

intelligence, such as reasoning, learning, problem-solving, perception, and language understanding. AI systems can be classified into narrow AI (designed for specific tasks like voice recognition or image classification) and general AI (designed to perform any intellectual task that a human can do). There are essentially two levels of AI in terms of their levels of complexity and learning capabilities, namely machine learning (ML) and deep learning (DL) (see Figure 5.4).

Machine learning (ML), as a subset of AI, focuses on the development of algorithms and statistical models that enable computers to learn from and make predictions or decisions based on data. Unlike traditional programming, where explicit instructions are provided, ML algorithms identify patterns and relationships in data to build models. Deep learning (DL) is a subset of ML that uses neural networks with many layers (hence 'deep') to model and understand complex patterns in data. These neural networks, inspired by the structure and function of the human brain, consist of interconnected nodes (neurons) that process input data through multiple layers. Each layer extracts higher-level features from the raw input, enabling the network to learn intricate patterns and representations. Deep learning can complete tasks involving unstructured data, such as image and speech recognition and natural language processing. It requires large amounts of data and computational power for training but can achieve remarkable accuracy and performance in various applications. Within DL, one application is generative AI (GenAI) (such as Claude AI (Anthropic), ChatGPT (OpenAI), Google Gemini, Microsoft Copilot, Midjourney), which leverages deep learning techniques to create models capable of generating new data.

**Artificial Intelligence (AI)**

*To mimic human behaviour via digital computers*

**Machine Learning (ML)**

*To identify patterns and automate analytical model building*

**Deep Learning (DL)**

*To model and understand complex patterns in data via multiple ("deep") layers (an application of DL is generative AI)*

**Figure 5.4** Simplified AI terminology

Source: Own illustration following Sarker (2021, p. 420); Turing (1950); and Wietschel et al. (2022)

---

## Alan Turing and the Dawn of AI

---

In his seminal paper 'Computing Machinery and Intelligence' from 1950, Alan Turing explores the question 'Can machines think?' He proposes replacing this question with a more practical one: 'Can machines play the imitation game?' This game involves an interrogator who must distinguish between a human and a machine based solely on their responses to questions. If the machine's responses are indistinguishable from the

human's, it can be said to 'think.' Turing discusses the capabilities and limitations of digital computers, which he defines as discrete-state machines capable of performing any operation a human computer can. He emphasises the universality of digital computers, meaning they can mimic any other human 'machine' given sufficient storage and speed.

Important quotes from Turing's (1950) paper:

- 'The idea behind digital computers may be explained by saying that these machines are intended to carry out any operations which could be done by a human computer.'
- 'The reader must accept it as a fact that digital computers (...) can in fact mimic the actions of a human computer very closely.'
- 'Sometimes such a machine is described as having free will.'
- 'The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century (...) one will be able to speak of machines thinking without expecting to be contradicted.'
- 'We may hope that machines will eventually compete with men in all purely intellectual fields (...) and then teach it to understand and speak English'.

Source: Turing (1950)

## The Future of AI and its Impact to Humanity —————

Grace et al. (2024) presented the findings of a large-scale survey conducted among 2,778 AI researchers, focusing on their predictions regarding the future of AI. The survey revealed that there is at least a 50% chance of AI systems achieving significant milestones by 2028, such as autonomously creating a payment processing site, composing songs indistinguishable from those by popular musicians, and fine-tuning large language models. Researchers estimate a 10% chance of AI outperforming humans in all tasks by 2027 and a 50% chance by 2047. However, the full automation of all human occupations is expected to reach a 10% probability by 2037 and 50% by 2116.

The survey highlights substantial uncertainty about the long-term impacts of AI, with 68% of respondents believing that good outcomes are more likely than bad ones. Nonetheless, 48% of these optimists still consider a 5% chance of extremely bad outcomes, such as human extinction. Additionally, 51% of respondents give at least a 10% chance to AI leading to catastrophic outcomes.

The survey also identifies several AI-related scenarios that warrant substantial or extreme concern, including the spread of false information, authoritarian population control, and increased inequality. There is a lack of consensus on whether faster or slower AI progress would be better for humanity, but there is broad agreement on the need to prioritise research aimed at minimising potential risks from AI systems.