

Prologue

Lest we forget, the seeds of innovation are often sown in the most unexpected times, from minds that dare to dream beyond the present. In the grand tapestry of human endeavor, some threads shimmer with an almost prophetic luminescence, foretelling futures that once seemed to be the stuff of fantasy. Join us as we embark on a captivating journey through the annals of Machine Learning, a field that has transformed the once far-fetched dream of the 18th century into the palpable reality of the 21st century. From the imaginative foresight of Jonathan Swift to the wartime brilliance of Alan Turing and his code-breaking team at Bletchley Park, we trace the lineage of ideas that led to the creation of the very first programmable computer, "Colossus." This is a story of how humanity, even in its darkest hours, pushed the boundaries of what was thought possible, ushering in a new era of automation and intelligence. We will witness the pioneering work of individuals like Arthur Samuel, who coined the term "machine learning" in the 1950s, laying the groundwork for machines that could learn from experience and challenging the very notion of human exclusivity in skill and adaptation. This prologue invites you to explore the fascinating evolution of Machine Learning, from its philosophical inception to its tangible applications, and to ponder the profound questions it raises about the nature of learning itself.

Introduction

LET me begin with a quote taken from a famous children's classic, 'Gulliver's Travels', by Jonathan Swift (1726) [5].

Everyone knew how laborious the usual method is of attaining to arts and sciences; whereas, by his contrivance, the most ignorant

person, at a reasonable charge, and with a little bodily labour, might write books in philosophy, poetry, politics, laws, mathematics, and theology, without the least assistance from genius or study.

This may be the first formal definition of modern computer applications and the visualization of the 21st century. This once-far-fetched dream of the 18th century is now a reality of the 21st century. You name it — ChatGPT, Co-pilot, Gemini, Meta-AI; the list is long and fulfills the prophecy of Jonathan Swift.

Now we travel a bit from the 18th century to the mid-19th century, when the world was shaking under Nazi boots, and humanity was fighting for survival. Everyone in the free world was giving their best effort to contribute to various fields of science, technology, agriculture, and industrial production. At that time, a name emerged with a different aura, **Alan Turing**.

In the 18th-century English country house of Bletchley Park, Alan Turing and his team, including Harry Golombek, John Tiltman, Gordon Welchman, Dilwyn Knox, Hugh Alexander, Donald Michie, Bill Tutte, and Stuart Milner-Barry, aimed to break the secret German communication using *Enigma* and *Lorenz*. They devised an automatic machine that deciphered and decrypted German communication, shortening the war by at least a year or two and saving many lives. The culmination of that great effort led to the world's first programmable computer "*Colossus*". The prophecy of Jonathan Swift came to light in 1946. In 200 years, a prophecy ushered us into a new era.

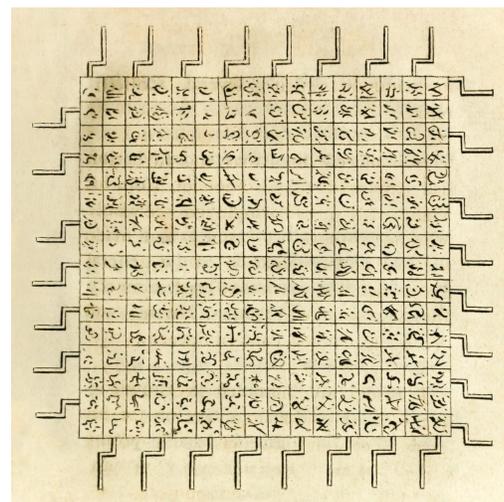


Figure 1: Illustration of The Engine from an edition of Gulliver's Travels (Photo Courtesy: Google)



Figure 2: The Bletchley Park Team. (Photo Courtesy: Google)

Alan Turing: A Brief History

ALAN Mathison Turing (1912 – 1954) was a British luminary whose profound insights forged the very foundation of modern computer science and artificial intelligence. His life, a tapestry woven with threads of dazzling intellect, crucial wartime service, and heartbreaking injustice, left an indelible mark on the 20th century and beyond. Even in his formative years in London, Turing's mind shone with extraordinary brilliance, particularly in the abstract realms of mathematics and science.

Unfettered by the rigid confines of traditional education, he embarked on a self-directed intellectual journey. This relentless pursuit of knowledge culminated in a first-class mathematics degree from King's College, Cambridge in 1934. It was within the esteemed halls of a subsequent fellowship that his genius truly blossomed, giving rise to his revolutionary work on computability and the now-iconic "Turing Machine." In his seminal 1936 treatise, "On Computable Numbers, with an Application to



Figure 3: Alan Turing. (Photo Courtesy: Google)

the *Entscheidungs problem*,” Turing unveiled his concept of a theoretical "Universal Turing Machine" — a stroke of pure intellectual alchemy that transmuted abstract logic into the very blueprint of the digital age [6, 7]. This elegant model not only illuminated the inherent limits and boundless potential of algorithmic computation but also furnished a rigorous definition of an algorithm and the conceptual genesis of the digital computer itself.

The exigencies of World War II thrust Turing’s exceptional abilities onto the global stage at Bletchley Park, the clandestine nerve center of British code-breaking operations. There, his intellectual prowess proved nothing short of transformative. He was the linchpin in cracking the seemingly impenetrable German Enigma code, the *lingua franca*¹ of Nazi military communications. His ingenious creation, the *Bombe* [2, 3] – an electromechanical marvel – shattered the Enigma’s complex encryptions with unprecedented speed. Turing’s wartime contributions are not merely significant; they are widely credited with altering the course of the war, potentially shortening its duration and sparing untold lives. His brilliance further extended to unraveling the intricacies of the German naval Enigma and the development of *Delilah*, a secure voice communication system. In recognition of his extraordinary service to the nation, he was awarded the Order of the British Empire in 1945.

Machine Learning: The term coined

THE evolution of Machine Learning started with the work of Alan Turing and his team at Bletchley Park and the automation they did to break the German Enigma and Lorenz. The term *Machine Learning* was coined by IBM’s **Arthur Samuel** around

¹A language that is adopted as a common language between speakers whose native languages are different.

the 1950s. Samuel's objective was to design a computer system that plays checkers better than its creator. His paper demonstrates that machine learning allows a computer to rapidly exceed a human programmer's checkers skill within hours [4]. Provided with just the rules, a guiding principle, and fundamental parameters, the system quickly masters the game, highlighting the versatile nature of these learning methods. The idea behind this work was that machines could learn from experience. This brings us to a more fundamental question of *Learning*. In the books, there are many formal explanations for this word. In a simple view, we can understand that learning is,

“adhere with the new situation and adjust accordingly”

This brings us some insight:

Can a machine adapt to and interact with the physical world like a human or any living being?

Let's begin by considering a few learning objectives. For this reason, we primarily use a broad definition of learning, subsequently incorporating computer programs that enhance performance through experience.

Data Science and Machine Learning:

DATA Science stands as a comprehensive, multidisciplinary field, acting as an umbrella term that integrates various cutting-edge concepts. These include the management and analysis of *big data*, the development of artificial intelligence (AI) systems, the extraction of hidden patterns through *data mining*, and the enabling of systems to learn from data via *machine learning*. The discipline of data science, focused on studying large

volumes of information, is a relatively recent development. Its emergence and growth are intrinsically linked to the evolution and widespread adoption of computers. Before the digital era, data analysis was performed manually, a practice that fell under the domain of statistics, or what we might now term classical statistics.

Consider, for instance, the calculation of cricket batting averages. This statistical measure existed long before computers became commonplace. Anyone equipped with a pencil, a notepad, and fundamental arithmetic skills could determine a player's seasonal batting average using classical statistical methods. I can cite another exciting incident from the life of Dr. Prannoy Roy, the former executive co-chairperson of NDTV and one of its co-founders. In the 1960s, his dedicated effort of sitting by the radio with a pen and diary to manually note down election updates was, in essence, a form of real-time data collection and analysis, albeit on a personal scale and reliant on manual methods. This painstaking process mirrors the "classical statistics" approach – data was valuable, insights were sought, but the tools were rudimentary by today's standards. This earlier form of manual data collection for election analysis sharply contrasts with current practices, where computers now rapidly generate sophisticated insights, including what we know as exit polls.

Indeed, statistics continue to be a staple in school curricula, much as they have been for centuries. While advanced levels of classical statistics exist, the datasets generally remain within a scope that humans can manage and process. However, the landscape changes dramatically when the need arises to calculate data at a significantly higher velocity (frequency), in much larger volumes, and with greater intrinsic value or complexity.

Imagine attempting to perform continuous calculations on one's heartbeat, and further, analyzing how that heartbeat responds

to variables such as temperature fluctuations and caloric intake. Such an endeavor is far beyond the capabilities of mental arithmetic or manual paper-based calculations. Moreover, the practicalities of collecting such detailed and voluminous data without technological aid would be prohibitive. This is precisely where the information age and the advent of powerful computers have profoundly transformed the field of statistics. Modern computing technology now provides the essential infrastructure to collect, store, and, most importantly, extract meaningful insights from massive and complex datasets, heralding the era of data science.

Any computer program is learning from an experience (ϵ), with respect to some specific task (τ) that is measured by performance (ρ), improves with experience (ϵ).

Some Classical Learning Problems:

LET US go back to the checkers problem as proposed by Samuel Land and try to understand this from the perspective of the definition of learning mentioned above.

The tasks involved here are to play the checkers games, through an experience acquired by playing games against itself. The perspective of defining a learning problem is having a task (objective) (τ), a performance measure (ρ), and training experience (ϵ). We can then formally write the playing checkers as:

- Tasks (τ) denotes "Playing Checkers".
- Performance Measure (ρ) denotes "percentage of games".
- Experience (ϵ) denotes "Playing practice games".

Handwriting recognition, especially at the word level, presents

a complex challenge in machine learning. This process involves training models to accurately interpret handwritten text, which requires addressing issues such as variability in writing styles, noise, and the inherent ambiguity of handwritten characters. Machine learning algorithms can segment words, recognize individual characters, and ultimately reconstruct the written text [1]. Handwritten word recognition problem can also be viewed from

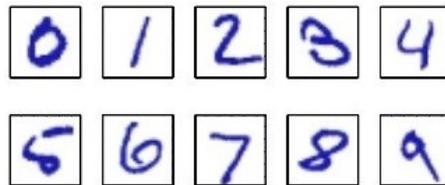


Figure 4: Hand-writing recognition, which is a kind of pattern recognition.

the perspective of the above definition:

- **Tasks (τ)** denotes “identify and classify handwritten words”.
- **Performance Measure (ρ)** denotes “percentage of accurately classified handwritten words”.
- **Experience (ϵ)** denotes “database of handwritten words with given classification”.

Arthur Samuel, A brief bio:

A RTHUR Lee Samuel (1901–1990) fundamentally shaped the nascent fields of computer gaming and artificial intelligence, famously coining the term “machine learning” in 1959. His early career saw him contribute to the development of vacuum tube technology and advancements in radar at Bell Laboratories, including a crucial gas-discharge transmit-receive switch during World War II.

After a period at the University of Illinois, where he launched the ILLIAC project, Samuel joined IBM in 1949. There, his innovations truly flourished. He is credited with one of the first software hash tables and played a pivotal role in early research on transistorized computers. His most sensational achievement was the development of the first checkers program on the IBM 701, a groundbreaking demonstration that famously caused IBM's stock to surge overnight. Samuel's pioneering work in non-numerical programming was instrumental in shaping processor instruction sets. A gifted communicator, he was celebrated for demystifying complex technical subjects. Following his retirement from IBM in 1966, Samuel continued his impactful work at Stanford University, even contributing to Donald Knuth's TeX project well into his late 80s.

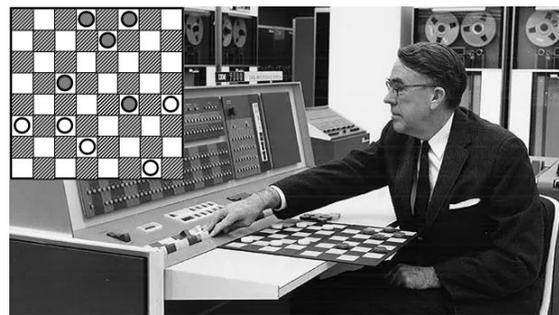


Figure 5: Arthur Samuel with his checker box problem (Photo Courtesy: Google).

In the next issue, we will delve deeper into the understanding of the concept of machine learning from the perspective of regression analysis.

Thought of the day:

Find the Task, Performance measure, Training experience with proper explanation for Autonomously Driven Car Problem!

References

- [1] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*. Vol. 4. 4. Springer, 2006.
- [2] Frank Carter. *The turing bombe*. Bletchley Park Trust, 2008.

- [3] William B Hayler and Hugh Sebag-Montefiore. “Enigma: The Battle for the Code”. In: *Naval War College Review* 54.4 (2001), p. 26.
- [4] Arthur L Samuel. “Some studies in machine learning using the game of checkers”. In: *IBM Journal of research and development* 3.3 (1959), pp. 210–229.
- [5] Jonathan Swift and Jonathan Swift. *Gulliver’s travels*. Springer, 1995.
- [6] Alan Turing. “Turing machine”. In: *Proc London Math Soc* 242 (1936), pp. 230–265.
- [7] Alan Mathison Turing et al. “On computable numbers, with an application to the Entscheidungsproblem”. In: *J. of Math* 58.345-363 (1936), p. 5.