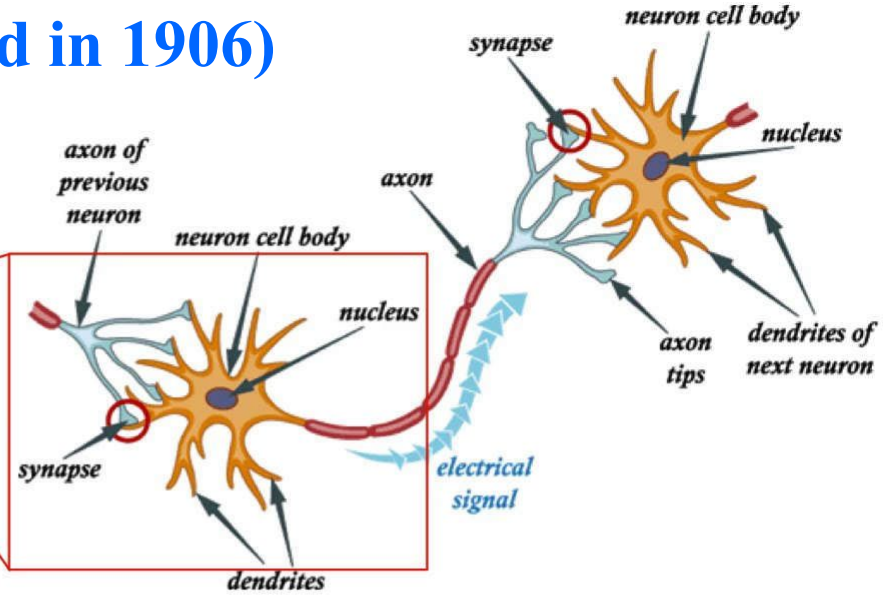


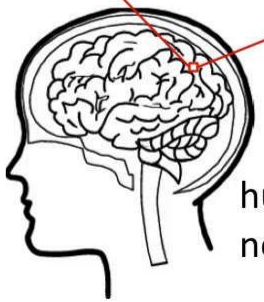
Technique to give computer brain like learning ability through progressively update with data, without being explicitly programmed.

Neuron : Learning Element in Brain

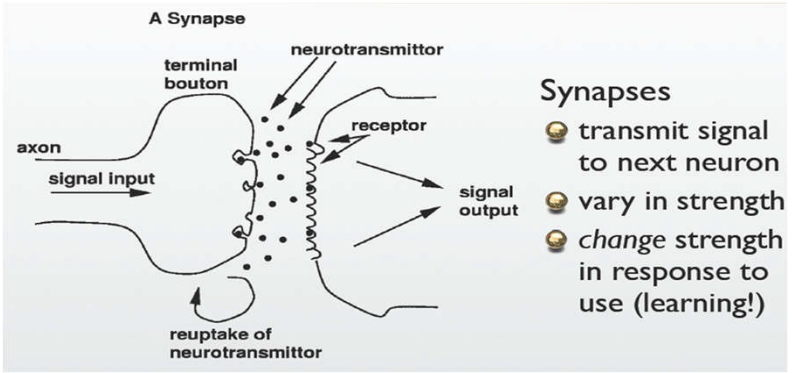
(Discovered in 1906)







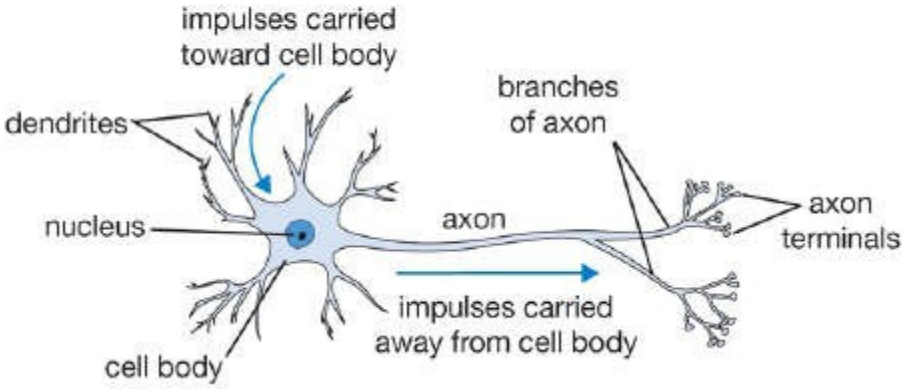
humans don't need features

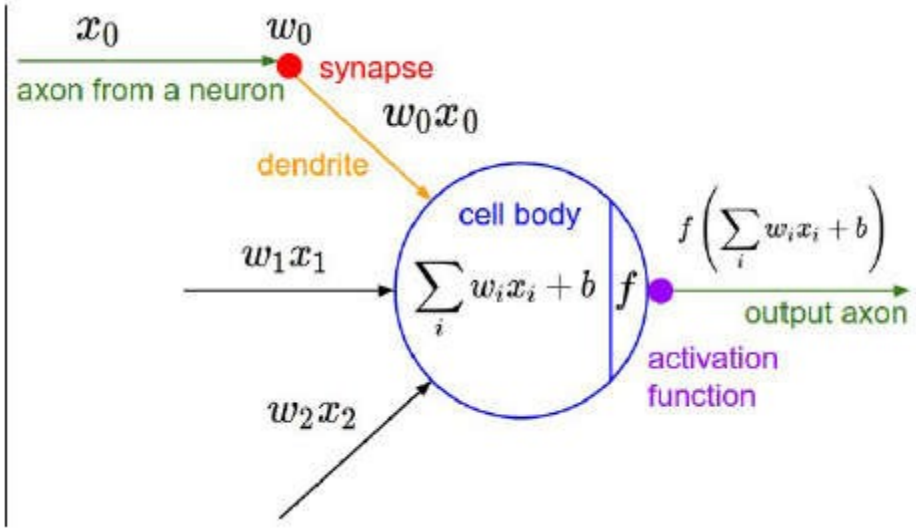


Synapses
- transmit signal to next neuron
- vary in strength
- *change* strength in response to use (learning!)

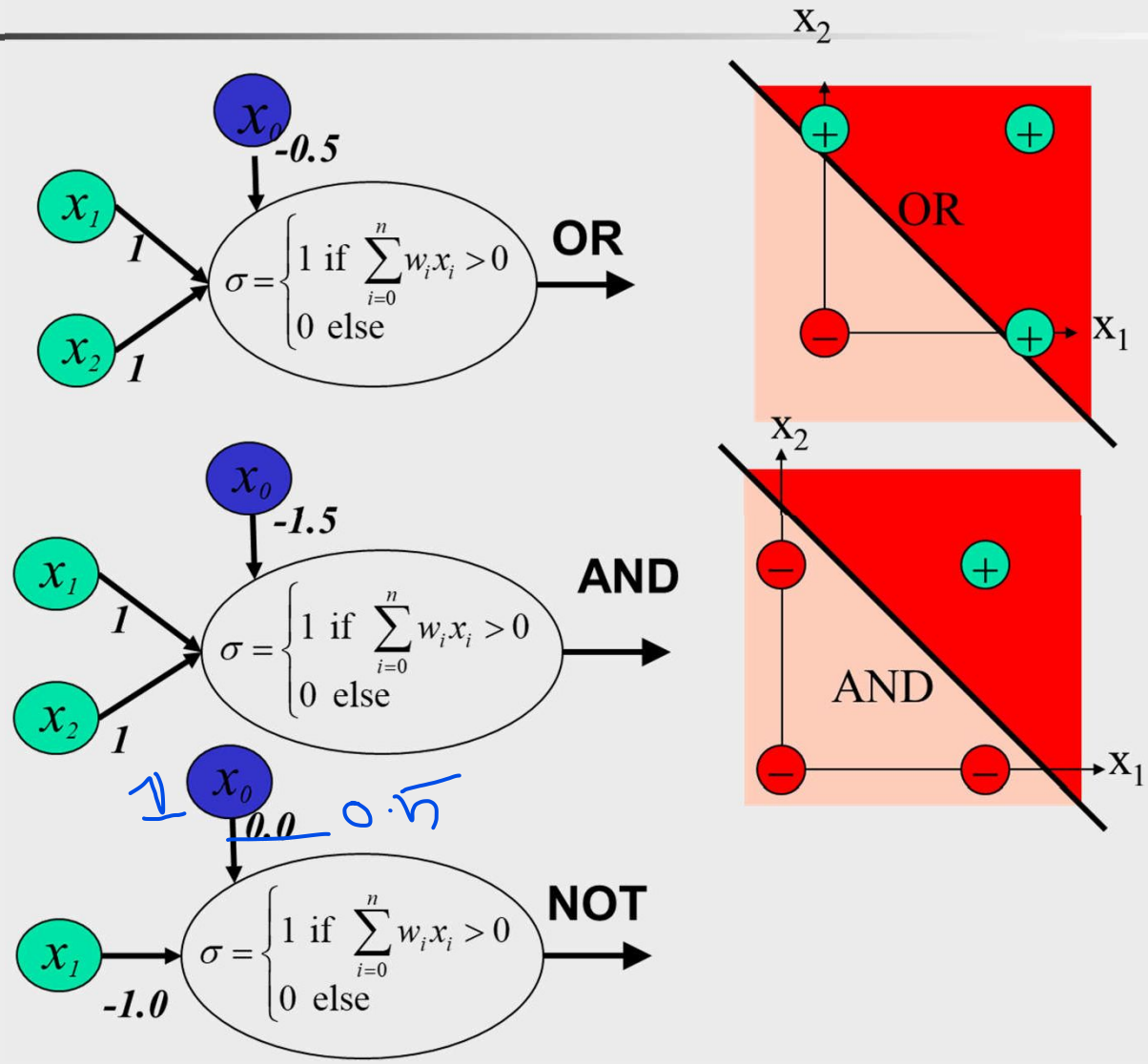
McCulloch-Pitts model of a Neuron

(Proposed in 1943)

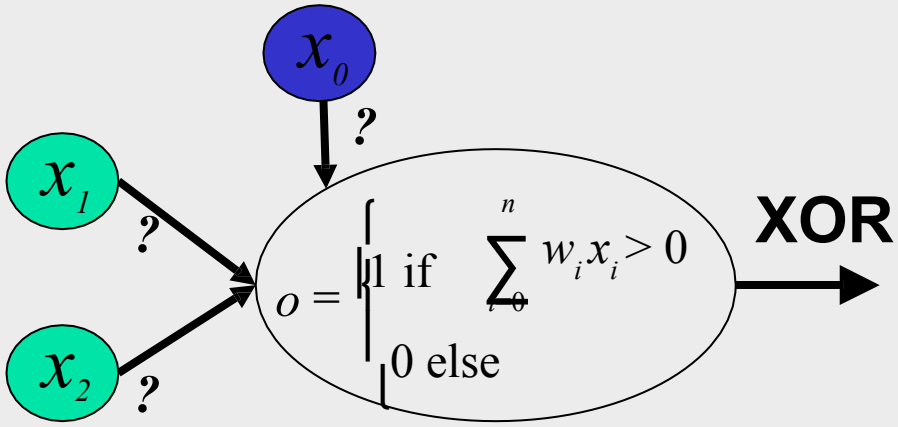


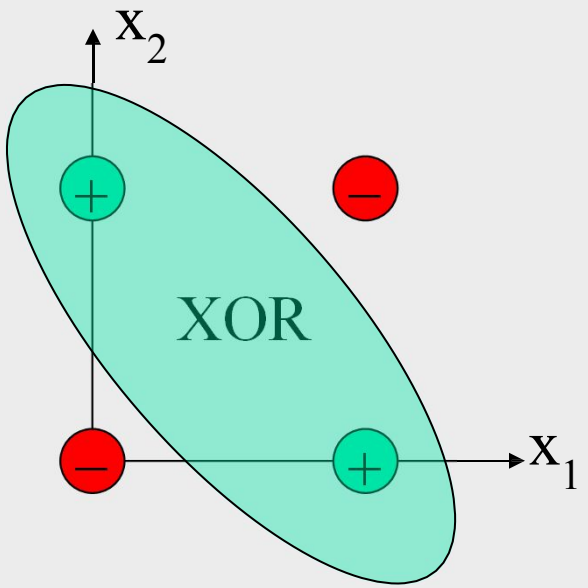


Ability of Single Neuron : Solving Linear Problem



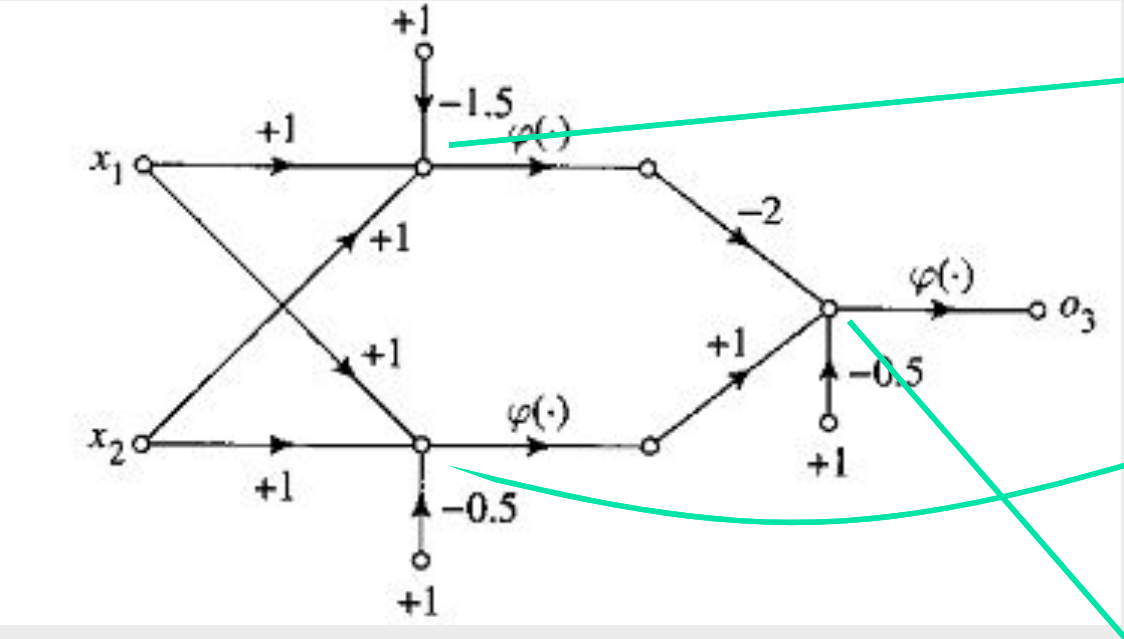
Ability of Single Neuron: Unable to Solve Nonlinear Problem

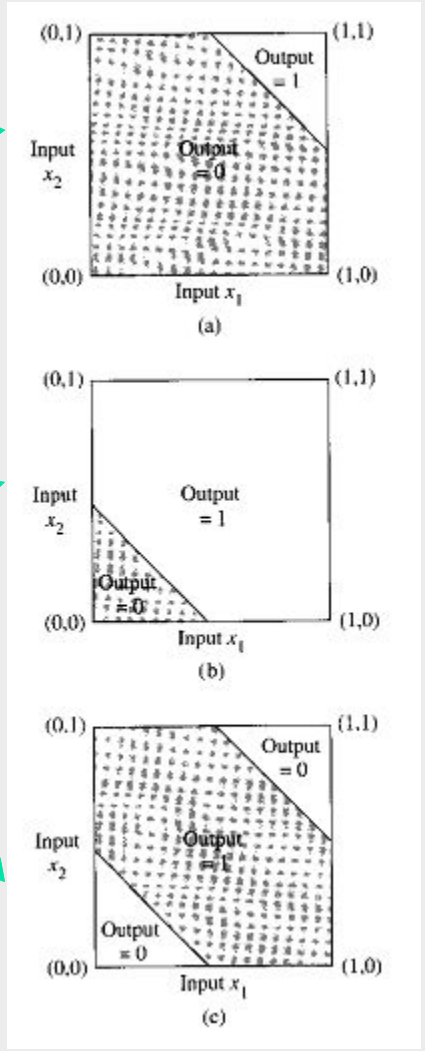




No arrangement work, not linearly separable. Require two boundary lines.

XOR solution with Hidden Neuron





AND

OR

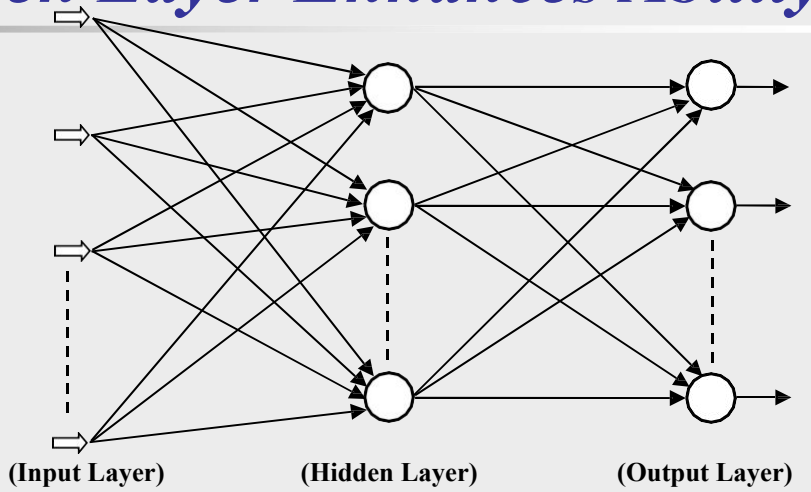
XOR

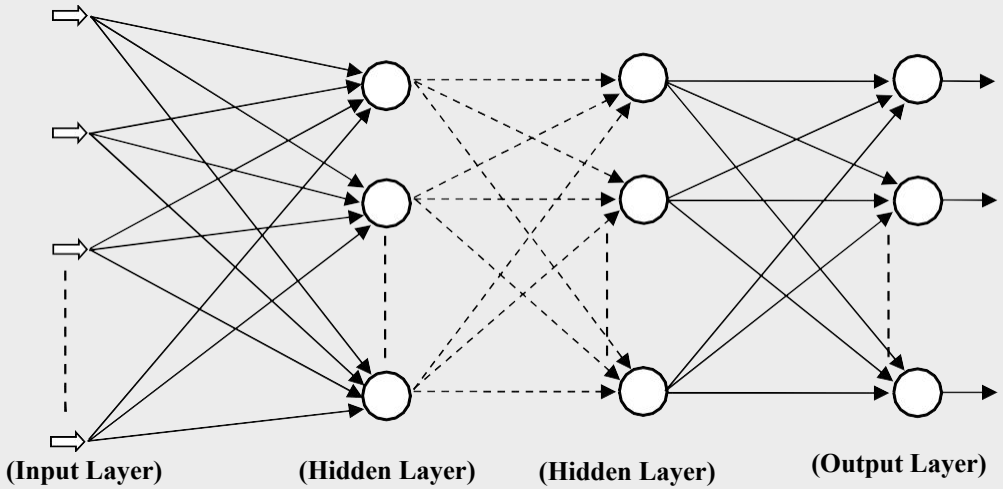
XOR logic through manipulation by hidden neurons

Open doors to Solve Real life Nonlinear Problem

Neural Network: Hidden Layer Enhances Ability

Hidden Layer(s) Enhances Ability to Solve Real Life Nonlinear Problems



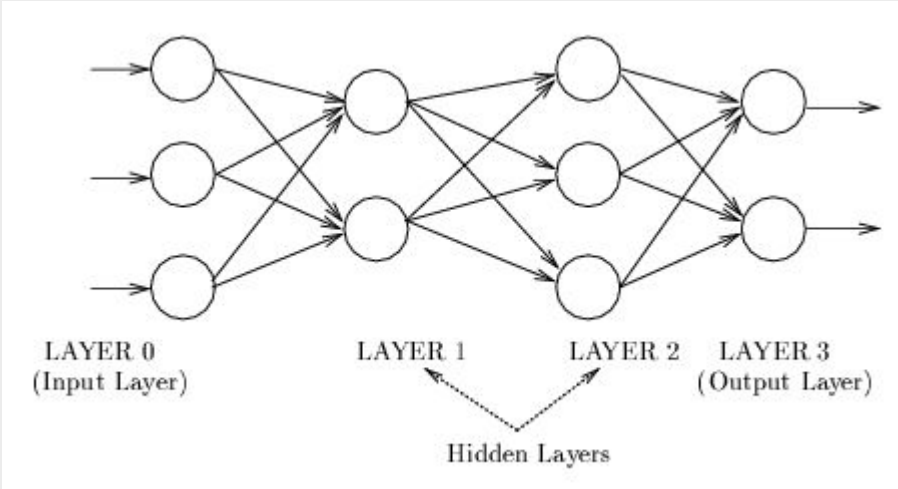


How to Select Appropriate Weight ? Learning?

How to get appropriate weight set?

Back-propagation is the famous algorithm for training feed-forward networks

(Proposed in 1986)



http://en.wikipedia.org/wiki/Backpropagation

Backpropagation, or propagation of error, is a common method of teaching artificial neural networks how to perform a given task.
It was first described by Paul Werbos in 1974, but it wasn't until 1986, through the work of David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams, that it gained recognition, and it led to a **"renaissance"** in the field of artificial neural network research.

letters to nature

Nature **323**, 533 - 536 (09 October 1986); doi:10.1038/323533a0

Learning representations by back-propagating errors

DAVID E. RUMELHART*, GEOFFREY E. HINTON† & RONALD J. WILLIAMS*

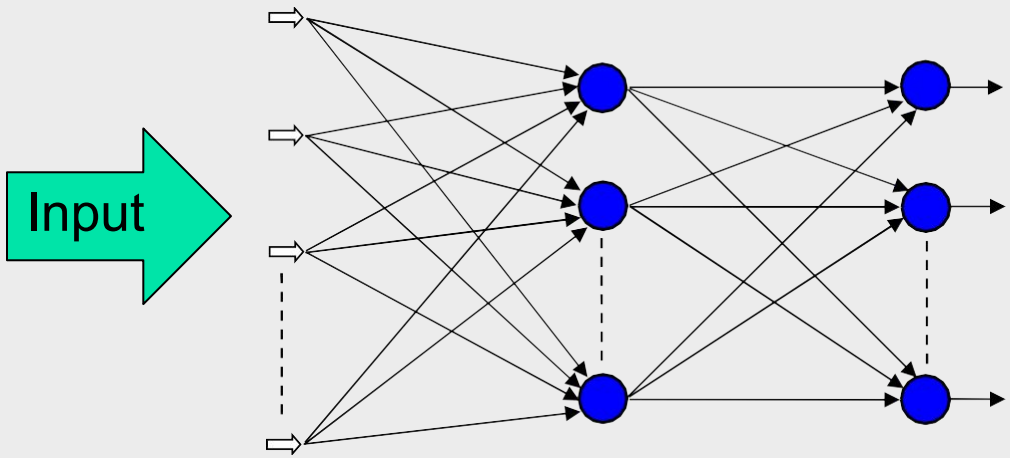
*Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA
†Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA
†To whom correspondence should be addressed.

We describe a new learning procedure, back-propagation, for networks of neurone-like units. The procedure repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector. As a result of the weight adjustments, internal 'hidden' units which are not part of the input or output come to represent important features of the task domain, and the regularities in the task are captured by the interactions of these units. The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure[1].

References

1. Rosenblatt, F. *Principles of Neurodynamics* (Spartan, Washington, DC, 1961).
2. Minsky, M. L. & Papert, S. *Perceptrons* (MIT, Cambridge, 1969).
3. Le Cun, Y. *Proc. Cognitiva* **85**, 599–604 (1985).

Back-propagation(BP)

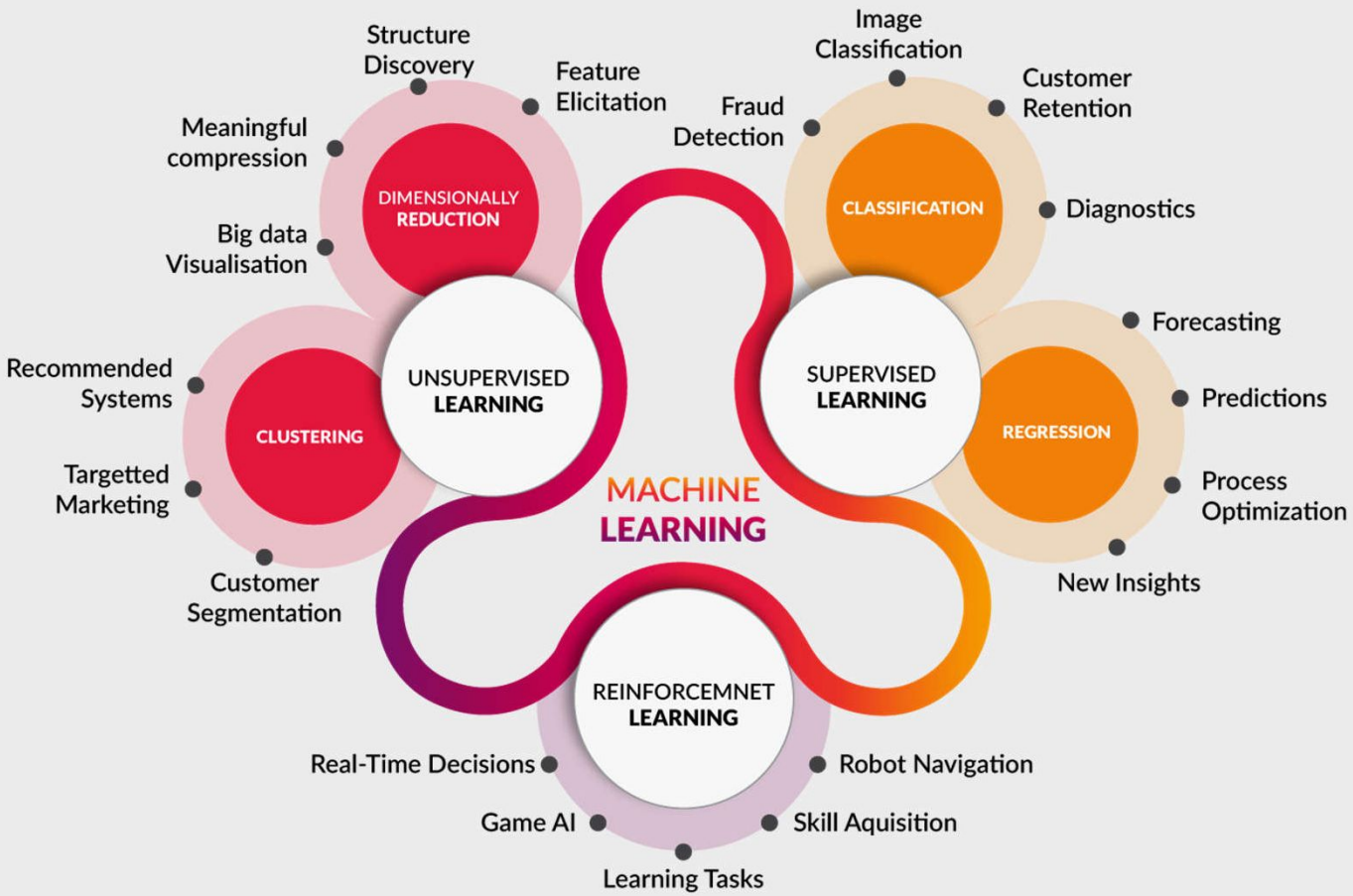


Error =
Desired Output
- Actual Output

Forward Pass: Error is calculated based on desired and actual output.

Backward Pass: Synaptic weights are adjusted based on calculated error.

Applications of Machine Learning



UCI Machine Learning Repository - Windows Internet Explorer

http://archive.ics.uci.edu/ml/

UCI Machine Learning Repository

About Citation Policy Donate a Data Set Contact

Search

Repository Web

View ALL Data Sets

Machine Learning Repository

Center for Machine Learning and Intelligent Systems

Welcome to the UC Irvine Machine Learning Repository!

We currently maintain 174 data sets as a service to the machine learning community. You may **view all data sets** through our searchable interface. Our old web site is still available, for those who prefer the old format. For a general overview of the Repository, please visit our About page. For information about citing data sets in publications, please read our citation policy. If you wish to donate a data set, please consult our donation policy. For any other questions, feel free to contact the Repository librarians. We have also set up a mirror site for the Repository.

Supported By: In Collaboration With: Rexa.info

Latest News:

07-23-2008: Repository mirror has been set up.
03-24-2008: New data sets have been added!
06-25-2007: Two new data sets have been added: UJI Pen Characters, MAGIC Gamma Telescope
04-13-2007: Research papers that cite the repository have been associated to specific data sets.
04-09-2007: Three new data sets have been added: Poker Hand, Callt2 Building People Counts, Dodgers Loop Sensor.
09-08-2006: The Beta site has been launched.
09-01-2006: SPECTF.test has been modified by the donor.

Featured Data Set: Statlog (Shuttle)

Task: Classification
Data Type: Multivariate
 # Attributes: 9
 # Instances: 58000

The shuttle dataset contains 9 attributes all of which are numerical. Approximately 80% of the data belongs to class 1

Newest Data Sets:

06-26-2008: Parkinsons
04-21-2008: Ozone Level Detection
04-03-2008: Abscisic Acid Signaling Network
03-20-2008: Hill-Valley
03-12-2008: Bag of Words
03-08-2008: Reuters Transcribed Subset
02-29-2008: Gisette
02-29-2008: Dorothea
02-29-2008: Madelon

Most Popular Data Sets (hits since 2007):

39351: Iris
31585: Adult
26458: Wine
23553: Breast Cancer Wisconsin (Diagnostic)
18449: Abalone
18321: Poker Hand
13471: Yeast
12996: Internet Advertisements
11793: SPECT Heart

UCI Machine Learning Repository - Windows Internet Explorer

Thesis | idea of neural ne... | UCI Machine Lea... | FTP directory: /p... | Ensemble of Div... | H-Bibliography.d... | Desktop | 4:39 PM

Benchmark Problems

Problems Related to Human Life

Problem	Task
Breast Cancer Wisconsin	Predicts whether a tumor is benign (not dangerous to health) or malignant (dangerous) based on a sample tissue taken from a patient's breast.
BUPA Liver Disorder	Identify lever disorders based on blood tests along with other related information such as alcohol consumption.
Diabetes	Investigate whether the patient shows or not the signs of diabetes.
Heart Disease Cleveland	Predicting whether at least one of four major heart vessels is reduced in diameter by more than 50%.
Hepatitis	Anticipate status (i.e., live or die) of hepatitis patient.
Lymphography	Predict the situation of lymph nodes and lymphatic vessels.
Lungcancer	Identify types of pathological lung cancers.
Postoperative	Determine place to send patients for postoperative recovery.

Input Features of Diabetes

1. Number of times pregnant
2. Plasma glucose concentration
3. Diastolic blood pressure
4. Triceps skin fold thickness (mm)
5. 2-Hour serum insulin (mu U/ml)
6. Body mass index
7. Diabetes pedigree function
8. Age

Benchmark Problems

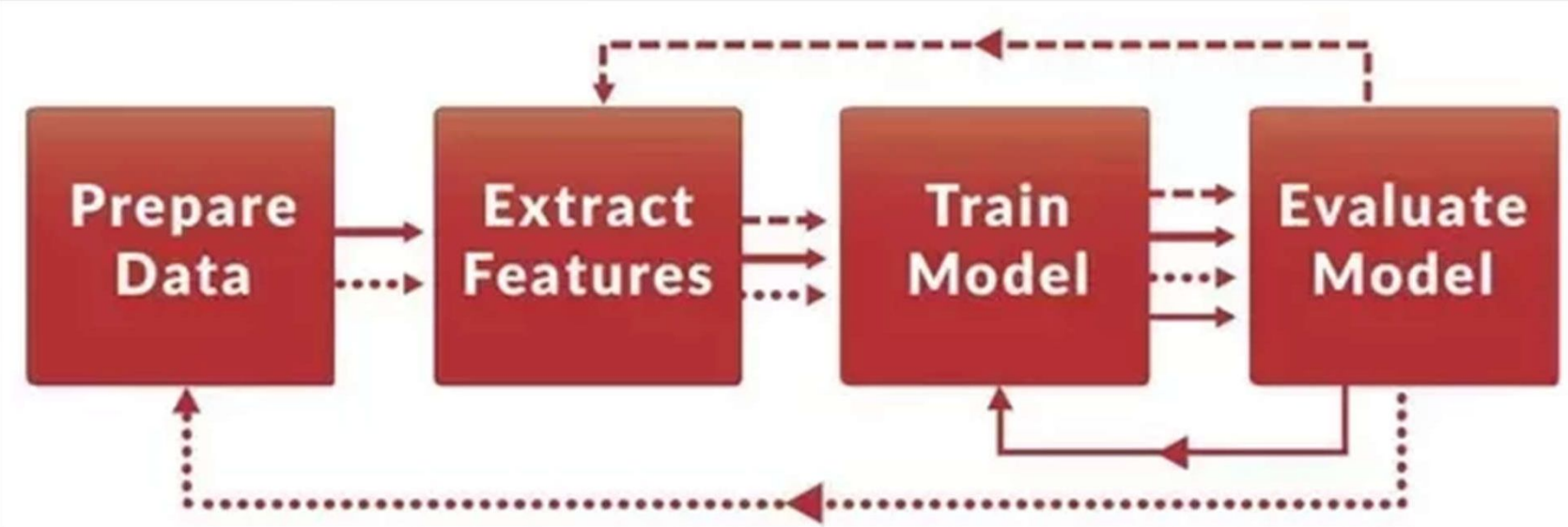
Problems Related to Finance

Problem	Task
Australian Credit Card	Classify people as good or bad credit risks depend on applicants' particulars.
Car	Evaluate cars based on price and facilities.
Labor Negotiations	Identify a worker as good or bad i.e., contract with him beneficial or not.
German Credit Card	Like AustralianCard, this problem also concerns to predict the approval or non-approval of a credit card to a customer.

Problems Related to Plants

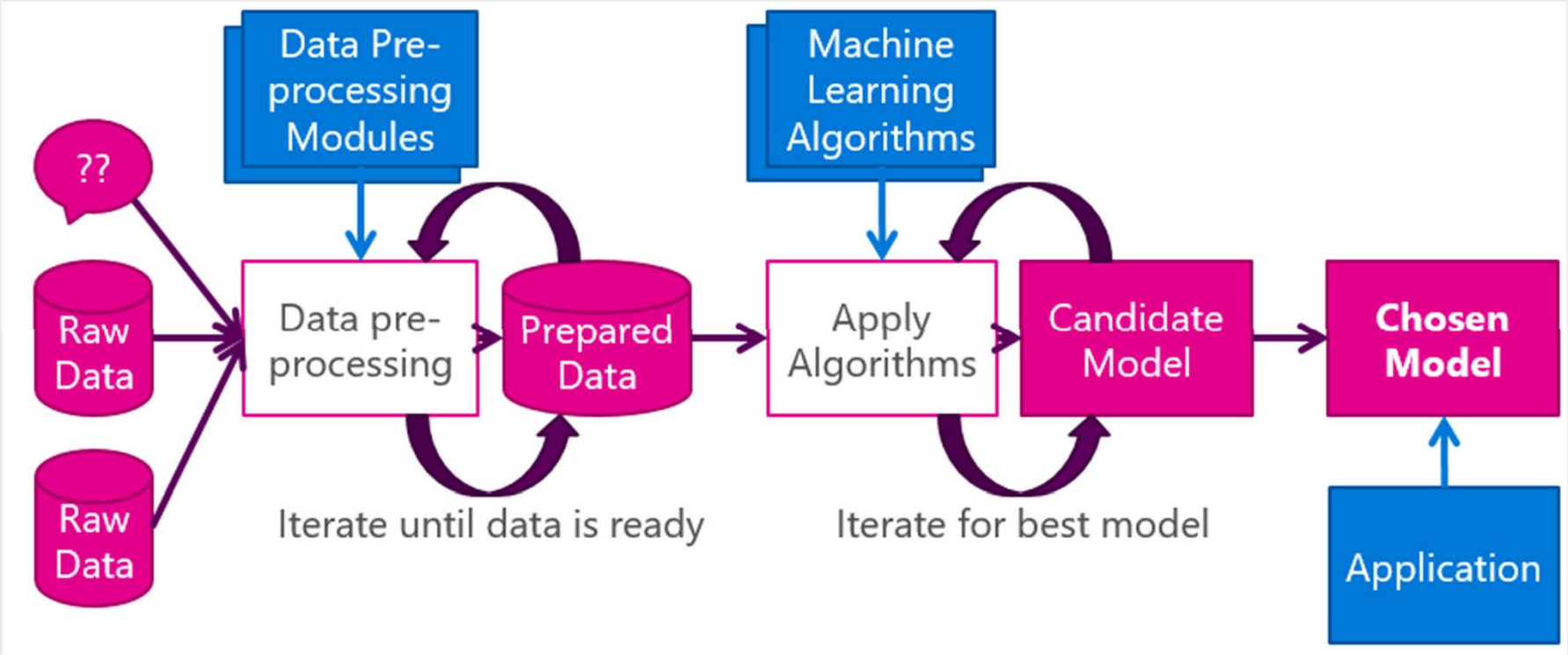
Problem	Task
Iris Plants	Classify iris plant types.
Mushroom	Identify whether a mushroom is edible or not based on a description of the mushroom's shape, color, odor, and habitat.
Soybean	Recognize 19 different diseases of soybeans.

Machine Learning Process



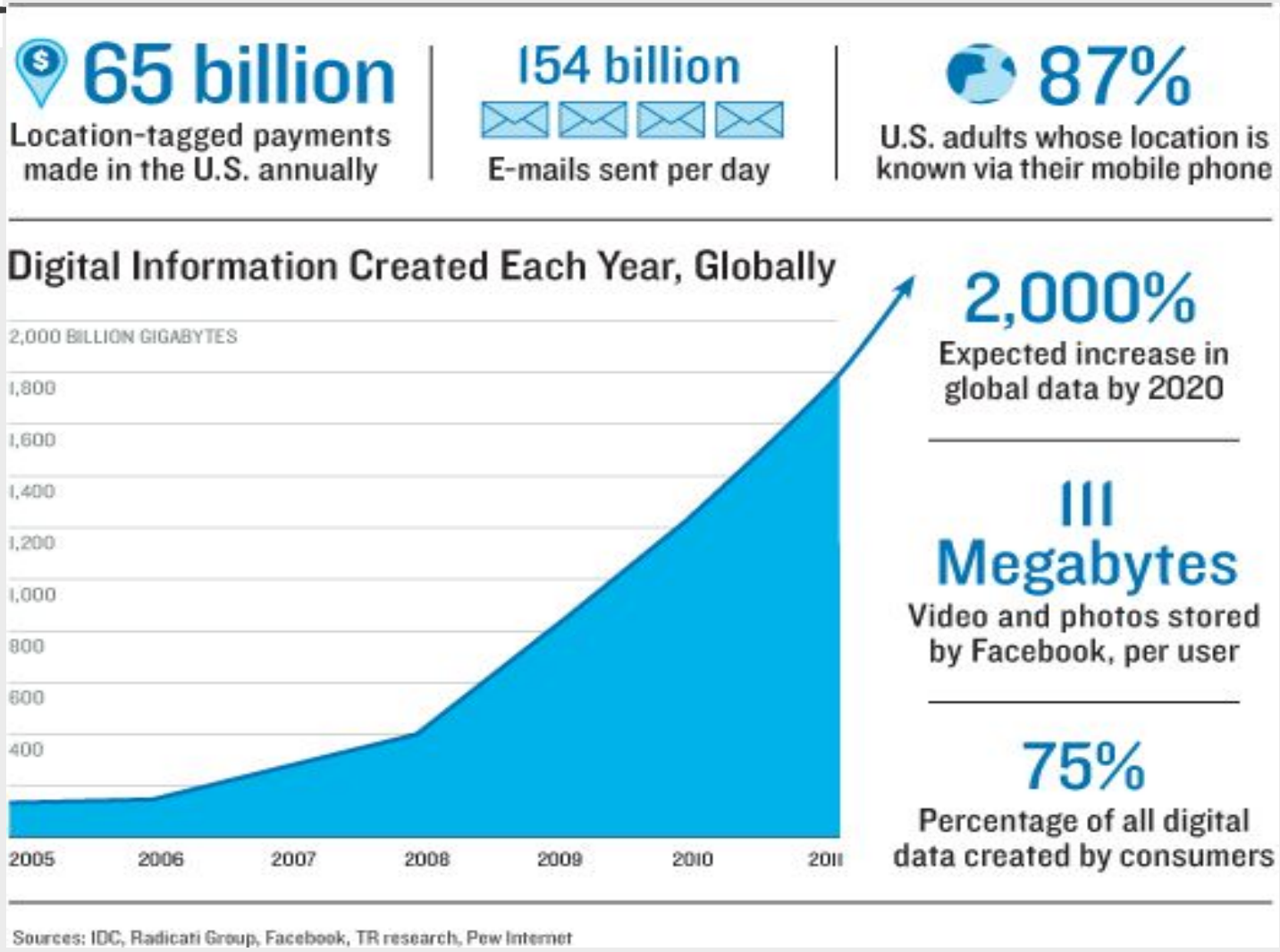
Performance depends on feature extraction which requires heuristics and, most notably, a great deal of time

Machine Learning Process (Different View)



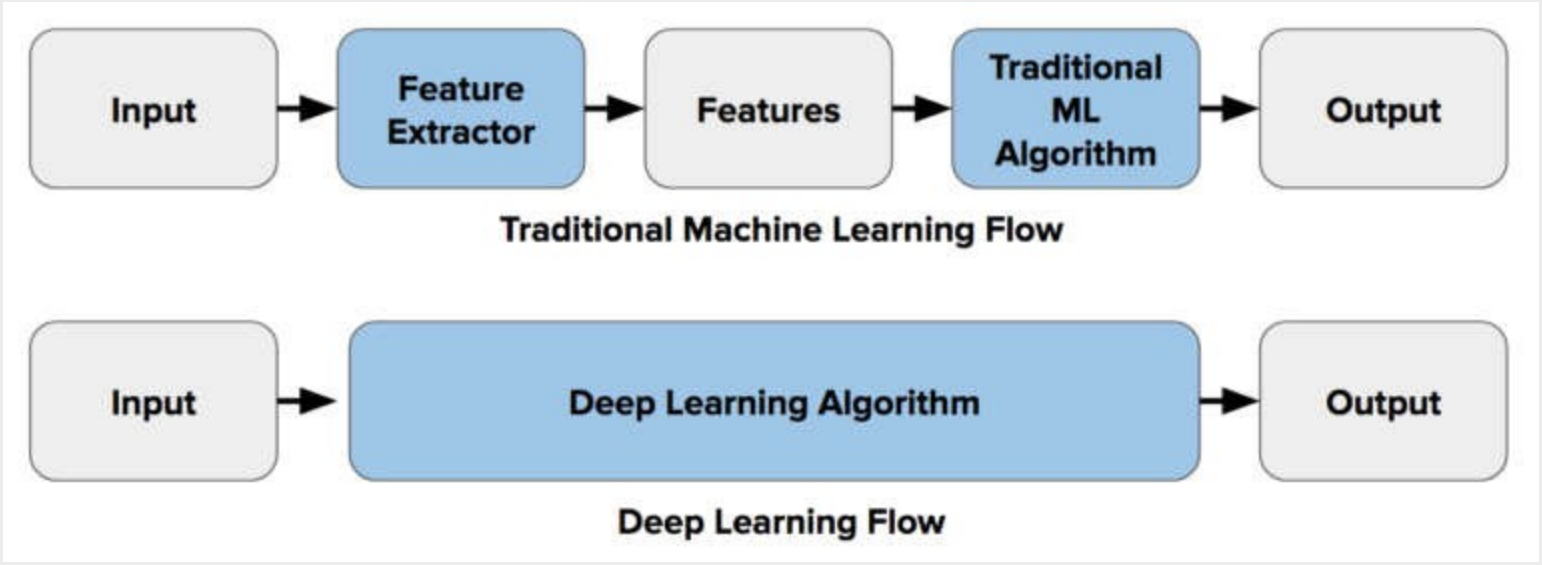
Performance depends on data pre-processing (i.e., feature extraction) which requires human expertize

Why Deep Architecture?

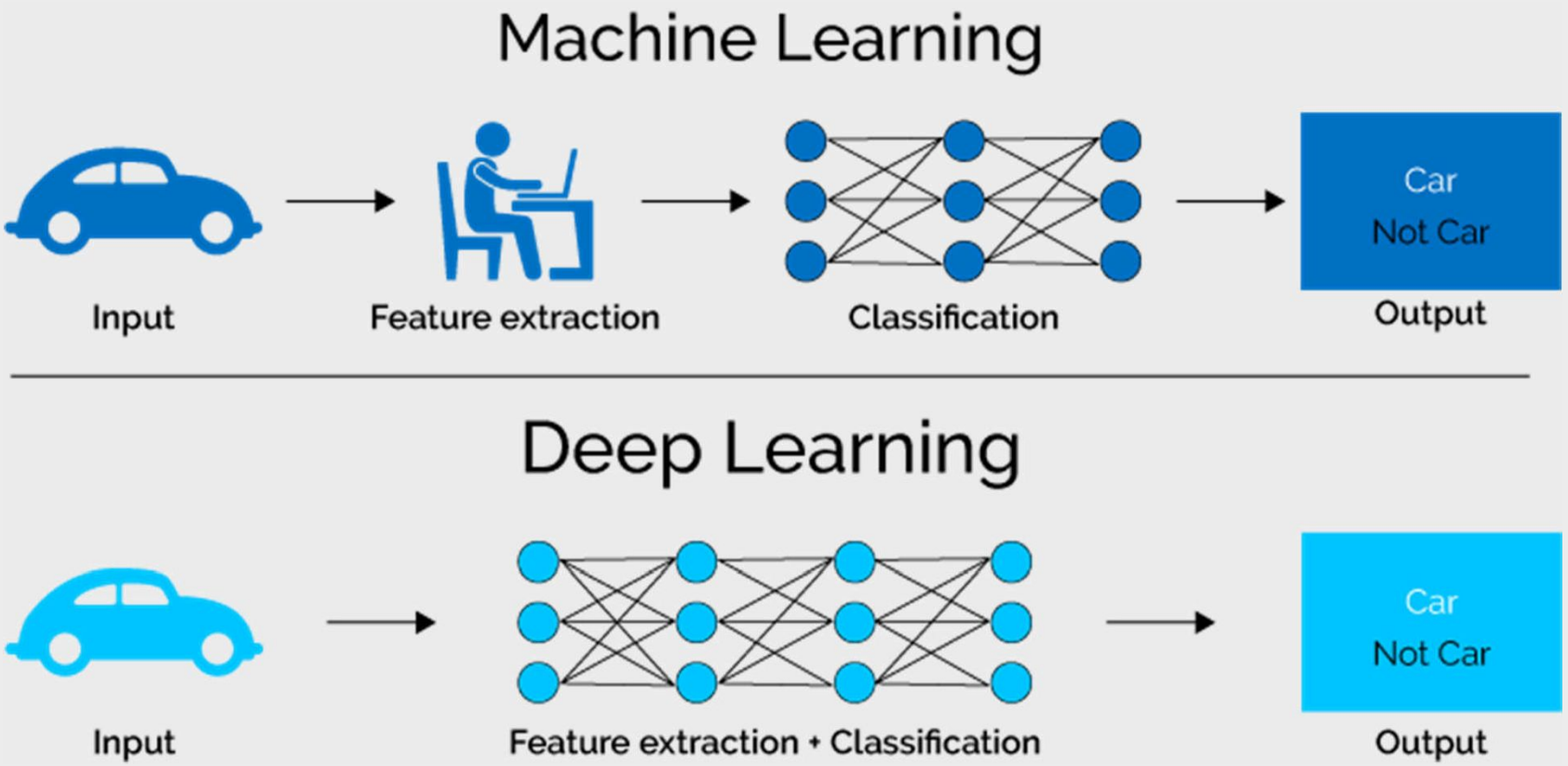


For huge heterogeneous data feature extraction is challenging and training time unacceptable

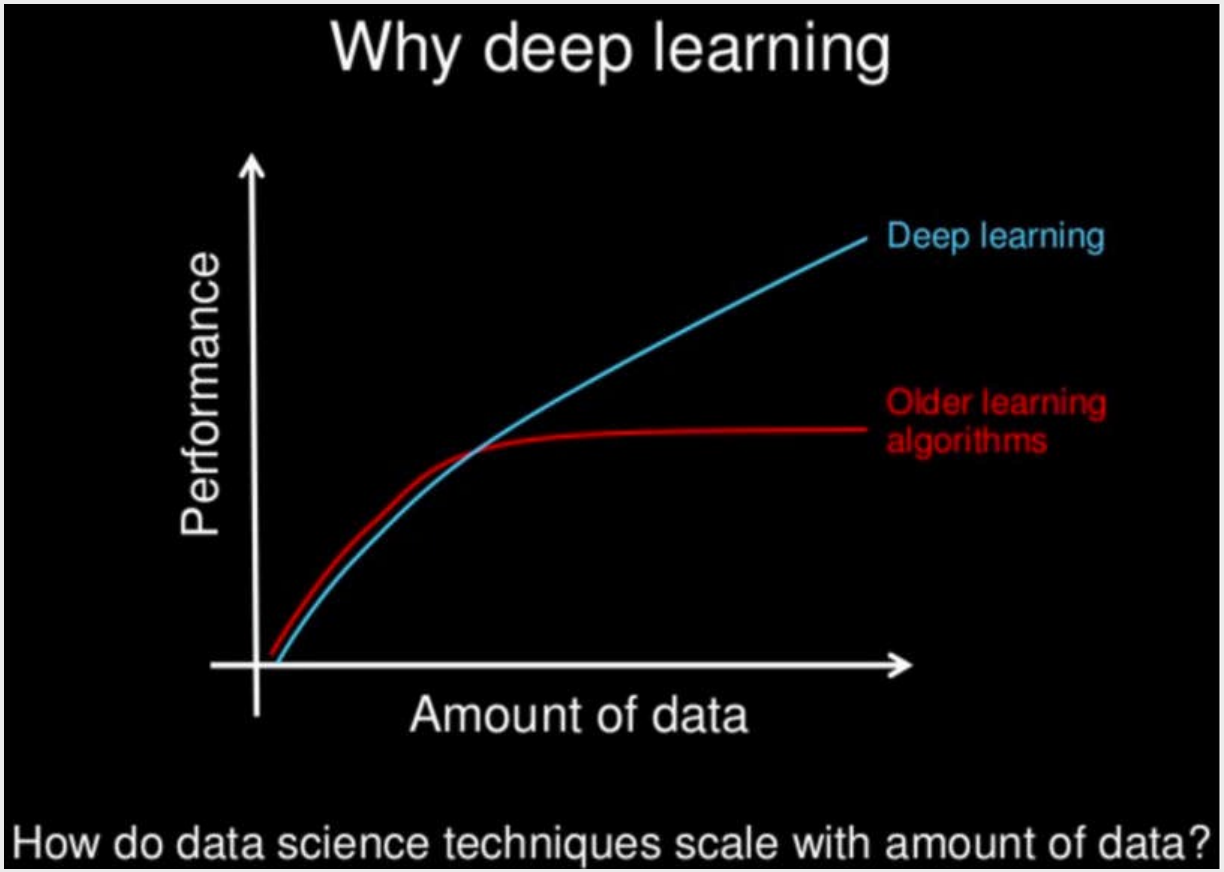
Motivation of Deep Learning



Machine Learning vs Deep Learning



Machine Learning vs Deep Learning



Understanding Deep Learning

1. ***what exactly is deep learning ? And,***
2. ***why is it generally better than other methods on image, speech and certain other types of data?***

The short answers

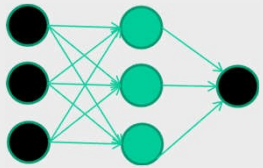
1. ***'Deep Learning' means using a neural network with several layers of nodes between input and output***

2.the series of layers between input & output do feature identification and processing in a series of stages, just as our brains seem to.

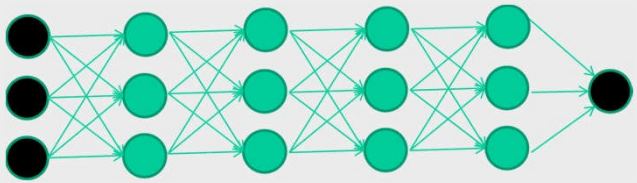
Understanding Deep Learning

3. multilayer neural networks have been around for 25 years. What's actually new?

we have always had good algorithms for learning the weights in networks with 1 hidden layer



but these algorithms are not good at learning the weights for networks with more hidden layers



what's new is: ***<u>algorithms for training many-layer networks</u>***

When to use Deep Learning (DL)?

- **DL** outperforms other techniques if the ***data size is large***.
- When there is ***lack of domain understanding for feature introspection***, **DL** techniques outshines others as you have to worry less about feature engineering.
- **DL** really shines when it comes to ***complex problems such as image classification, natural language processing, and speech recognition***.

60+ STARTUPS USING DEEP LEARNING

CORE AI: COMPUTER VISION

clarifai, Tractable, netra, HYPERVERGE, captricity, deepomatic, sentient, Sighthound, terraloupe, imagenii, pilot.ai

CORE AI: OTHER

TERADEEP, affectiva, Wave Computing, DigitalGenius, LEAPMIND, Arya, twentybn

BI, SALES & CRM

SKYMIND, TalkIQ, intranetum, ripjar, FRACTAL INDUSTRIES, MARIANA, DEEPGRAM, True AI, KOVID, ditto, Gridspace

CORE AI: VOICE INTERFACE

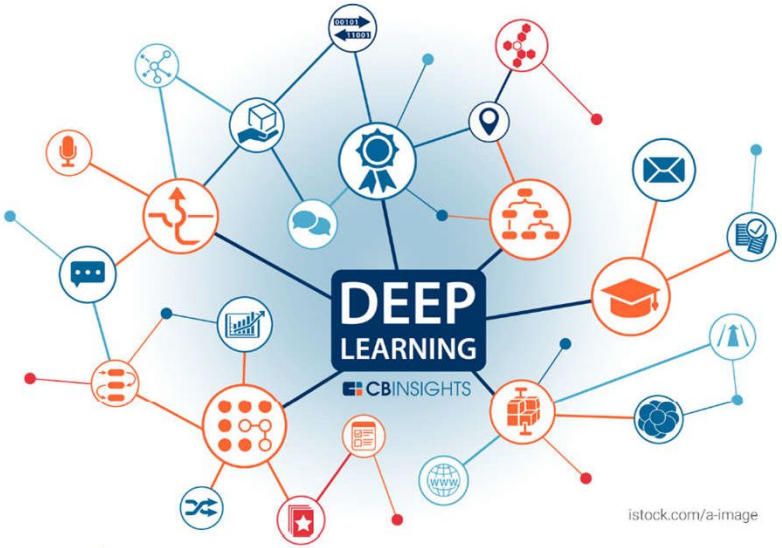
VIV, Mobvoi 出门问问, MaluubA, OBEN

ROBOTICS & AUTO

netradyne, AdasWorks, drive.ai, Rokid, 图灵机器人 TURING ROBOT, comma.ai

HEALTHCARE

imagia, Mindshare MEDICAL, BAYLABS, BUTTERFLY Network, Inc., deep genomics, Atomwise, zebra MEDICAL VISION, PATHWAY GENOMICS, AVALON AI, Behold.ai, MedyMatch, ADVENIO, enlitic, Lunit, SIGTUPLE



istock.com/a-image

SECURITY

deepinstinct, umbo CV, signalsense

OTHER

Alpaca, indico, Iris Automation

E-COMMERCE

Reflektion, CORTEXICA, VISENZE, STAQU

ACQUIRED

(2014-2016YTD)

Etsy Blackbird, Orbeus, IBM AlchemyAPI, intel Movidius, nervana SYSTEMS, Perceptio, vocaliq, turi, DEEPMIND, DNNresearch, Dark Blue Labs, Vision Factory, MAGIC PONY TECHNOLOGY, MADBITS, MetaMind

CBINSIGHTS

প্রথম আলো

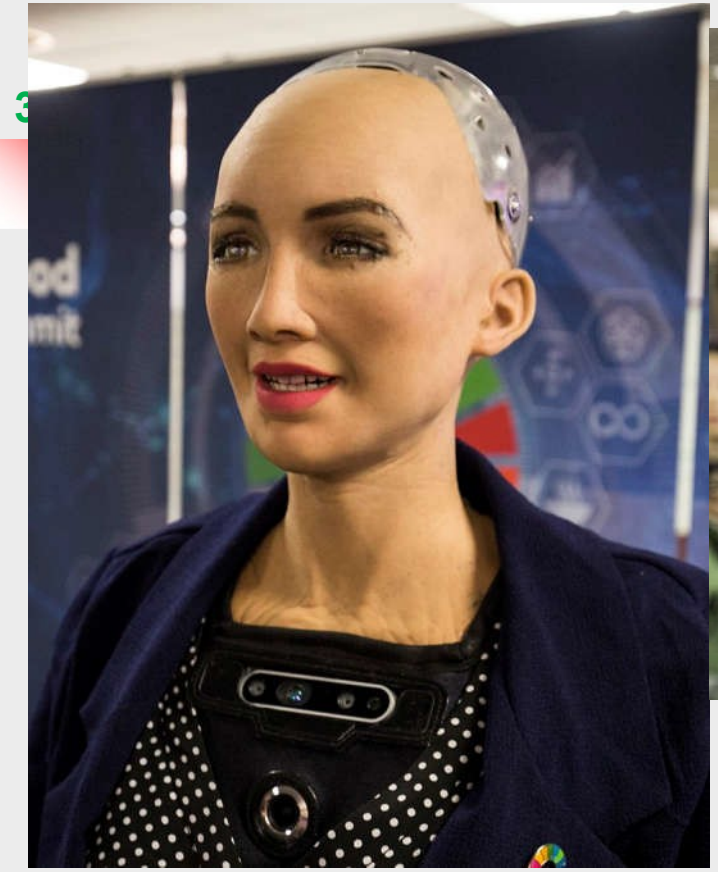
সাংবাদিক ছাড়াই সংবাদ

৩০ মে ২০১৮, ১৯ :৫০

আপডেট : ০২ জুন ২০১৮, ১১ :৪০

6/21/2018 About JX PRESS - English - JX通信社 https://jxpress.net/about/about-jx-press/ 2/5

The major technological areas where our company excels in the development of this kind of service are **machine learning, deep learning, and natural language processing technology**. Using these technologies, we aim to achieve qualitative improvements in journalism while simultaneously reducing costs through this “mechanization of the news



Sophia uses voice recognition (speech-to-text) technology from Alphabet Inc of Google.

--- Sophia does utilize AI methods including face tracking, emotion recognition, and robotic movements generated by deep neural networks. Sophia's dialogue is generated via a decision tree, but is integrated with these outputs uniquely.

Deep Learning Review in Nature

rnal/v521/n7553/full/nature14539.html

nature International weekly journal of science

Search Advanced search

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive | Volume 521 | Issue 7553 | Insights | Reviews | Article

NATURE | INSIGHT | REVIEW

Deep learning

Yann LeCun, Yoshua Bengio & Geoffrey Hinton

Affiliations | Corresponding author

Nature **521**, 436–444 (28 May 2015) | doi:10.1038/nature14539
Received 25 February 2015 | Accepted 01 May 2015 | Published online 27 May 2015

PDF | Citation | Rights & permissions | Article metrics

Abstract

Abstract · Supervised learning · Backpropagation to train multilayer architectures · Convolutional neural networks · Image understanding with deep convolutional networks · Distributed representations and language processing · Recurrent neural networks · The future of deep learning · References · Acknowledgements · Author information

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

Subject terms: Mathematics and computing · Computer science

At a glance

nl#insight

Editors' pick

Image credit: Gabriela Hasbun for Nature

The rise and fall and rise again of 23andMe ›

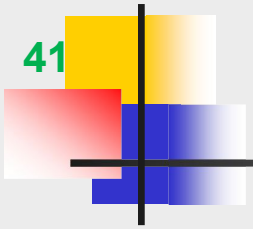
Science jobs | Science events

naturejobs.com

Professor and Faculty Positions at the Academy of Medical Sciences (AMS), Zhengzhou University
The Academy of Medical Sciences of Zhengzhou University

Associate Editor / Senior Editor roles, Nature Research - Talent Pool 2017
Springer Nature

Postdocs, Key Lab for Neuroinformation, University of Electronic Sciences and



Prominent Deep Learning Methods

- Stacked Denoising Autoencoder (SDAE)
- Deep Belief Network (DBN)

} Final Architecture same as NN but preparation differs

- Convolutional Neural Network (CNN) — Efficient for 2D data e.g., Image
- Long Short-Term Memory (LSTM) — Developed for Sequential Data

No stopping of Deep Learning