# Broad Research on the History of Machine Learning

Ilana Chaia Finger, Ananda Julia Galvão Campelo

November 26, 2023

**Abstract**

Machine learning, a dynamic field at the intersection of computer science and artificial intelligence, has undergone transformative phases since its inception. This article provides a thorough examination of the historical milestones in machine learning, from the conceptualization of neural networks in the 1940s to contemporary breakthroughs in deep learning. The narrative encompasses key figures such as Turing, Samuel, and Hsu, highlighting their contributions to shaping the landscape of intelligent machines. The notorious "AI winter" is explored as a period of skepticism and subsequent resurgence in the 1990s. As the discussion progresses to recent developments, the impact of open-source initiatives, like Torch, is underscored. The article concludes with a reflection on the current state of machine learning, emphasizing its ability to handle vast datasets and its promising trajectory as a pivotal field of study.

## 1 Introduction

Machine learning is a field of computer science that focuses on creating algorithms capable of learning from data. Over the last few decades, this subject has brought about a significant transformation in various research fields, such as natural language processing (NLP) and image recognition. This is possible because machine learning algorithms can enhance efficiency and accuracy, especially in tasks like predicting outcomes or interpreting data.

Machine learning utilizes various algorithm types, each with distinct strengths and weaknesses. Among these are algorithms that rely on a set of training data and the association of positive and negative feedback behaviors. It's essential to recognize that the field has undergone many phases of development since the inception of computers.

# 2  Early Developments

The first neural network was developed by neurophysiologist Warren McCulloch and mathematician Walter Pitts in 1943. They published a paper on how neurons might work, modeling a simple network with electrical circuits. The demonstration showed the possibility of two computers communicating without any human intervention.

## 2.1  Turing and the Imitation Game

Later, in the 1950s, mathematician Alan Turing proposed a solution to the question of whether machines can think. He created the 'Turing Test' in his paper 'Computing Machinery and Intelligence' while working at the University of Manchester. Turing chose to replace this question with another closely related one, expressed in relatively unambiguous terms, by describing a problem in the form of a game called 'The Imitation Game.'

Turing initially wrote about a simple game between a man and a woman in which they both go into separate rooms, and a third party tries to tell them apart by questioning each one. In this game, both the man and the woman try to convince the third party that they are the other:

> "We now ask the question, 'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?' " [6]

To test a digital computer's ability to mimic intelligent behavior equivalent to that of a human, the Turing Test is sometimes employed to establish specific conditions for the manifestation of mind, thought, or intelligence. The original concept involved a three-person game, wherein a human interrogator aimed to differentiate between a person and a machine. Each participant would be isolated in a room, and the interrogator would identify them using labels such as "A" and "B".

During the game, the interrogator would be allowed to question both the person and the machine through a text-only channel, so the result would not depend on the computer's ability to speak. If the interrogator could not distinguish between the human and the machine, the machine would have then passed the test.

The questions would be formulated not to assess whether the machine is capable of generating correct answers, but to analyze how closely its answers resemble those of a human. Questions such as "Does 'A' know how to play chess?" would be asked, and "A", whether it's the computer or the human, would respond. The person's objective would be to assist the interrogator in correctly identifying the computer, and the computer's objective would be to try to make the interrogator conclude that the machine is the other person. Regarding the game, Turing says:

"I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about $10^9$, to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning. ... I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted." [6]

## 2.2 Following achievements

### 2.2.1 IBM 704: Advancements in Alpha-Beta Pruning

Following that, in 1952, Professor Emeritus Arthur L. Samuel utilized tube-based computing for logic and memory in IBM's first stored program computer, the 704, based on Williams tubes. These tubes stored bits as charged spots on the screen of a cathode ray tube, significantly enhancing the memory capacities of machines. Samuel had the vision to develop a sophisticated program to play checkers:

"I became so intrigued with this general problem of writing a program that would appear to exhibit intelligence that it was to occupy my thoughts during almost every free moment for the entire duration of my employment by IBM and indeed for some years beyond." [5]

The Samuel Checkers-Playing Program implemented the first alpha-beta pruning algorithm, with a loss function that would calculate the probability of winning the game based on the current position. That is a very early example of a method now commonly used in artificial intelligence (AI) research.

### 2.2.2 Perceptron

In 1958, psychologist Frank Rosenblatt used an IBM 704 to create a machine that is "capable of having an original idea", the "Perceptron". The computer was fed a series of punch cards, and after 50 trials, it taught itself to distinguish cards marked on the left from cards marked on the right.

"A machine capable of perceiving, recognizing and identifying its surroundings without any human training or control", as Rosenblatt wrote, drew keen interest from reporters and the growing computing community. It was designed to improve the accuracy of computer predictions, learning from data by adjusting its parameters until reaching an optimal solution.

### 2.2.3 The Nearest Neighbor Algorithm

In 1967, the idea of nearest neighbor pattern classification was first introduced by Cover and Hart, with a method in which the nearest neighbor decision rule

assigns to an unclassified sample point the classification of the nearest of a set of previously classified points.

The Nearest Neighbor Algorithm was developed as a means to automatically identify patterns within large datasets. With the goal of finding similarities between two items and determining which one is closer to the pattern found in the other item, it can be used to discover mirror patterns between different pieces of data or predict future events based on past events.

## 2.3   AI Winter

What followed was a period during which interest, research activities, and funding for AI projects significantly decreased. The term "AI winter" began as funds dried up following unmet early promises. The term explicitly referenced nuclear winters, suggesting that AI research was in such disarray that it would not receive attention for many years.

The AI winter occurred between 1970 and 1990. Initially happening during the Cold War, with high expectations to automate the task of language translation, but no outcomes were achieved. The consequences of many unfulfilled promises caused this dark era in machine learning development.

Criticism came from everywhere, and the AI community became very pessimistic about developments, with one of the major reasons being the collapse of the LISP-based hardware business. However, research later found the reasons for such failure in this early stage of AI, including the lack of experience with algorithms, computing capacity, and, most importantly, the availability of data.

## 2.4   The Resurgence of AI

The early 1990s showed impressive strides forward in AI research, and it was a period of rapid growth and interest. The surge in interest was followed by a surge in funding for research, allowing even more progress to be made.

### 2.4.1   AI in chess and Voice Recognition

Notably, in 1997, then-student Feng-hsiung Hsu created a chess-playing expert system, run on a unique IBM supercomputer, the Deep Blue. It was the first time a machine had beaten a world champion at chess, making it a landmark event that showed that AI systems could surpass human understanding in complex tasks.

Later that same year, Dr. James Baker released Dragon Naturally Speaking, the first voice recognition software to detect continuous speech. The system was released to the public as Dragon Dictate for DOS-based computers and made use of the Markov Model for voice recognition, and it has come a long way since then.

4

### 2.4.2  First Open-Source Library

In 2002, the first open-source machine learning library, Torch, was released, providing interfaces to deep learning algorithms implemented in C. It was created by computer scientists Geoffrey Hinton, Pedro Domingos, and Andrew Ng with the belief that their specific needs were not met by other libraries. Hinton would then publish a paper on "A Fast Learning Algorithm for Deep Belief Nets", which many consider to be the birth of deep learning, describing the first deep learning algorithm that can achieve human-level performance on difficult and complex pattern recognition tasks.

## 3  Current State and Future Prospects

The present Machine Learning algorithms are able to handle large amounts of data with accuracy in a relatively short amount of time. Used in many different fields and able to learn patterns and relationships between data, find predictive insights for complex problems, and extract information that is otherwise too difficult to find, it has seen rapid progression since its inception and continues to be one of the most promising fields of study.

The speed of Machine Learning development will continue to increase as technology becomes more sophisticated, with the rate of progress depending on the amount of investment and research applied. With the right dataset, Machine Learning can be used to achieve almost anything.

## 4  Conclusion

In summary, this analysis traces the evolution of machine learning, emphasizing its impact on computational intelligence. From early neural networks to contemporary deep learning, the field showcases technological advancements and ongoing scientific commitment. Propelled by open-source collaboration and deep learning breakthroughs, machine learning's future seems promising. Its role as a driving force in shaping artificial intelligence is underlined by its ability to extract insights from complex data and adapt to advancing technology.

## References

[1] Alexander L. Fradkov. Early history of machine learning. *IFAC-PapersOnLine*, 53(2):1385–1390, 2020. 21st IFAC World Congress.

[2] Robert M French. The turing test: the first 50 years. *Trends in Cognitive Sciences*, 4(3):115–122, 2000.

[3] Giuliano Pezzolo Giacaglia. *Making Things Think: How AI and Deep Learning Power the Products We Use.* Holloway, e1.0.2 edition, 2022.

[4] Robert Koch. History of machine learning – a journey through the timeline. *clickworker*, 2022.

[5] A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):210–229, 1959.

[6] Alan M. Turing. Computing machinery and intelligence. *Mind*, 59(236):433–460, 1950.