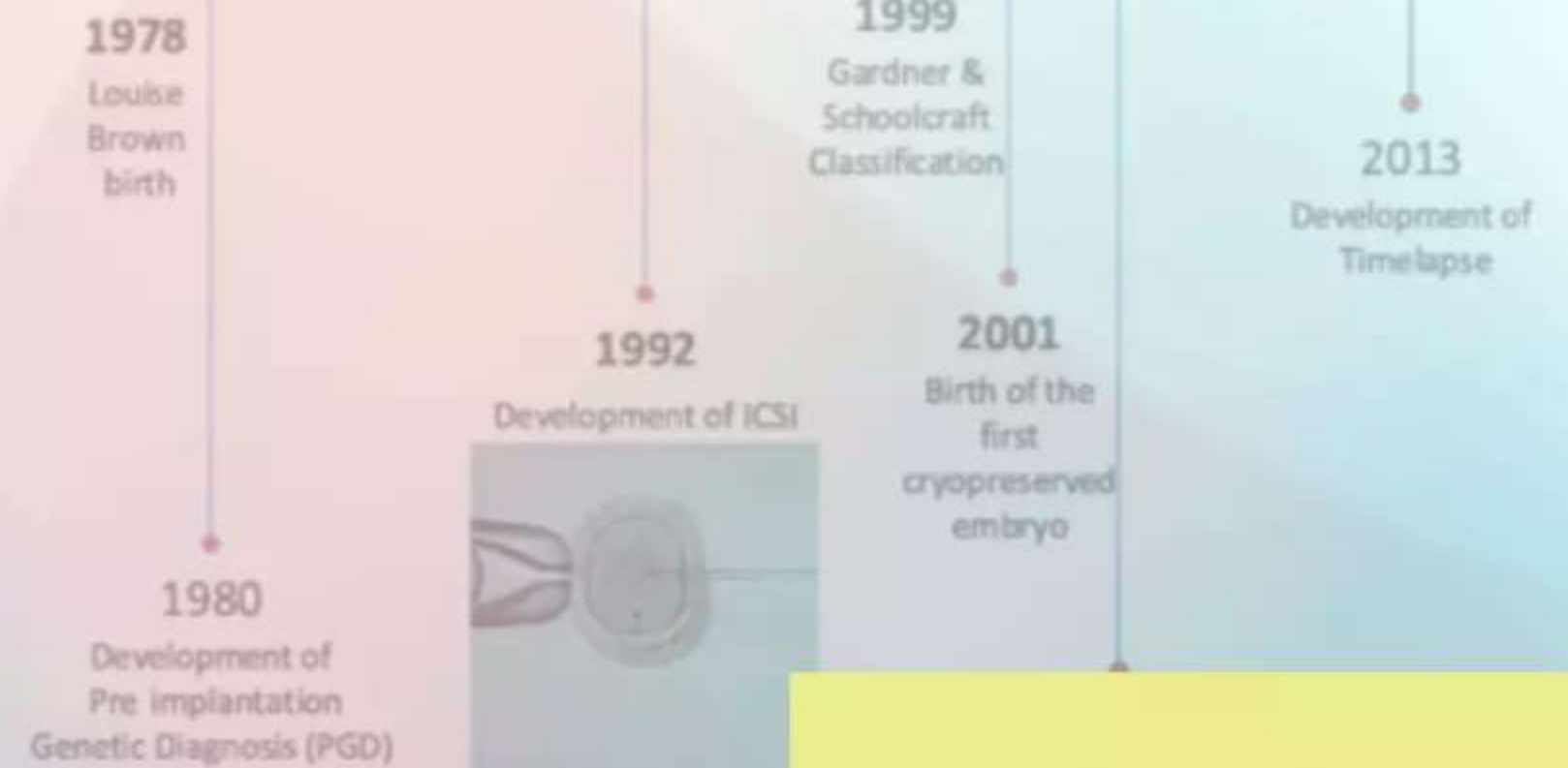


# History of AI

## From Birth Till Date



**AI**  
SHORTS





Can  
Machine  
Think?

## Symbolic AI

- The programs developed in the years after the Dartmouth Workshop were,
- Computers were solving algebra word problems, proving theorems in geometry and learning to speak English.

## Birth of AI

1952-1956

In the 1940s and 50s, a handful of scientists from a variety of fields (mathematics, psychology, engineering, economics and political science) began to discuss the possibility of creating an artificial brain.

## Boom

- Rise of Expert Systems
- The Knowledge Revolution
- The Money Return

1980-1987

1974-1980

## AI Winter

- Limited Computer Power
- There are many problems that can probably only be solved in exponential time
- Can still only handle trivial versions of the problems.
- The end of funding

1987-1993

## Bust : 2nd AI Winter

The collapse was due to the failure of commercial vendors to develop a wide variety of workable solutions. As dozens of companies failed, the perception was that the technology was not viable



## AI Renaissance

- AI was both more cautious and more successful than it had ever been.
- Deep Blue became the first computer chess-playing system

1993-2011

## Deep learning, big data and artificial general intelligence (AGI)

2011- Present

OpenAI is founded by a group of entrepreneurs

OpenAI releases a language model known as GPT-2

- OpenAI releases a new version of GPT-3
- OpenAI releases a tool known as DALL-E

OpenAI releases a new version of GPT-3 known as GPT-3 Prime

2015

2017

2018

2019

2020

2021

2022

OpenAI releases a language model known as GPT-1

Microsoft backed OpenAI with \$1Billion

OpenAI announces plans to develop and release GPT-3 under Open Source Licence



# 1936 - Alan Turing

## Conception

Alan Turing publishes "On Computable Numbers, with an Application to the Entscheidungsproblem,"

Introduces the concept of a universal machine capable of performing any computation that a human being can.



230

[Nov. 12,

### ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHEIDUNGSPROBLEM

*By* A. M. TURING.

[Received 28 May, 1936.—Read 12 November, 1936.]

The "computable" numbers may be described briefly as the real numbers whose expressions as a decimal are calculable by finite means. Although the subject of this paper is ostensibly the computable *numbers*, it is almost equally easy to define and investigate computable functions of an integral variable or a real or computable variable, computable predicates, and so forth. The fundamental problems involved are, however, the same in each case, and I have chosen the computable numbers for explicit treatment as involving the least cumbersome technique. I hope shortly to give an account of the relations of the computable numbers, functions, and so forth to one another. This will include a development of the theory of functions of a real variable expressed in terms of computable numbers. According to my definition, a number is computable if its decimal can be written down by a machine.

[https://www.cs.virginia.edu/~robins/Turing\\_Paper\\_1936.pdf](https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf)



# 1943 - Warren McCulloch and Walter Pitts

## First mathematical model of a neural network

Building on ideas by Alan Turing they built M-P model - it is considered to be one of the earliest examples of an artificial neural network which is a key technique used in machine learning.

The M-P model is not capable of learning from data. However, it provided an early inspiration for the development of more advanced neural network models.



BULLETIN OF  
MATHEMATICAL BIOPHYSICS  
VOLUME 5, 1943

### A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE,  
DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE,  
AND THE UNIVERSITY OF CHICAGO

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

<https://home.csulb.edu/~cwallis/382/readings/482/mcculloch.logical.calculus.ideas.1943.pdf>



# 1949 - Donald Hebb

## Hebbian Learning

1949: Donald Hebb publishes "The Organization of Behavior," which proposes the idea of Hebbian learning, a form of **unsupervised learning** that involves strengthening connections between neurons that fire together.



### *The Organization of Behavior*

A NEUROPSYCHOLOGICAL THEORY

D. O. HEBB

*McGill University*



# 1950 - Alan Turing

## Can Machine Think?

1950: Alan Turing publishes "Computing Machinery and Intelligence" which proposes the Turing Test, a method for determining whether a machine can exhibit intelligent behaviour equivalent to, or indistinguishable from, that of a human.



Computing Machinery and Intelligence

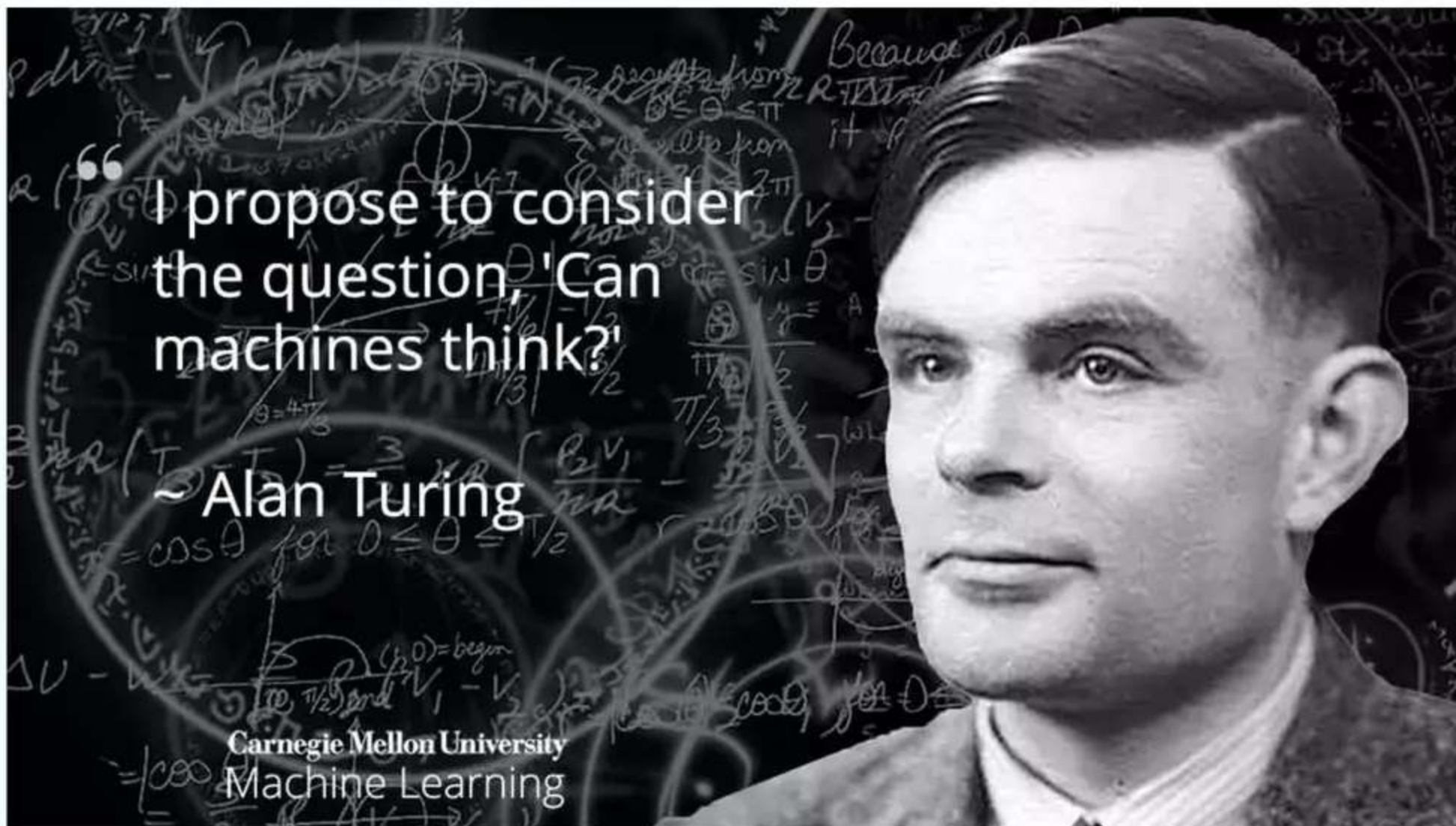
A. M. Turing

1950

## 1 The Imitation Game

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

<https://web.iitd.ac.in/~sumeet/Turing50.pdf>





# 1956 - John McCarthy

## Birth of AI

1956: John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon organize the **Dartmouth Conference**, which is considered to be the birth of AI as a field. The conference defines the goals of AI research and lays the groundwork for the development of early AI systems.

### 1956 Dartmouth Conference: The Founding Fathers of AI



John MacCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More





# 1958 - John McCarthy

## Lisp - Designed for AI

1958: John McCarthy invents Lisp, a programming language designed specifically for AI research.



Recursive Functions of Symbolic Expressions  
and Their Computation by Machine, Part I

John McCarthy, Massachusetts Institute of Technology, Cambridge, Mass. \*

April 1960

### 1 Introduction

A programming system called LISP (for LISt Processor) has been developed for the IBM 704 computer by the Artificial Intelligence group at M.I.T. The system was designed to facilitate experiments with a proposed system called the Advice Taker, whereby a machine could be instructed to handle declarative as well as imperative sentences and could exhibit “common sense” in carrying out its instructions. The original proposal [1] for the Advice Taker was made in November 1958. The main requirement was a programming system for manipulating expressions representing formalized declarative and imperative sentences so that the Advice Taker system could make deductions.

<http://www-formal.stanford.edu/jmc/recursive.pdf>





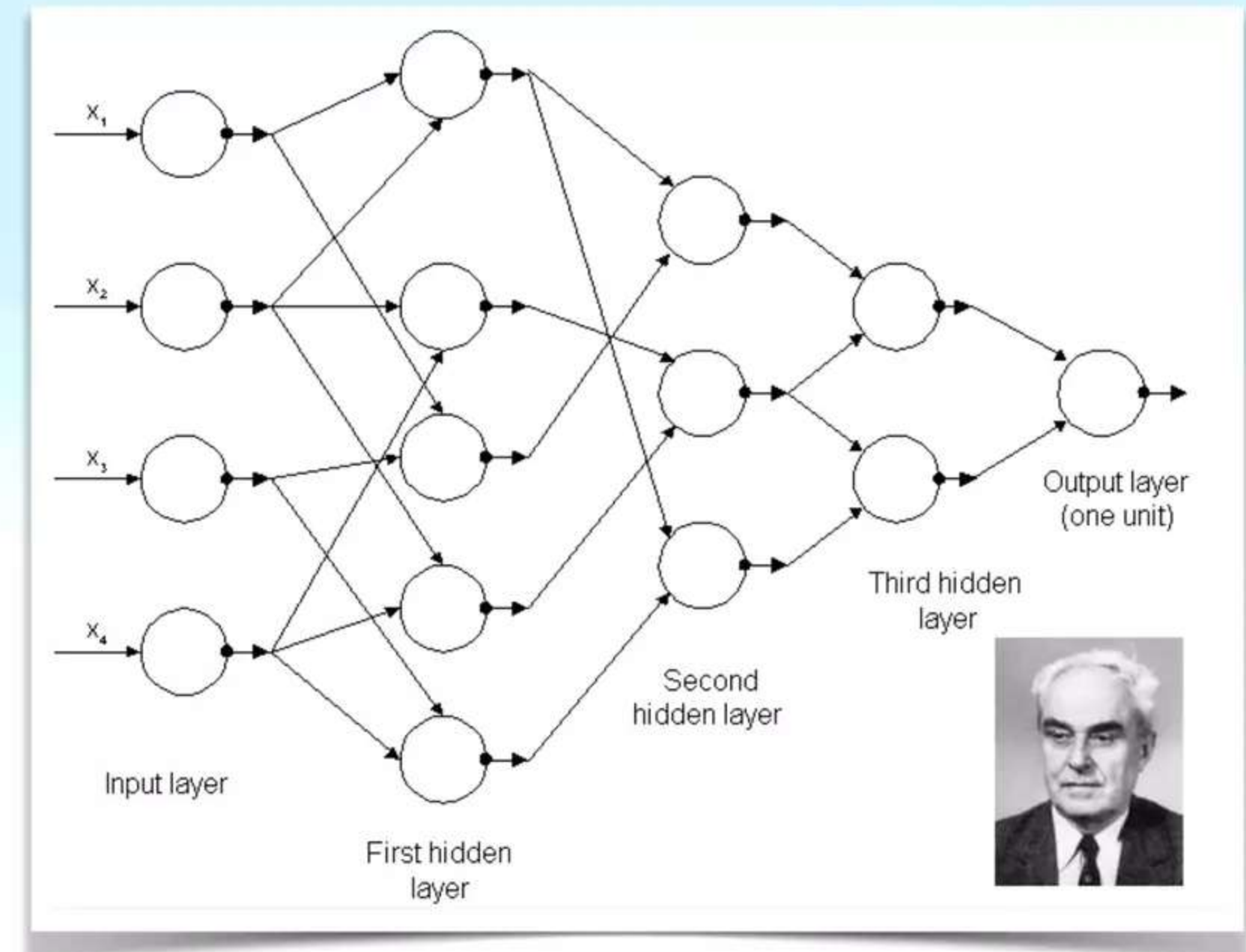
# 1965 - Anatolii Gershman, Alexey Ivakhnenko, Valentin Lapa

## Deep Learning - Multi Layered Perceptron

1965 : **The Group Method of Data Handling** (GMDH) is a data-driven approach to modeling that is based on a multi-layered architecture of interconnected polynomial models.

A multi-layer perceptron (MLP) is a type of artificial neural network (ANN) that is composed of multiple layers of interconnected nodes, or "neurons". MLPs are typically used for supervised learning tasks, such as classification and regression.

Ivakhnenko is often considered as the father of **deep learning**.





# 1966 - Arthur Samuel

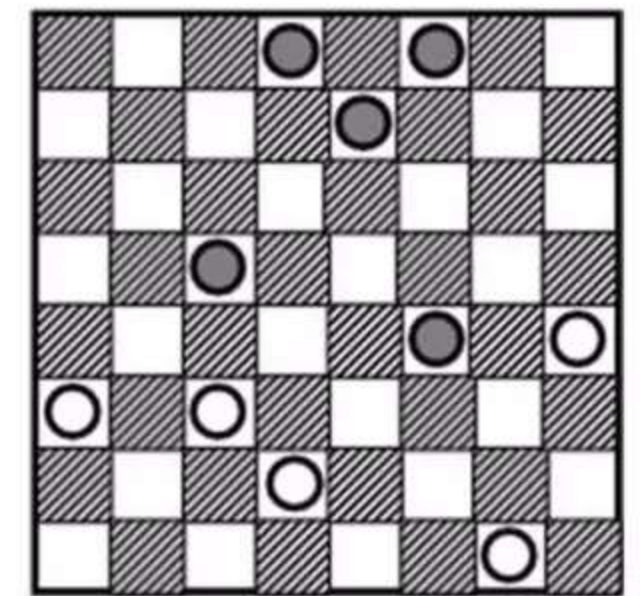
## Machine Learning

1966: Arthur Samuel is the first person to come up with and popularize the term "machine learning".

He defines it as the "field of study that gives computers the ability to learn without being explicitly programmed."



- Arthur Samuel (1959) wrote a program that **learnt** to play checkers well enough to beat him.





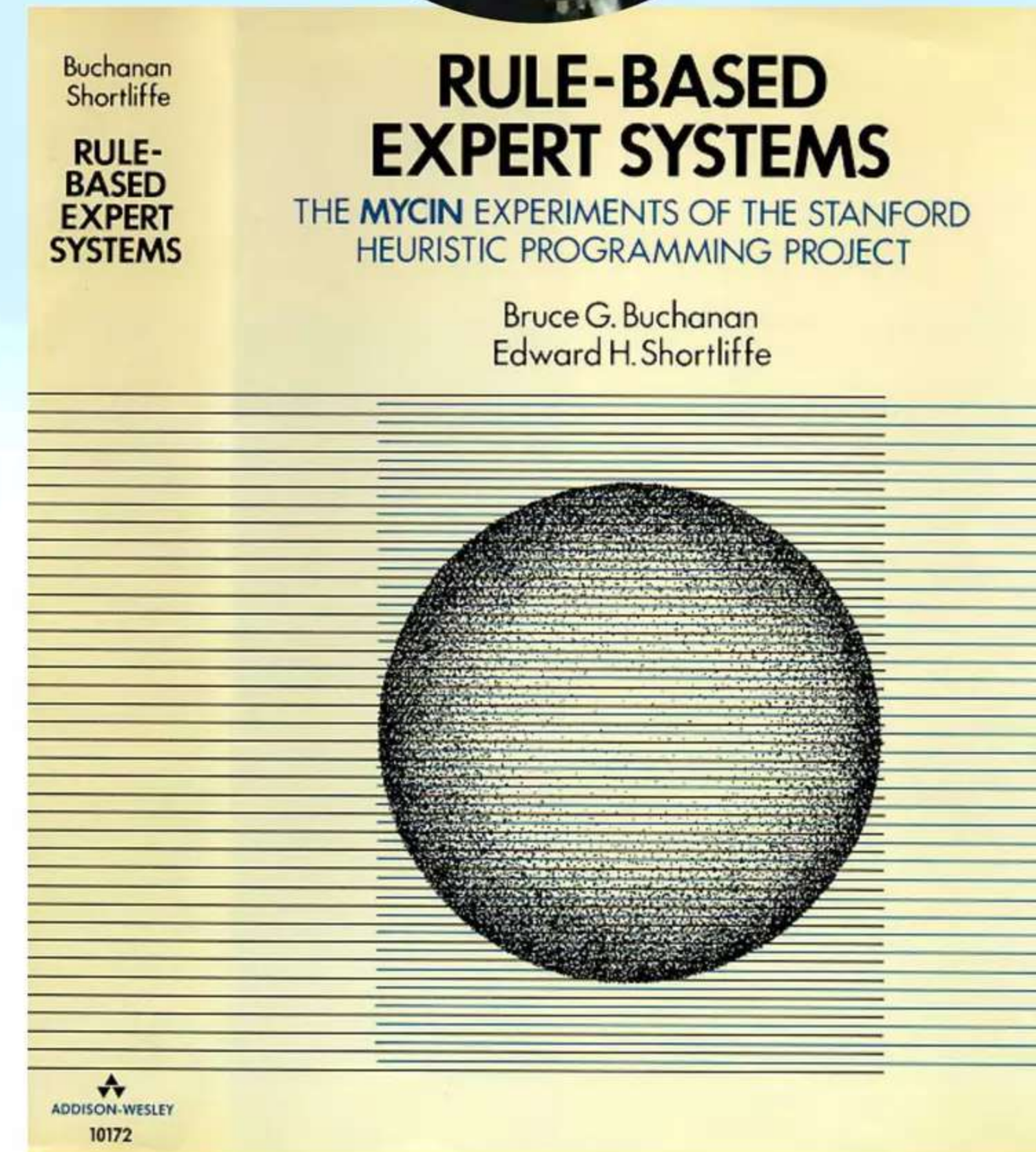
# 1974 - Edward Shortliffe

## First Application of AI

1974: The first AI program, known as the MYCIN system, is developed to assist physicians in diagnosing bacterial infections.

It was one of the first successful applications of AI in the field of medicine, and it helped to inspire further research and development in this area.

1980s: Expert systems, which are AI systems designed to emulate the decision-making abilities of a human expert in a specific field, become popular in business and industry.



*During AI winter*

**1974 - 1980 — 1st AI Winter**



# 1986 - Deep Neural Networks

## Backpropagation Algorithm

1986: Geoffrey Hinton, David Rumelhart, and Ronald Williams publish a paper on the backpropagation algorithm, which revolutionizes the field of neural network research and makes it possible to train deep neural networks.

### The Authors

- David E. Rumelhart  
- Stanford University
  - Geoffrey E. Hinton  
- University of Toronto
  - Ronald J. Williams  
- Northeastern University
- in
- 1986



E/15/325 Chalani

### 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

## 1987 - 1993 — 2nd AI Winter



# 1997 - Machine Defeated Human

## Deep Blue by IBM

1997: IBM's Deep Blue defeats world chess champion Garry Kasparov in a six-game match.

The chess machine won the second game after Kasparov made a mistake in the opening, becoming the first computer system to defeat a reigning world champion

A documentary film, Game Over - suggested that Deep Blue's victory was a trick by IBM to lift its stock value.





# 2006 - Geoffrey Hinton

## Improvement in Speech and Image Recognition

2006: Geoffrey Hinton and his team develop deep learning algorithms that significantly improve speech recognition and image recognition.

Deep Belief Networks, which allows for efficient and effective training of large-scale neural networks for machine learning tasks.



### A fast learning algorithm for deep belief nets \*

**Geoffrey E. Hinton and Simon Osindero**  
Department of Computer Science University of Toronto  
10 Kings College Road  
Toronto, Canada M5S 3G4  
{hinton, osindero}@cs.toronto.edu

**Yee-Whye Teh**  
Department of Computer Science  
National University of Singapore  
3 Science Drive 3, Singapore, 117543  
tehyw@comp.nus.edu.sg

### Abstract

We show how to use “complementary priors” to eliminate the explaining away effects that make inference difficult in densely-connected belief nets that have many hidden layers. Using complementary priors, we derive a fast, greedy algorithm that can learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm is used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm. After fine-tuning, a network with three hidden layers forms a very good generative model of the joint distribution of handwritten digit images and their labels. This generative model gives better digit classification than the best discrimi-

remaining hidden layers form a directed acyclic graph that converts the representations in the associative memory into observable variables such as the pixels of an image. This hybrid model has some attractive features:

1. There is a fast, greedy learning algorithm that can find a fairly good set of parameters quickly, even in deep networks with millions of parameters and many hidden layers.
2. The learning algorithm is unsupervised but can be applied to labeled data by learning a model that generates both the label and the data.
3. There is a fine-tuning algorithm that learns an excellent generative model which outperforms discriminative methods on the MNIST database of hand-written digits.
4. The generative model makes it easy to interpret the dis-



# 2011 - Thomas J. Watson

## Machine Beats Human in a game

2011: IBM's Watson defeats human champions in the game show Jeopardy!

Watson defeated Jennings and Rutter by a significant margin, winning a grand prize of \$1 million.



Two "Jeopardy!" champions, Ken Jennings, left, and Brad Rutter, competed against a computer named Watson, which proved adept at buzzing in quickly. Carol Kaelson/Jeopardy Productions Inc., via Associated Press



# 2012 - Google Brain

## Achieving Record Low Error Rate in Image Recognition

2012: Google Brain is a deep learning artificial intelligence research team under the umbrella of Google AI.

Google's deep learning algorithm, known as Google Brain, achieves a record-low error rate on an image recognition task.

2015: **TensorFlow** is an open source software library powered by Google Brain



Jeff Dean  
Google Brain Team



Rajat Monga  
CoFounder Tensorflow



# 2014-18 : Google acquires DeepMind

## Development of Deep Learning Algorithms

**2014:** Facebook creates its AI research division, and Google acquires DeepMind, an AI company that later develops the AlphaGo system that beats the world champion in the game of Go.

**2016:** Google's AlphaGo defeats world champion Lee Sedol in a five-game match.

**2018:** The development of deep learning algorithms for natural language processing leads to significant improvements in machine translation and other language-based AI applications.



AlphaGo's ultimate challenge: a five-game match against the legendary Lee Sedol

### Google Acquires Artificial Intelligence Startup DeepMind For More Than \$500M

Catherine Shu @catherineshu / 6:50 AM GMT+5:30 • January 27, 2014

Comment



Google will buy London-based artificial intelligence company [DeepMind](#). [The Information](#) reports that the acquisition price was more than \$500 million, and that Facebook was also in talks to buy the startup late last year. DeepMind confirmed the acquisition to us, but couldn't disclose deal terms.

<https://techcrunch.com/2014/01/26/google-deepmind/>



# 2015-2022 - Open API & ChatGPT3

## Large Language Model

2015: OpenAI is founded by a group of entrepreneurs - Elon Musk, Sam Altman, Reid Hoffman etc - they pledged \$1Billion

2017: OpenAI releases GPT-1

2018: OpenAI releases GPT-2

2019: Microsoft backed OpenAI with \$1Billion

2020: OpenAI releases a new version of GPT-3

2020: OpenAI releases a tool known as DALL-E

2021: OpenAI announces plans to develop and release GPT-3 under an open-source license.

2022: OpenAI releases GPT-3 Prime

115M

GPT1

1.5B

GPT2

175B

GPT3

~ In Trillion  
GPT-4

### Significance of Number of Params

These are tuneable variables that the model has learned during the training process.

More params means more flexibility in the model's ability to generate diverse and coherent text output