

What is Artificial Intelligence?

Objectives

- The Three eras of AI with key individuals, their contribution and timelines
- Role of math, machines and neuroscience in AI
- Basic understanding of;
 - Perceptrons and Neural nets
 - Classification
 - Linear regression, Logistic regression, K-means
- Machine Learning
 - Supervised vs Unsupervised vs Reinforcement Learning
 - Convolutional Neural Nets
 - Recurrent Neural nets
 - Tensor Flow

Marvin Minsky, MIT



Artificial intelligence is the science
of making machines do things that
would require intelligence if done by
men.

— Marvin Minsky —

AZ QUOTES

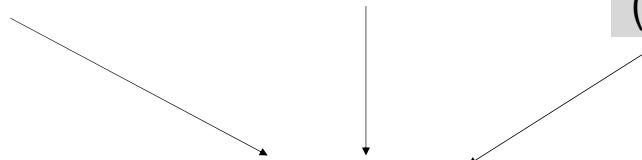
The Three Dimensions of AI

Math & Logic

Machines

Biology
(Neuroscience)

AI



Three Major Phases of AI

Phase I: 1600 – 1945
AI Before Digital Computers

Phase 2: 1945 – 2001
AI and Computers

Phase 3: 2001 –
AI meets Big Data

Math & Logic

1600

Calculus

1800

**Taylor series
(sin, cos, log)**

1900

**Axiomatic
Logic & Set
Theory**

1945

Machines

Leibnitz (1620) - Calculator

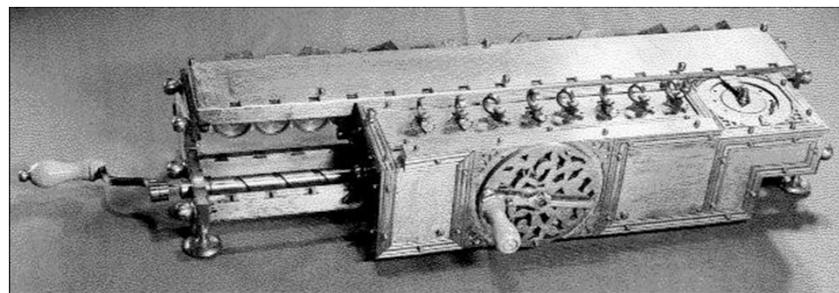
Charles Babbage (1820) –
Difference Engine

Alan Turing (1930s) – Turing Machine

John von Neumann (1940s)

Phase I

Leibnitz – Step Reckoner –
could add, subtract, multiply
and divide



Math & Logic

1600

Calculus

1800

**Taylor series
(sin, cos, log)**

1900

**Axiomatic
Logic & Set
Theory**

1945

Machines

Leibnitz (1620) - Calculator

Charles Babbage (1820) –
Difference Engine

Alan Turing (1930s) - TM

John von Neumann (1940s)



1820

Charles Babbage



Ships in the British Navy were running aground and sailors were dying because the logarithm tables used in calculations contained errors – even though the calculations were done by COMPUTERS !

What kind of computer existed in 1820?



Computers were people who did arithmetic for a living

Taylor Series – can compute sin, cos, tangent and e^x using an infinite series
= A big polynomial

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$

$$= \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Logarithms let you multiplication and division by doing addition and subtraction

$$\log_b(MN) = \log_b(M) + \log_b(N)$$

$$\log_b\left(\frac{M}{N}\right) = \log_b(M) - \log_b(N)$$

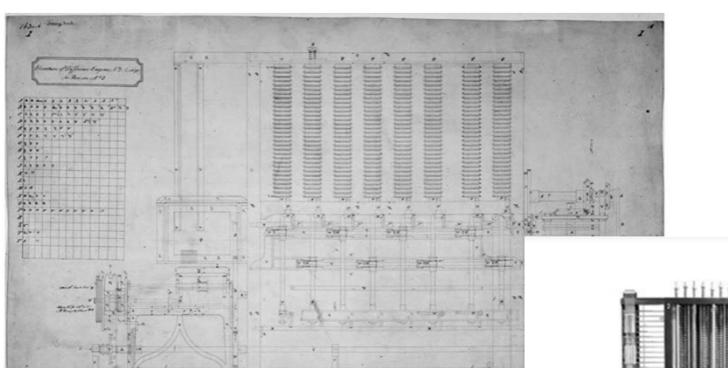
$$\log_b(M^p) = p \log_b(M)$$

BUT, errors in log tables resulted in numerous ships lost at sea.

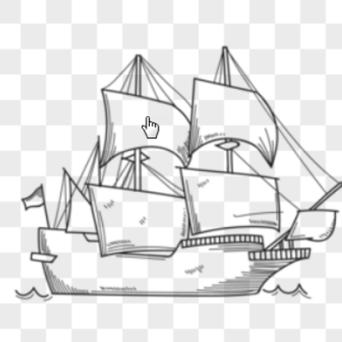
COMMON LOGARITHMS log x			
x	y	z	ADD
10.0000	0000.0000	0000.0000	0000.0000
11.0000	0413 0403 0531	0000 0000 0000	0000 0000 0000
12.0000	0842 0846 0849	0000 0000 0000	0000 0000 0000
13.0000	1273 1266 1259	0000 0000 0000	0000 0000 0000
14.0000	1704 1693 1684	0000 0000 0000	0000 0000 0000
15.0000	2135 2126 2114	0000 0000 0000	0000 0000 0000
16.0000	2566 2556 2544	0000 0000 0000	0000 0000 0000
17.0000	3000 2986 2974	0000 0000 0000	0000 0000 0000
18.0000	3430 3414 3393	0000 0000 0000	0000 0000 0000
19.0000	3870 3850 3829	0000 0000 0000	0000 0000 0000
20.0000	4300 4274 4249	0000 0000 0000	0000 0000 0000
21.0000	4730 4697 4664	0000 0000 0000	0000 0000 0000
22.0000	5160 5124 5087	0000 0000 0000	0000 0000 0000
23.0000	5590 5546 5499	0000 0000 0000	0000 0000 0000
24.0000	6020 5963 5892	0000 0000 0000	0000 0000 0000
25.0000	6450 6381 6304	0000 0000 0000	0000 0000 0000
26.0000	6879 6803 6719	0000 0000 0000	0000 0000 0000
27.0000	7310 7227 7136	0000 0000 0000	0000 0000 0000
28.0000	7739 7646 7547	0000 0000 0000	0000 0000 0000
29.0000	8168 8054 7946	0000 0000 0000	0000 0000 0000
30.0000	8597 8471 8349	0000 0000 0000	0000 0000 0000
31.0000	9026 8844 8699	0000 0000 0000	0000 0000 0000
32.0000	9454 9213 8975	0000 0000 0000	0000 0000 0000
33.0000	9882 9627 9365	0000 0000 0000	0000 0000 0000
34.0000	10309 10013 9645	0000 0000 0000	0000 0000 0000
35.0000	10736 10414 9975	0000 0000 0000	0000 0000 0000
36.0000	11163 10801 10502	0000 0000 0000	0000 0000 0000
37.0000	11590 11384 11081	0000 0000 0000	0000 0000 0000
38.0000	12017 11864 11561	0000 0000 0000	0000 0000 0000
39.0000	12444 12344 12044	0000 0000 0000	0000 0000 0000
40.0000	12871 12744 12444	0000 0000 0000	0000 0000 0000
41.0000	13300 13144 12844	0000 0000 0000	0000 0000 0000
42.0000	13727 13544 13244	0000 0000 0000	0000 0000 0000
43.0000	14154 13944 13644	0000 0000 0000	0000 0000 0000
44.0000	14581 14344 14044	0000 0000 0000	0000 0000 0000
45.0000	15008 14744 14444	0000 0000 0000	0000 0000 0000
46.0000	15435 15144 14844	0000 0000 0000	0000 0000 0000
47.0000	15862 15544 15244	0000 0000 0000	0000 0000 0000
48.0000	16289 16344 16044	0000 0000 0000	0000 0000 0000
49.0000	16716 16744 16444	0000 0000 0000	0000 0000 0000
50.0000	17143 17144 16844	0000 0000 0000	0000 0000 0000
51.0000	17570 17544 17244	0000 0000 0000	0000 0000 0000
52.0000	18000 17944 17644	0000 0000 0000	0000 0000 0000
53.0000	18427 18344 18044	0000 0000 0000	0000 0000 0000
54.0000	18854 18744 18444	0000 0000 0000	0000 0000 0000
55.0000	19281 19144 18844	0000 0000 0000	0000 0000 0000
56.0000	19708 19644 19344	0000 0000 0000	0000 0000 0000
57.0000	20135 20044 19744	0000 0000 0000	0000 0000 0000
58.0000	20562 20444 19944	0000 0000 0000	0000 0000 0000
59.0000	20989 20844 20544	0000 0000 0000	0000 0000 0000
60.0000	21416 21344 21044	0000 0000 0000	0000 0000 0000
61.0000	21843 21744 21444	0000 0000 0000	0000 0000 0000
62.0000	22270 22144 21844	0000 0000 0000	0000 0000 0000
63.0000	22697 22544 22244	0000 0000 0000	0000 0000 0000
64.0000	23124 23044 22744	0000 0000 0000	0000 0000 0000
65.0000	23551 23444 23144	0000 0000 0000	0000 0000 0000
66.0000	23978 23844 23544	0000 0000 0000	0000 0000 0000
67.0000	24405 24244 23944	0000 0000 0000	0000 0000 0000
68.0000	24832 24644 24344	0000 0000 0000	0000 0000 0000
69.0000	25259 25044 24744	0000 0000 0000	0000 0000 0000
70.0000	25686 25444 25144	0000 0000 0000	0000 0000 0000
71.0000	26113 25844 25544	0000 0000 0000	0000 0000 0000
72.0000	26540 26244 25944	0000 0000 0000	0000 0000 0000
73.0000	26967 26644 26344	0000 0000 0000	0000 0000 0000
74.0000	27404 27144 26844	0000 0000 0000	0000 0000 0000
75.0000	27831 27544 27244	0000 0000 0000	0000 0000 0000
76.0000	28258 27944 27644	0000 0000 0000	0000 0000 0000
77.0000	28685 28344 28044	0000 0000 0000	0000 0000 0000
78.0000	29112 28844 28544	0000 0000 0000	0000 0000 0000
79.0000	29539 29244 28944	0000 0000 0000	0000 0000 0000
80.0000	29966 29644 29344	0000 0000 0000	0000 0000 0000
81.0000	30393 30044 29744	0000 0000 0000	0000 0000 0000
82.0000	30820 30444 29144	0000 0000 0000	0000 0000 0000
83.0000	31247 30944 29844	0000 0000 0000	0000 0000 0000
84.0000	31674 31544 29544	0000 0000 0000	0000 0000 0000
85.0000	32101 31844 30244	0000 0000 0000	0000 0000 0000
86.0000	32528 32344 30944	0000 0000 0000	0000 0000 0000
87.0000	32955 32744 31644	0000 0000 0000	0000 0000 0000
88.0000	33382 33144 31344	0000 0000 0000	0000 0000 0000
89.0000	33809 33544 32044	0000 0000 0000	0000 0000 0000
90.0000	34236 34044 32744	0000 0000 0000	0000 0000 0000
91.0000	34663 34444 33444	0000 0000 0000	0000 0000 0000
92.0000	35090 34844 34144	0000 0000 0000	0000 0000 0000
93.0000	35517 35344 34844	0000 0000 0000	0000 0000 0000
94.0000	35944 35744 35444	0000 0000 0000	0000 0000 0000
95.0000	36371 36144 35144	0000 0000 0000	0000 0000 0000
96.0000	36808 36644 35844	0000 0000 0000	0000 0000 0000
97.0000	37235 37044 36144	0000 0000 0000	0000 0000 0000
98.0000	37662 37444 36344	0000 0000 0000	0000 0000 0000
99.0000	38089 37844 37044	0000 0000 0000	0000 0000 0000
100.0000	38516 38344 37344	0000 0000 0000	0000 0000 0000
No.	log		
$x = 3.14159$	0.49125	$\log x = \log_e x \cdot (\log_e 10)^{-1}$	No.
$x = 3.141592653589793$	0.491253383279489	$\log_e x = M \cdot F$	$\log_e 10 = 2.302585092994046$
$x = 3.141592653589793$	0.491253383279489	$M = 0.491253383279489$	$F = 0.000000000000000$
$x = 3.141592653589793$	0.491253383279489		
$x = 3.141592653589793$	0.491253383279489		

gggerard@singnet.com

Babbage Difference Engine

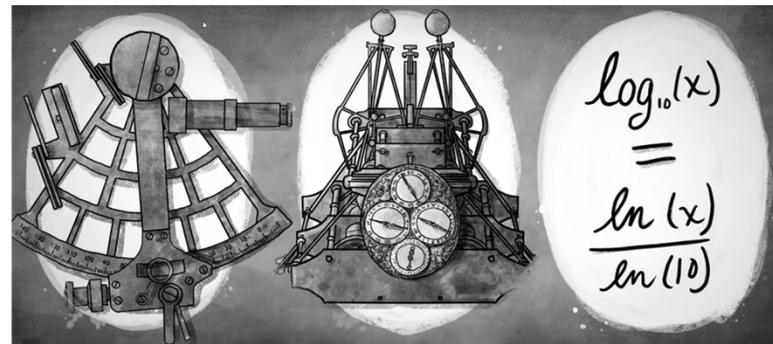


Computer History Museum



Navigation needs:

1. sextants
2. clocks
3. logarithms for math calculations



Easy way to compute logarithms using polynomials – based on difference tables

$$f(x) = 2x + 5$$

x	$f(x)$	Δ
1	7	
2	9	$9 - 7 = 2$
3	11	$11 - 9 = 2$
4	13	$13 - 11 = 2$

$$f(x) = 2x^2 + 3x + 1$$

x	$f(x)$	Δ_1	Δ_2
1	6		
2	15	$15 - 6 = 9$	
3	28	$28 - 15 = 13$	$13 - 9 = 4$
4	45	$45 - 28 = 17$	$17 - 13 = 4$
5	66	$66 - 45 = 21$	$21 - 17 = 4$

$$f(x) = 2x^2 + 3x + 1$$

x	$f(x)$	Δ_1	Δ_2
1	6	9	4
2	15	13	4
3	28	17	4
4	45	21	4

$$f(x) = x^7 + x$$

x	$f(x)$	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	Δ_7
3	2190	14198	47544	92526	109200	77280	30240	5040
4	16388	61742	140070	201726	186480	107520	35280	5040

$$f(x) = x^7 + x$$

x	$f(x)$	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	Δ_7
1	2							
2	130	128						
3	2190	2060	1932					
4	16388	14198	12138	10206				
5	78130	61742	47544	35406	25200			
6	279942	201812	140070	92526	57120	31920		
7	823550	543608	341796	201726	109200	52080	20160	
8	2097160	1273610	730002	388206	186480	77280	25200	5040
9	4782978	2685818	1412208	682206	294000	107520	30240	5040
10	10000010	5217032	2531214	1119006	436800	142800	35280	5040

Babbage's Difference Engine



Math & Logic

1600

Calculus

1800

**Taylor series
(sin, cos, log)**

1900

**Axiomatic
Logic & Set
Theory**

1945

Machines

Leibnitz (1620) - Calculator

Charles Babbage (1820) –
Difference Engine

Alan Turing (1930s) - TM

John von Neumann (1940s)

What are the limits of computation?

For Church it was the lambda-calculus ([λ calculus](#))ⁱⁱⁱ, what some have called the smallest universal programming language. It introduced nameless functions called lambda expressions, and had lots of parenthesis and recursion and took some getting accustomed to.

$$Y \equiv (\lambda y.(\lambda x.y(xx)))(\lambda x.y(xx)))$$

John von Neumann

The Three Dimensions of AI

Math & Logic

Machines

Biology
(Neuroscience)

AI

Biology
(Neuroscience)

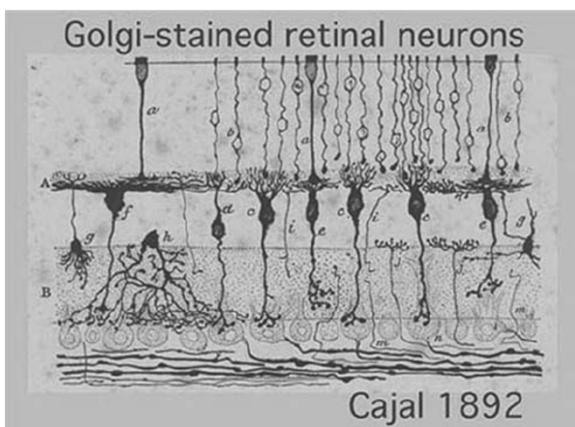
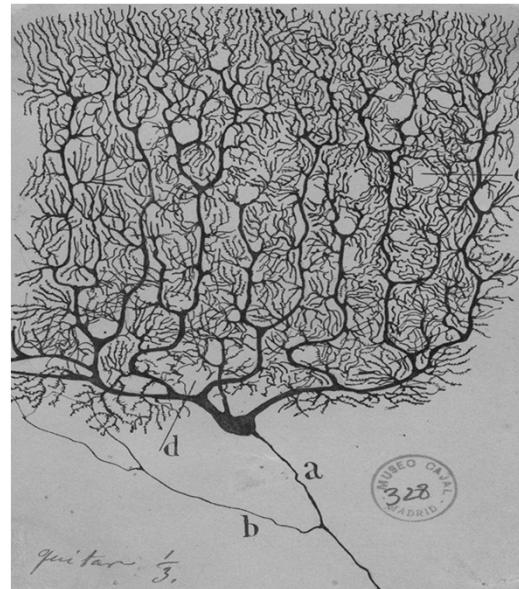


What's going
on inside ??



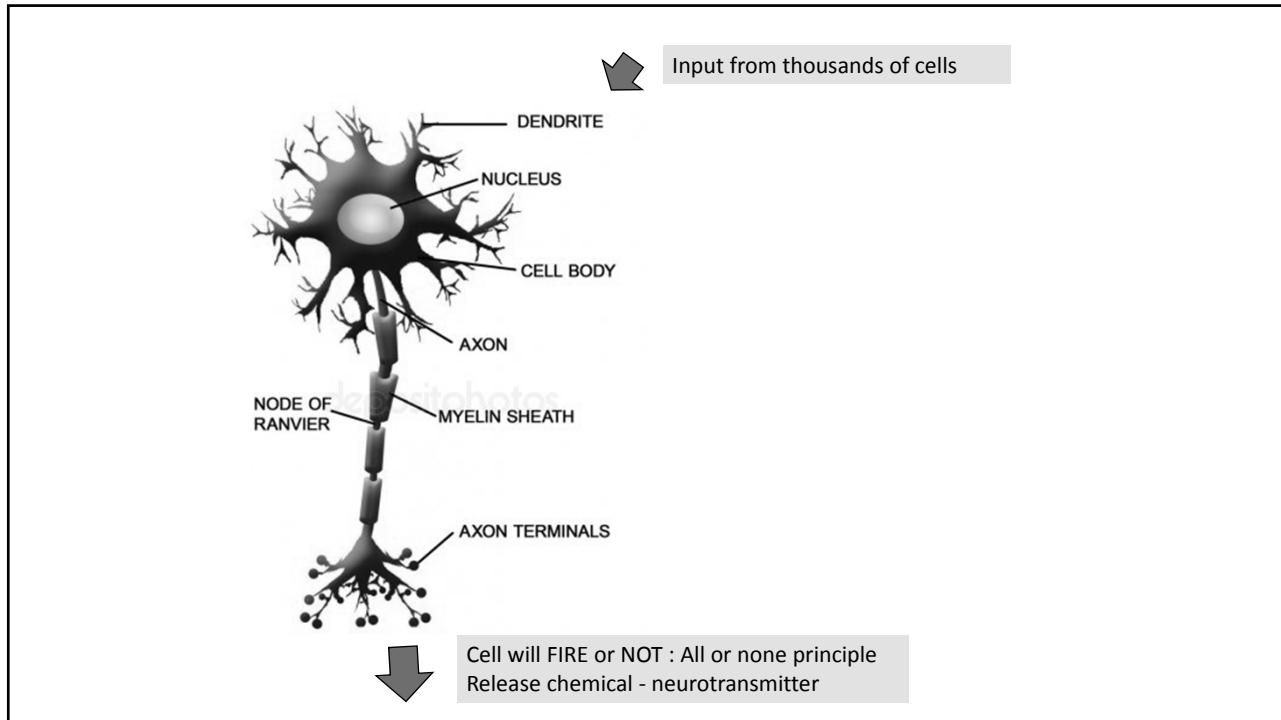
1890s Ramon Cajal

Dendrites – receptors – feed information into the cell body



BIG Question – how
do electrical signals
get from one cell to
another ?

Ramon Cajal discovered that nerve cells
were not continuous but contiguous



Three Major Phases of AI

**Phase I: 1600 – 1945
AI Before Digital Computers**

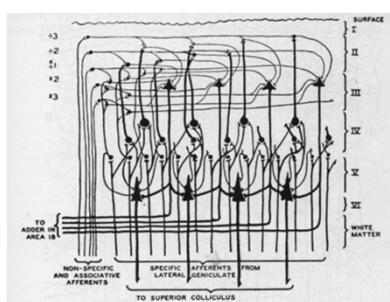
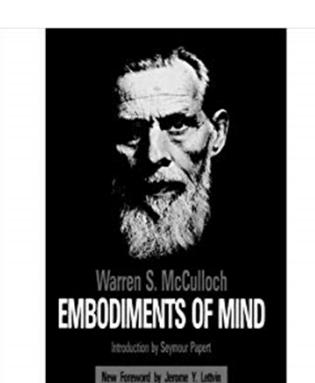
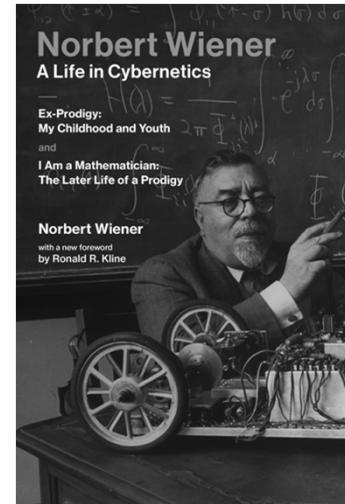
**Phase 2: 1945 – 2001
AI and Computers**

**Phase 3: 2001 –
AI meets Big Data**

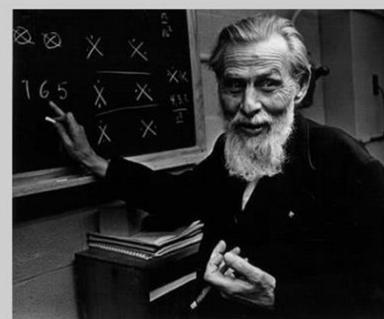
Cybernetics

Cybernetics is a transdisciplinary for exploring regulatory systems—their structures, constraints, and possibilities. Norbert Wiener defined cybernetics in 1948 as "the scientific study of control and communication in the animal and the machine."^[2] In the 21st century, the term is often used in a rather loose way to imply "control of any system using technology."

In other words, it is the scientific study of how humans, animals and machines control and communicate with each other.

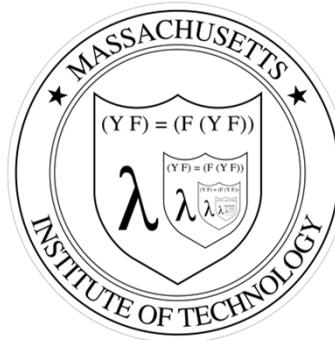


Warren McCulloch (16.11.1898 – 24.10.1969)



Warren Sturgis McCulloch was an American neurophysiologist and cybernetician, known for his work on the foundation for certain brain theories and his contribution to the movement cybernetics.

He is remembered for his work with Joannes Gregorius Dusser de Barenne from Yale and later with Walter Pitts from the University of Chicago. There he provided the foundation for certain brain theories in a number of classic papers, including "A Logical Calculus of the Ideas Immanent in Nervous Activity"



MIT 1950s

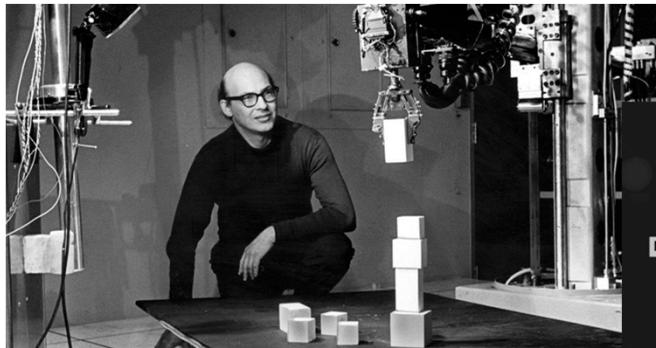


John McCarthy (MIT) : Invented the LISP programming language based on the lambda calculus to code chess programs

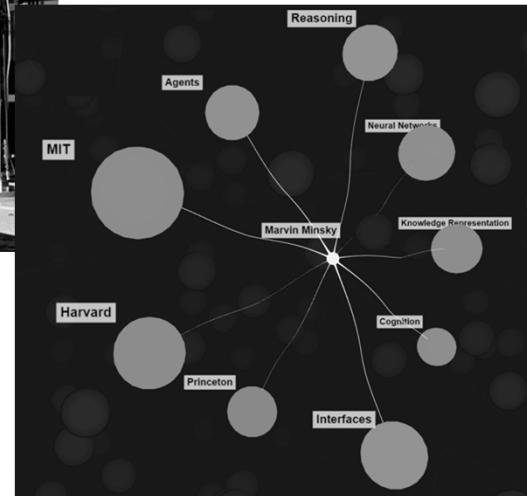
Computer Time-Sharing

- ⇒ “A bunch of people decided that time-sharing was clearly the way to work with a computer, but nobody could figure out how to make it work for general purpose computing – nobody except John”
- ⇒ “The Internet would not have happened nearly as soon as it did except for the fact that John initiated the development of time-sharing systems. We keep inventing new names for time-sharing. It came to be called servers.... Now we call it cloud computing. That is still just time-sharing. John started it.” (Les Earnest)





Marvin Minsky(MIT)



Perceptron – Frank Rosenblatt

Background on Perceptrons

In 1958 *Psychological Review* published "[The Perceptron: A Probabilistic Model For Information Storage and Organization in the Brain](#)", by Cornell researcher Frank Rosenblatt.

Perceptron is a single layer neural network and a multi-layer perceptron is called Neural Networks.

Perceptron is a linear classifier (binary). Also, it is used in supervised learning. It helps to classify the given input data. But how the heck it works ?

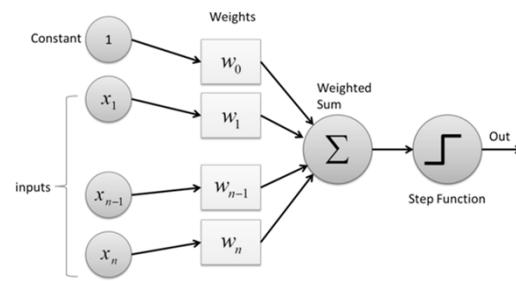
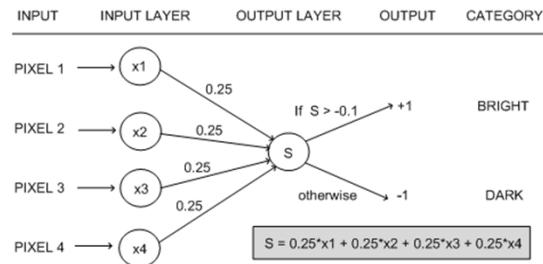


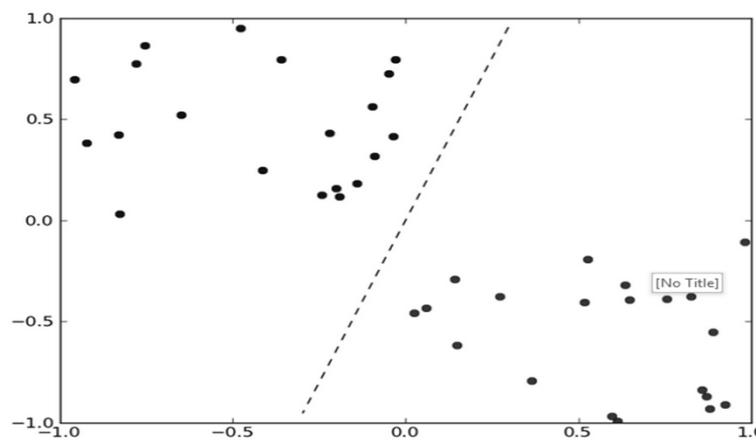
Fig : Perceptron



Why do we need Activation Function?

In short, the activation functions are used to map the input between the required values like $(0, 1)$ or $(-1, 1)$.

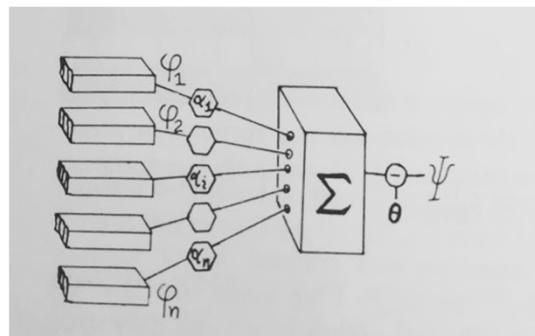
Perceptron is usually used to classify the data into two parts. Therefore, it is also known as a Linear Binary Classifier.



Soon after, media outlets including the New York Times and the New Yorker published articles about Rosenblatt's work. Interestingly, some of Rosenblatt's predictions regarding future of perceptrons proved to be surprisingly accurate.

"Later Perceptrons will be able to recognize people and call out their names and instantly translate speech in one language to speech or writing in another language, it was predicted."

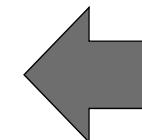
NY Times, "New Navy Device Learns By Doing" 7/7/1958

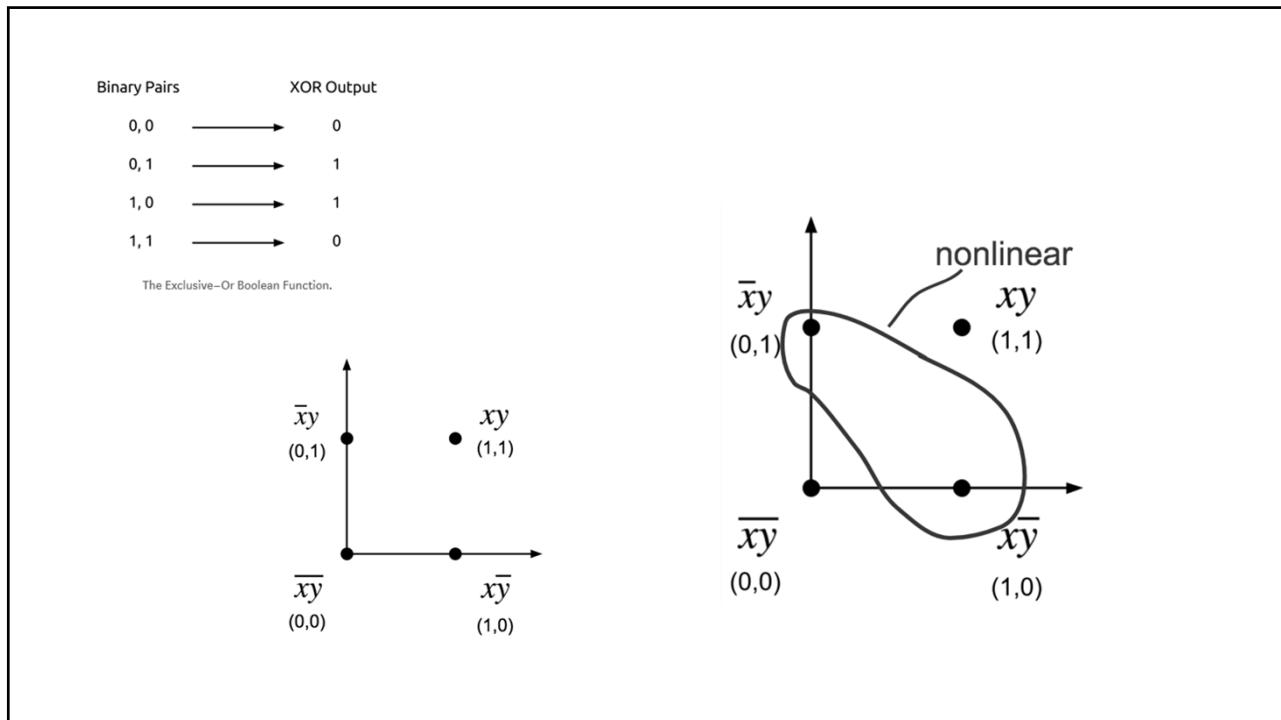


A perceptron is an approximator of linear functions (with an attached threshold function). Image source: "Perceptrons" Minsky, Papert.

In 1969, Marvin Minsky and Seymour Papert published *Perceptrons*—a historic text that would alter the course of artificial intelligence research for decades.

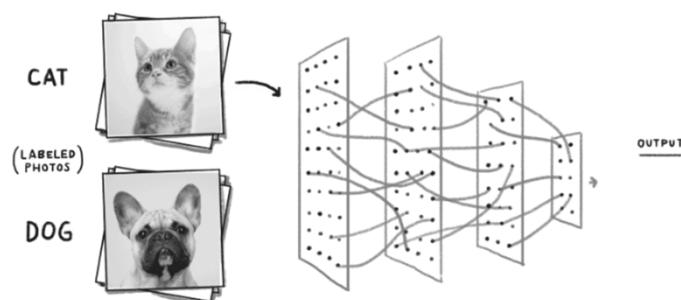
In the text, Minsky and Papert proved that a single perceptron—a grandparent to the computational units which compose modern neural networks—was incapable of learning the exclusive-or (aka XOR) function.





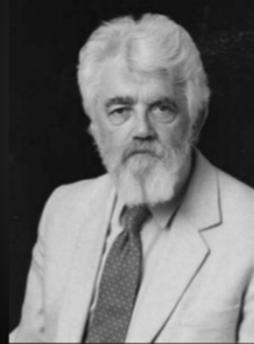
Because of Minsky's paper, researchers quit looking at Perceptrons.

Neural nets (perceptrons with multiple layers) were ignored until the 2000s with the availability of BIG DATA and GOOGLE.



HISTORY OF AI

JOHN MCCARTHY



- Precursors
- The birth of AI(1952-1956)
- The golden years(1956-1974)
- The first AI winter(1974-1980)
- Boom(1980-1987)
- Bust: second AI winter(1987-1993)
- AI(1993-2001)
- Deep learning, big data & artificial general intelligence(2001-current).

“Artificial Intelligence”: 1956

A.I. term was coined during a conference at Dartmouth College by 5 leading researchers



John McCarthy



Marvin Minsky



Herbert Simon



Arthur Samuel



Allen Newell

IN THIS BUILDING DURING THE SUMMER OF 1956

JOHN McCARTHY (DARTMOUTH COLLEGE), MARVIN L. MINSKY (MIT)
NATHANIEL ROCHESTER (IBM), AND CLAUDE SHANNON (BELL LABORATORIES)
CONDUCTED

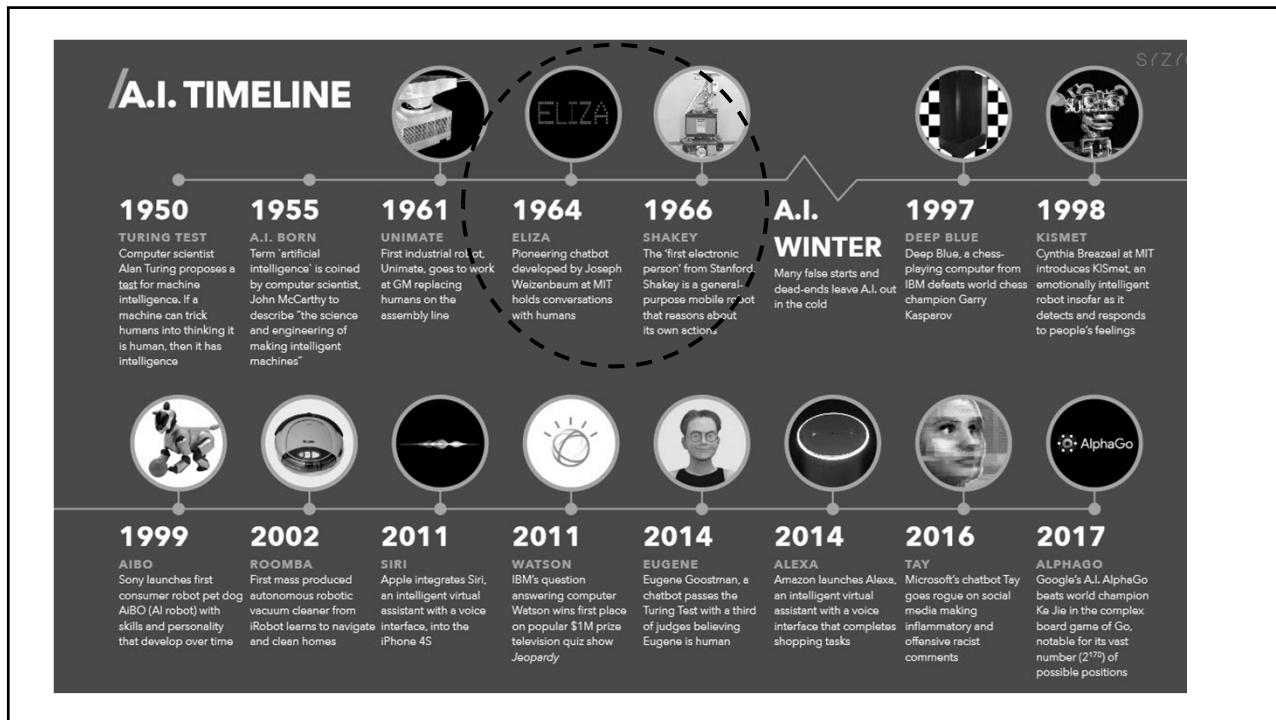
**THE DARTMOUTH SUMMER RESEARCH PROJECT
ON ARTIFICIAL INTELLIGENCE**

FIRST USE OF THE TERM "ARTIFICIAL INTELLIGENCE"

FOUNDING OF ARTIFICIAL INTELLIGENCE AS A RESEARCH DISCIPLINE

"To proceed on the basis of the conjecture
that every aspect of learning or any other feature of intelligence
can in principle be so precisely described that a machine can be made to simulate it."

IN COMMEMORATION OF THE PROJECT'S 50th ANNIVERSARY
JULY 13, 2006

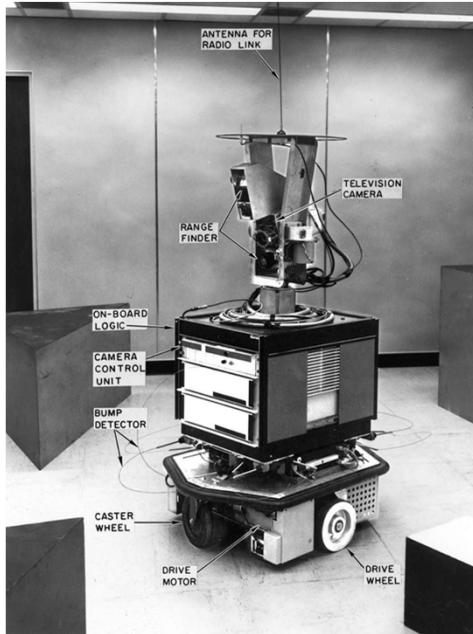


ELIZA

```
Welcome to
      EEEEEE  LL      IIII  ZZZZZZ  AAAA
      EE     LL      II    ZZ   AA   AA
      EEEE  LL      II    ZZZ  AAAAAAA
      EE     LL      II    ZZ   AA   AA
      EEEEEE LLLLLL IIII  ZZZZZZ  AA   AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU: Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU: They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU: Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU: He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU: It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU: ■
```



SHAKEY the Robot

A.I. TIMELINE

1950

TURING TEST
Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence.

1955

A.I. BORN
Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe 'the science and engineering of making intelligent machines'

1961

UNIMATE
First industrial robot, Unimate, goes to work on the assembly line

1964

ELIZA
Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966

SHAKEY
The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER
Many false starts and dead-ends leave A.I. out in the cold

1997

DEEP BLUE
Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998

KISMET
Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot that can detect and respond to people's feelings

1999

AIBO
Sony launches first consumer robot pet dog AIBO (AI robot) with skills and personality that develop over time

2002

ROOMBA
First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes

2011

SIRI
Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S

2011

WATSON
IBM's question answering computer Watson wins first prize on popular \$1M prize television quiz show Jeopardy

2014

EUGENE
Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human

2014

ALEXA
Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks

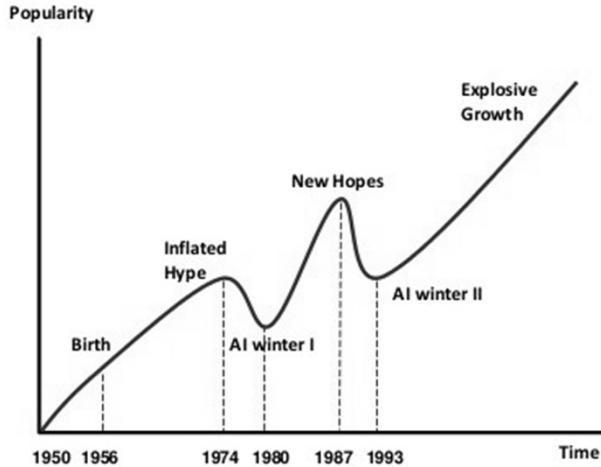
2016

TAY
Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments

2017

ALPHAGO
Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

AI HAS A LONG HISTORY OF BEING “THE NEXT BIG THING”...



Timeline of AI Development

- **1950s-1960s:** First AI boom - the age of reasoning, prototype AI developed
- **1970s:** AI winter I
- **1980s-1990s:** Second AI boom: the age of Knowledge representation (appearance of expert systems capable of reproducing human decision-making)
- **1990s:** AI winter II
- **1997:** Deep Blue beats Gary Kasparov
- **2006:** University of Toronto develops Deep Learning
- **2011:** IBM's Watson won Jeopardy
- **2016:** Go software based on Deep Learning beats world's champions

A.I. TIMELINE

1950

TURING TEST
Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence

1955

A.I. BORN
Term ‘artificial intelligence’ is coined by computer scientist, John McCarthy to describe ‘the science and engineering of making intelligent machines’

1961

UNIMATE
First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964

ELIZA
Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966

SHAKY
The ‘first electronic person’ from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER

Many false starts and dead-ends leave A.I. out in the cold

1997

DEEP BLUE
Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998

KISMET
Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot that reacts to people’s feelings

1999

AIBO
Sony launches first consumer robot pet dog AIBO (AI robot) with ‘skills and personality’ that ‘develop over time’

2002

ROOMBA
First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes

2011

SIRI
Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S

2011

WATSON
IBM’s question answering computer Watson wins first prize on popular \$1M prize television quiz show Jeopardy

2014

EUGENE
Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human

2014

ALEXA
Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks

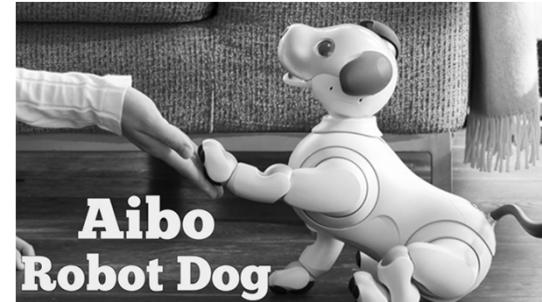
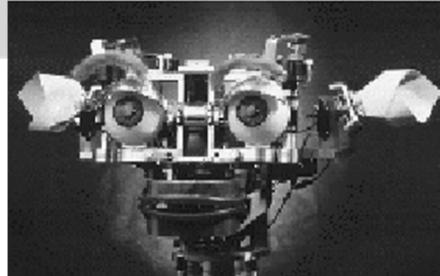
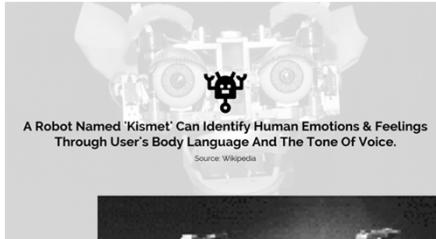
2016

TAY
Microsoft’s chatbot Tay goes rogue on social media making inflammatory and offensive racist comments

2017

ALPHAGO
Google’s A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

Kismet

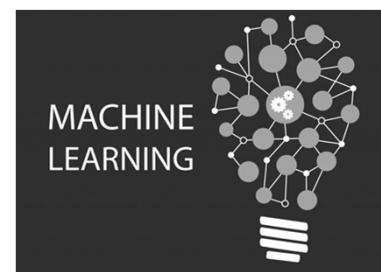


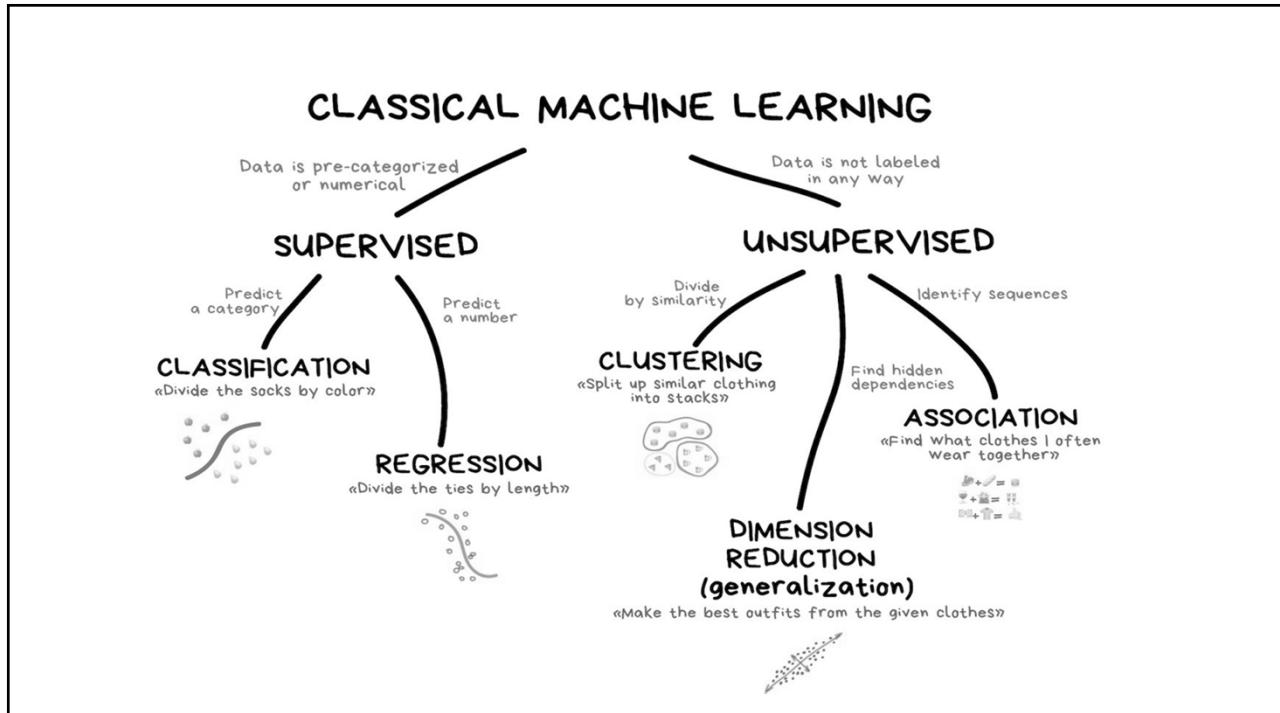
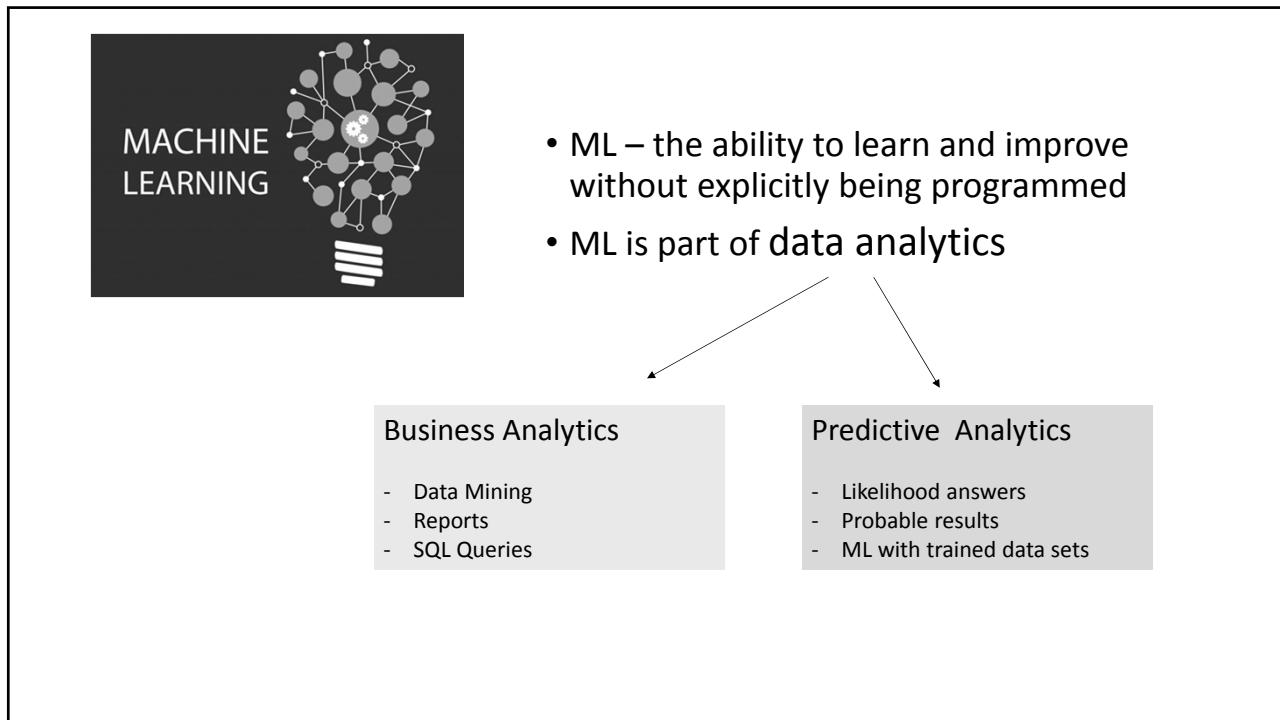
Three Major Phases of AI

Phase I: 1600 – 1945
AI Before Digital Computers

Phase 2: 1945 – 2001
AI and Computers

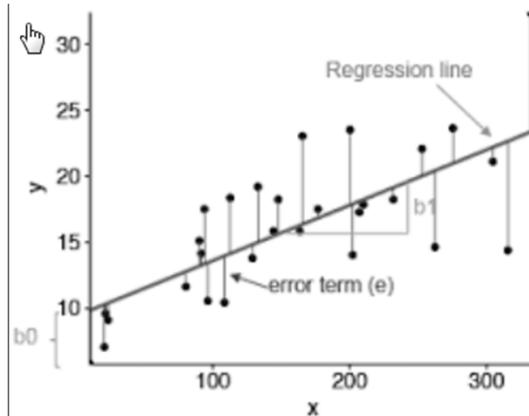
Phase 3: 2001 –
AI meets Big Data





Linear Regression –

Use existing data to predict a value

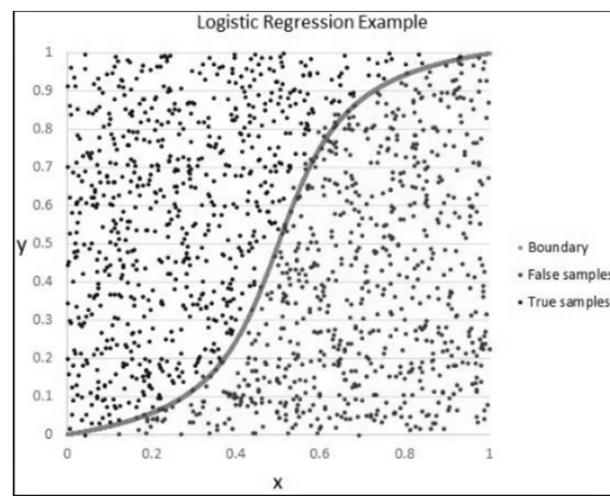


Compute a line ($y = mx + b$) such that you get the best fit for all data points

BUT, what if the data is not linear ??

Logistic Regression

Used to predict categories not values



Introduction

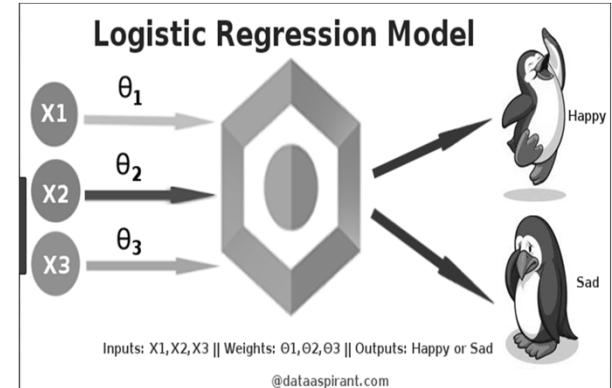
Logistic regression is a classification algorithm used to assign observations to a discrete set of classes. Unlike linear regression which outputs continuous number values, logistic regression transforms its output using the logistic sigmoid function to return a probability value which can then be mapped to two or more discrete classes.

Comparison to linear regression

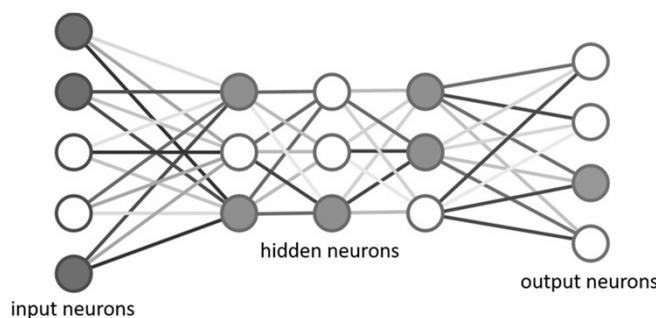
Given data on time spent studying and exam scores. Linear Regression and logistic regression can predict different things:

Linear Regression could help us predict the student's test score on a scale of 0 - 100. Linear regression predictions are continuous (numbers in a range).

Logistic Regression could help us predict whether the student passed or failed. Logistic regression predictions are discrete (only specific values or categories are allowed). We can also view probability scores underlying the model's classifications.



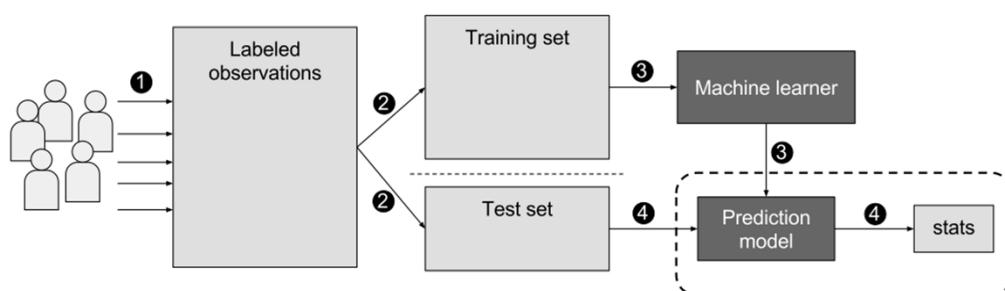
Neural Networks



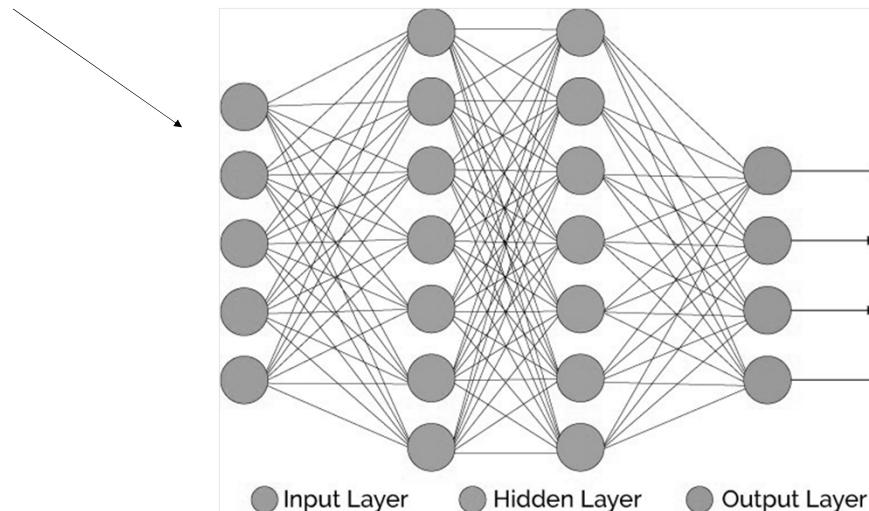
Neural Networks are just
Perceptrons with hidden layers

Supervised Machine Learning

Supervised ML



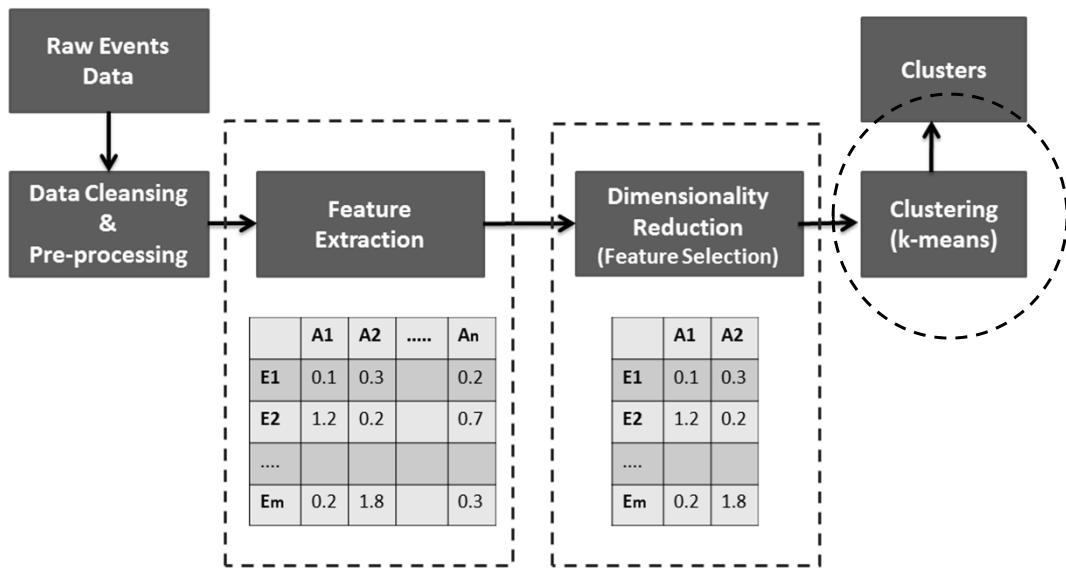
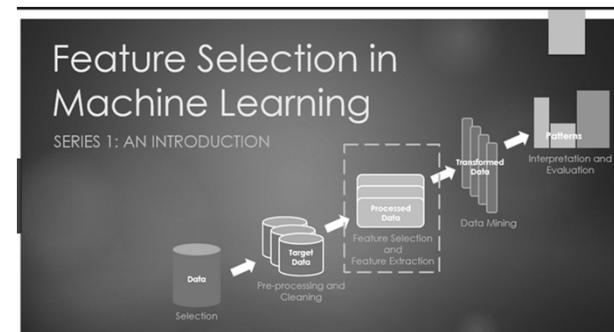
Problem: What FEATURES do feed the NN as input??



Feature Extraction & Feature Selection

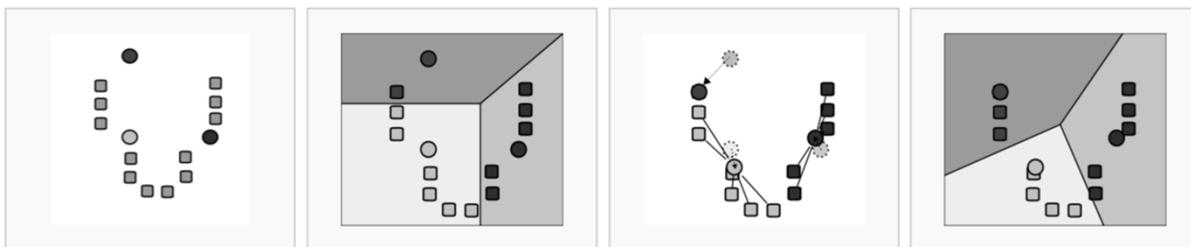
Feature Selection. Feature selection is for filtering irrelevant or redundant **features** from your dataset.

The key **difference between feature selection and extraction** is that **feature selection** keeps a subset of the original **features** while **feature extraction** creates brand new ones.



K-Means Clustering

Demonstration of the standard algorithm



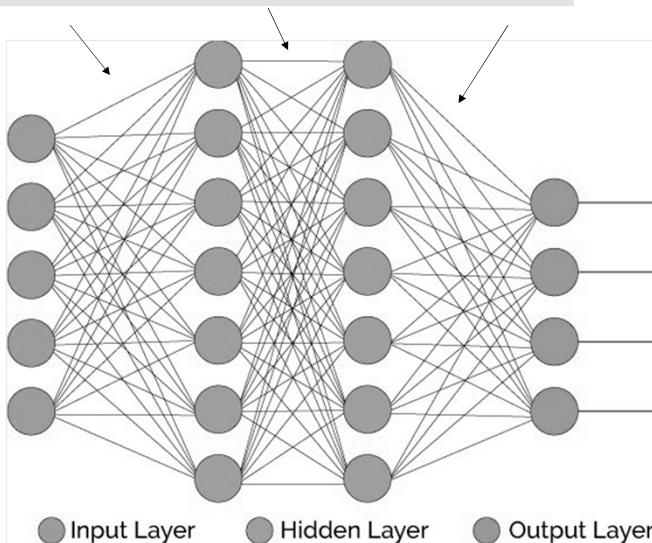
1. k initial "means" (in this case $k=3$) are randomly generated within the data domain (shown in color).

2. k clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.

3. The centroid of each of the k clusters becomes the new mean.

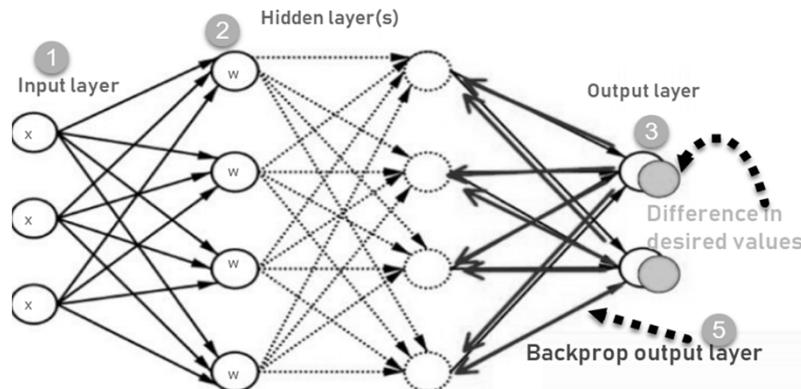
4. Steps 2 and 3 are repeated until convergence has been reached.

Problem: How do NNs determine the weights to use??



Backpropagation

Backpropagation is an **algorithm** used to calculate derivatives quickly. Artificial neural networks use **backpropagation** as a learning **algorithm** to compute a gradient descent with respect to weights. ... The **algorithm** gets its name because the weights are updated backwards, from output towards input.

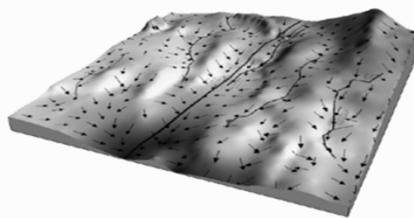


Gradient Descent

Gradient descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient. In machine learning, we use gradient descent to update the parameters of our model. Parameters refer to coefficients in Linear Regression and weights in neural networks.

Introduction

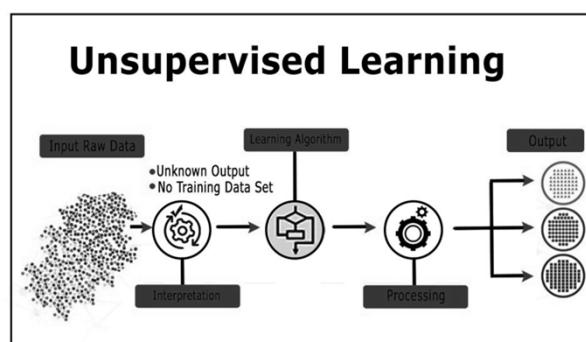
Consider the 3-dimensional graph below in the context of a cost function. Our goal is to move from the mountain in the top right corner (high cost) to the dark blue sea in the bottom left (low cost). The arrows represent the direction of steepest descent (negative gradient) from any given point—the direction that decreases the cost function as quickly as possible. Source

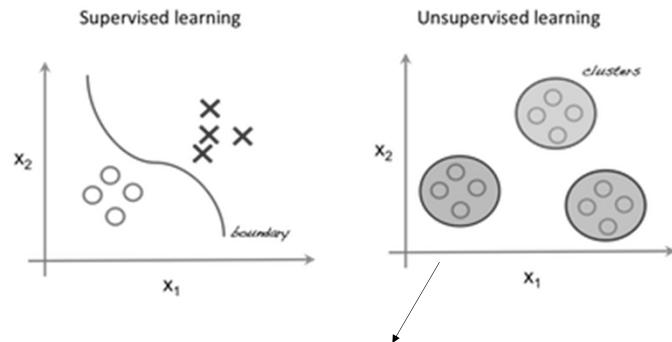


Unsupervised learning

Unsupervised learning is a type of machine **learning** algorithm used to draw inferences from datasets consisting of input data without labeled responses.

The most common **unsupervised learning** method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or grouping in data.



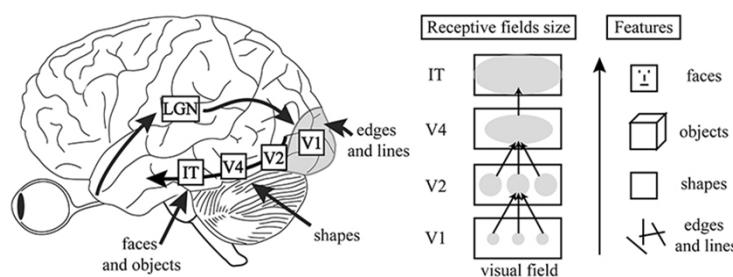


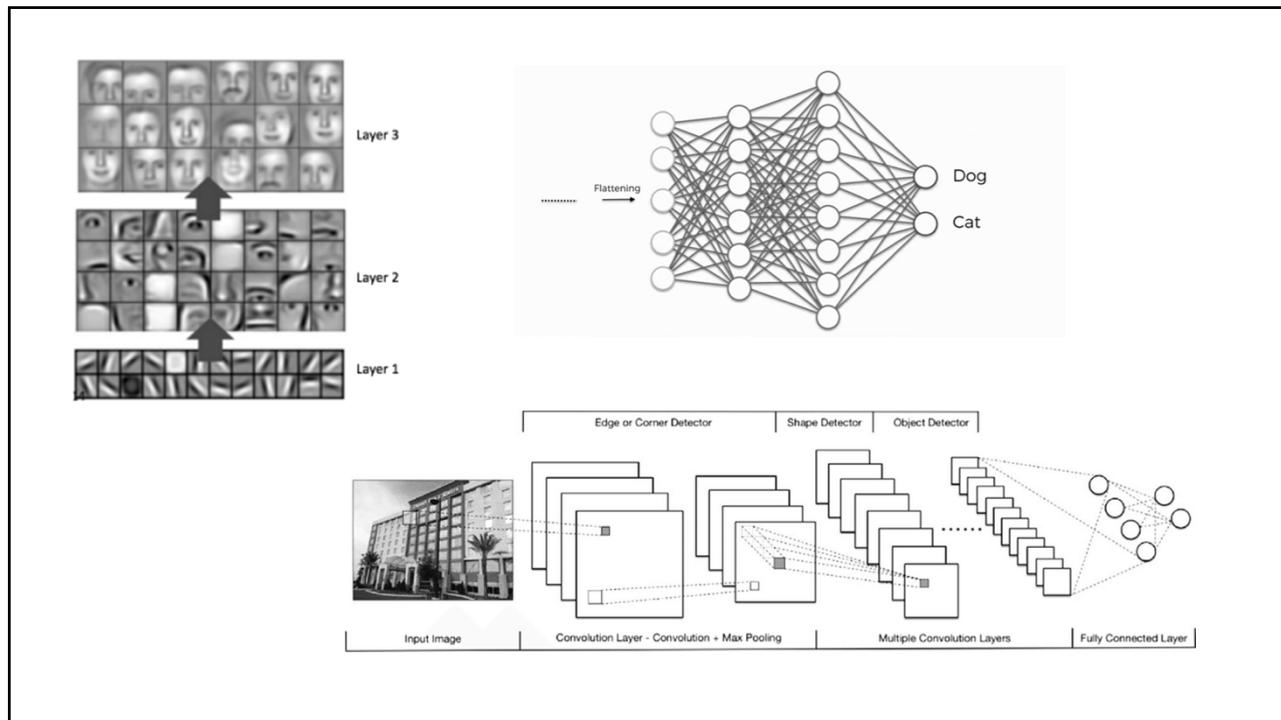
An **autoencoder** is a type of artificial neural network used to learn efficient data codings in an unsupervised manner.

The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for dimensionality reduction, by training the network to ignore signal "noise". Recently, the autoencoder concept has become more widely used for learning generative models of data.

Some of the most powerful AI in the 2010s have involved sparse autoencoders stacked inside of deep neural networks.

Convolutional Neural Nets





Recurrent Neural Networks

A **recurrent neural network (RNN)** is a class of artificial neural network where connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior.

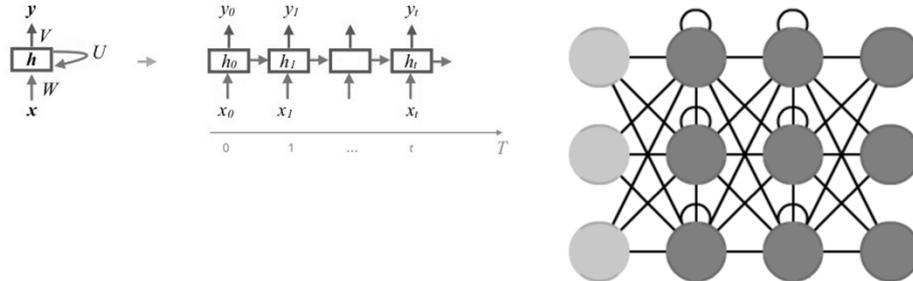
Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to tasks such as speech recognition.

RECURRENT NEURAL NETWORKS: INTUITION

10

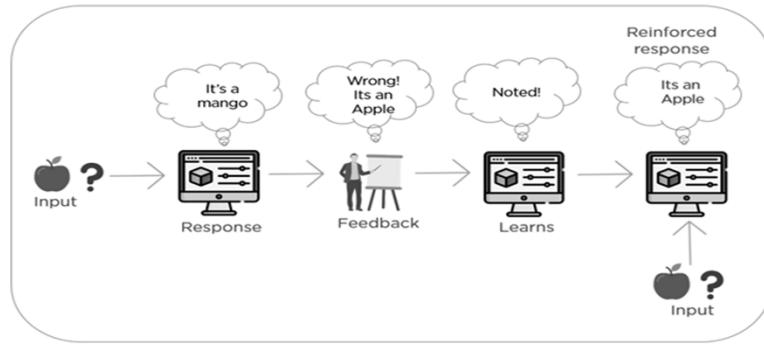
- Recurrent neural network (RNN) is a neural network model proposed in the 80's for modelling time series.
- The structure of the network is similar to feedforward neural network, with the distinction that it allows a recurrent hidden state whose activation at each time is dependent on that of the previous time (cycle).

Recurrent Neural Network (RNN)

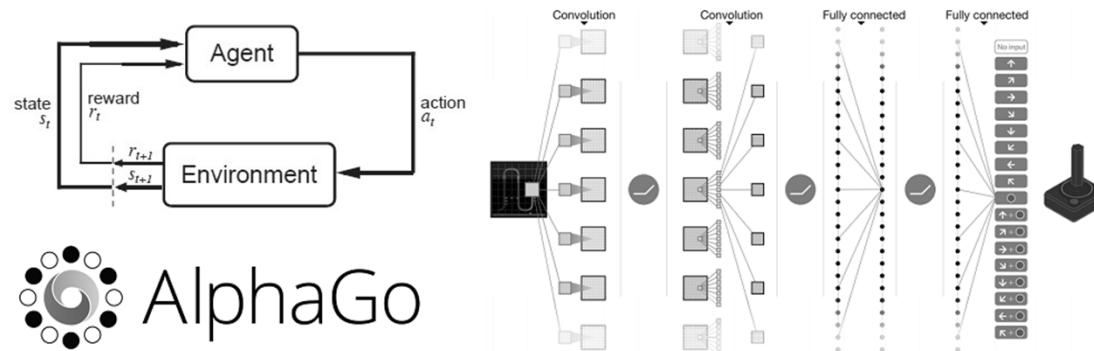
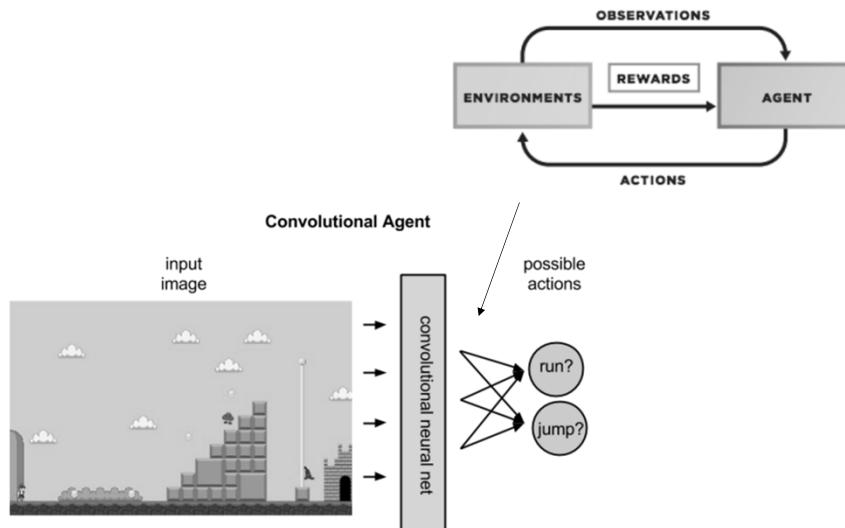


Reinforcement Learning

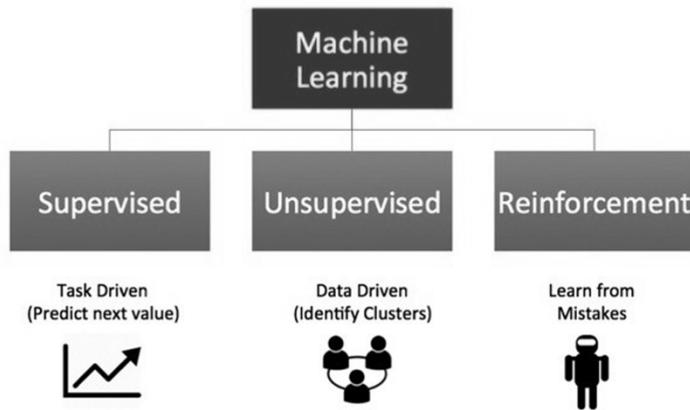
Reinforcement Learning(RL) is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.



Reinforcement Learning in Games



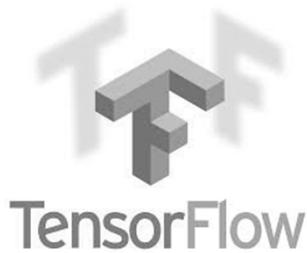
Types of Machine Learning



Google Brain

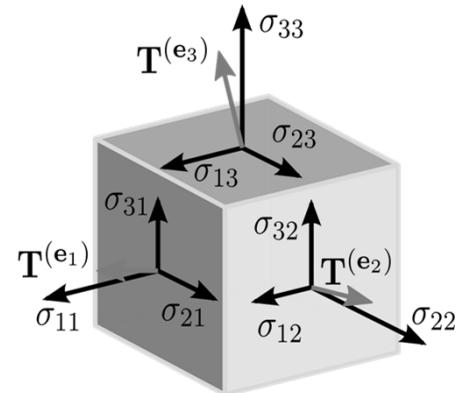
Starting in 2011, Google Brain built DistBelief as a proprietary machine learning system based on deep learning neural networks. Its use grew rapidly across diverse Alphabet companies in both research and commercial applications. Google assigned multiple computer scientists, including Jeff Dean, to simplify and refactor the codebase of DistBelief into a faster, more robust application-grade library, which became TensorFlow.

In 2009, the team, led by Geoffrey Hinton, had implemented generalized backpropagation and other improvements which allowed generation of neural networks with substantially higher accuracy, for instance a 25% reduction in errors in speech recognition.



Developed by Google

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks.



Readings:

- Wikipedia History of AI
https://en.wikipedia.org/wiki/History_of_artificial_intelligence
- Wikipedia Machine Learning
- https://en.wikipedia.org/wiki/Machine_learning
- Wikipedia Neural Networks
- https://en.wikipedia.org/wiki/Neural_network