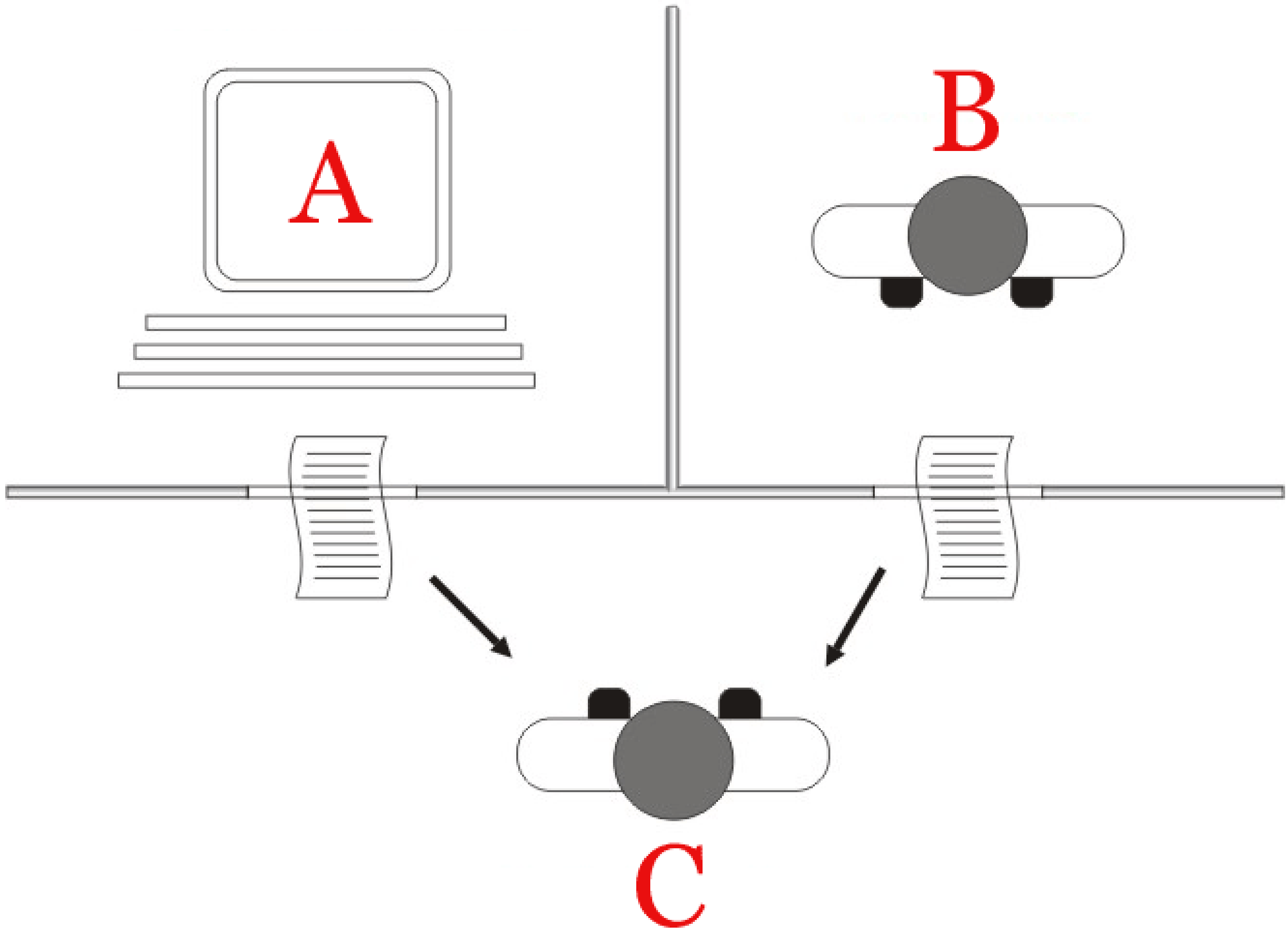
Artificial Intelligence

Artificial intelligence (AI) refers to computer systems capable of performing complex tasks that historically only a human could do, such as reasoning, making decisions, or solving problems.

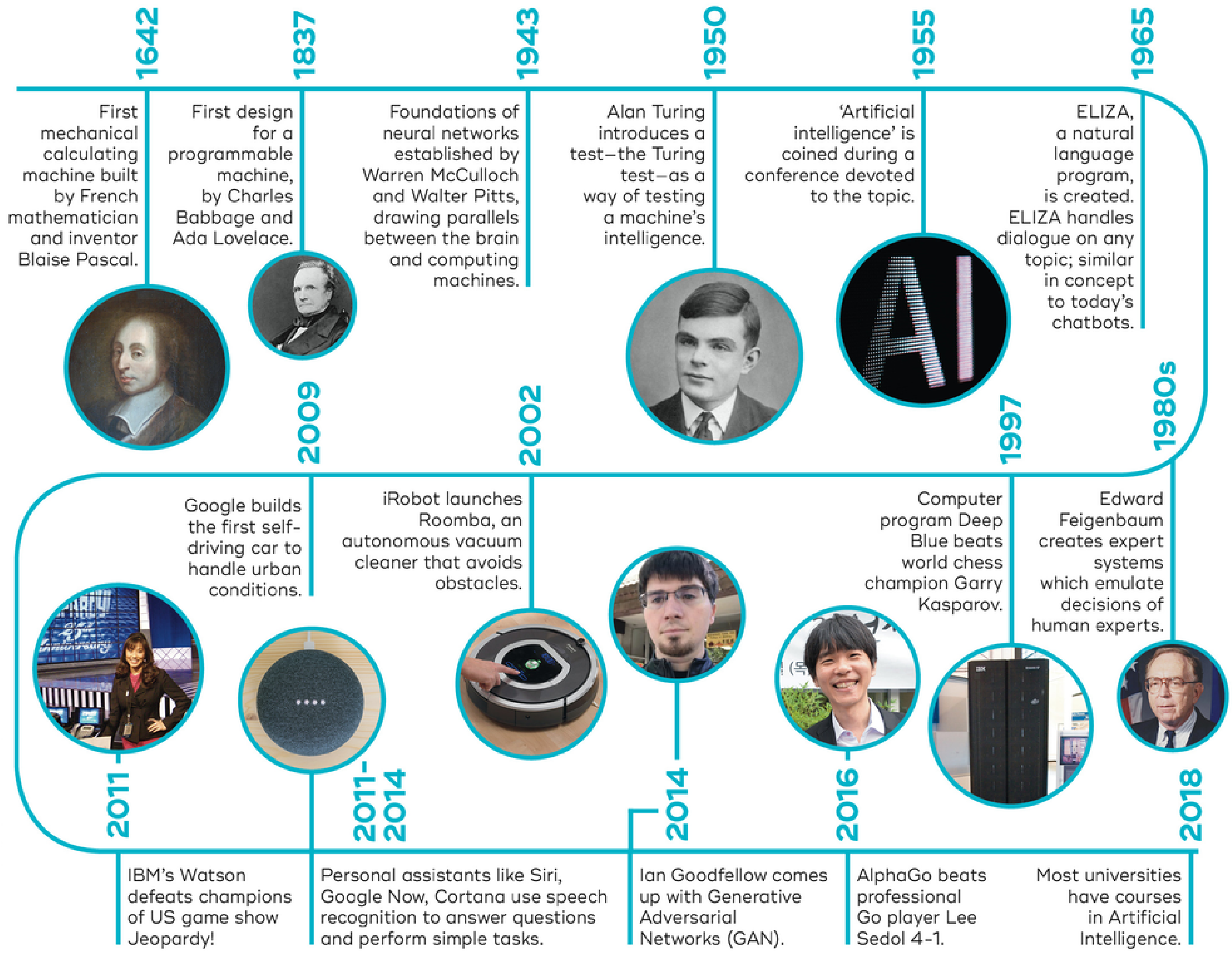
The concept of AI first emerged in the 1950s, but ideas about intelligent machines date back much further in human history, with mechanical automatons and imagined humanoid robots appearing in ancient myths and fiction.

The modern field of AI was largely sparked by Alan Turing's 1950 paper "Computing Machinery and Intelligence". In it, he poses the pivotal question: "Can machines think?" To evaluate whether machines can exhibit intelligent behavior, he devised the now famous Turing Test.



While no AI has fully passed the complex Turing Test yet, systems have come close recently. Thanks to rapid advances in the field over the past 20 years, AI and machine learning now power many everyday technologies and tools that would have seemed like science fiction just a decade ago.

A Brief history of AI



Pre-computer AI concepts

Ideas about developing artificial beings with intelligence date back many centuries. Notable examples include:

- **1200s**: Accounts of mechanical automatons able to sing, play instruments, or perform other lifelike actions.
- **1818**: Mary Shelley's Frankenstein explores the idea of artificially created life.
- **1950s**: As computing advanced, scientists start formally studying how to achieve artificial intelligence.

Founding of AI as a discipline

Key events in establishing AI as a scientific discipline include:

- **1950s**: Term "artificial intelligence" coined; Neural network concepts developed.
- **1956**: First academic AI conference held at Dartmouth College.
- **1997**: IBM's Deep Blue defeats world chess champion, demonstrating AI progress.

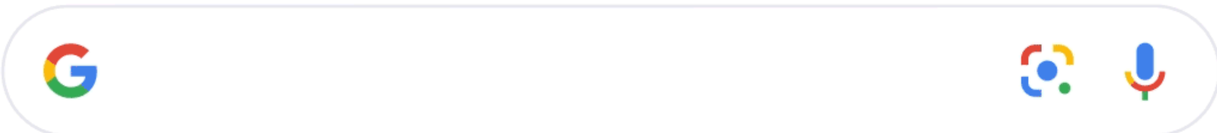
AI winter and resurgence

- **1974-1980s**: Disappointing early progress led to reduced funding and interest, known as the "AI winter".
- **1980s**: Expert systems and knowledge bases created for more specialised AI applications.
- **1990s**: Machine learning techniques helped AI progress beyond rule-based systems.
- **1997**: Deep Blue's chess victory regains interest in general AI goals.

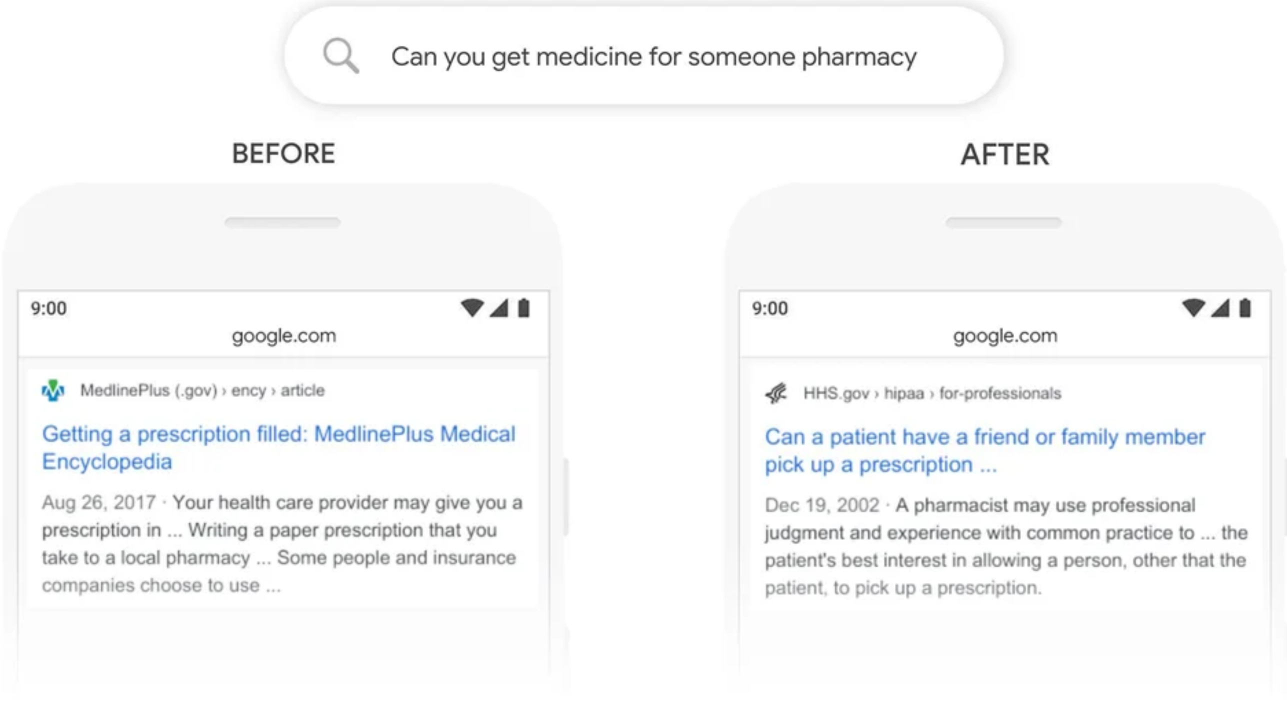
AI revolution powered by machine learning

Rapid advancement of AI since 2000 is largely driven by machine learning and deep neural networks, enabled by growth in data and computing power.

- **2001**: Machine learning helps Google Search users correct their spelling.



- **2011**: IBM Watson's Jeopardy! win highlights improvements in natural language processing.
- **2016**: AlphaGo program becomes first to defeat top human professional player in complex game of Go.
- **2019**: Natural language understanding helps Google to search better.



- **2020**: Protein-folding algorithms and large language models take major leaps forward.
- **2020**: OpenAI introduces GPT-3 model.
- **2022**: OpenAI launches ChatGPT.
- **2023**: Google launches Google BARD.

The pace of AI capabilities continues to accelerate.

Pioneers of modern AI

A diverse community of researchers and innovators has driven the rapid progress of artificial intelligence over the past few decades. Key breakthroughs trace back to the groundbreaking work of several brilliant minds who dedicated their careers to unlocking the possibilities of AI.

This section highlights five individuals who have been integral in propelling multiple waves of AI advancement from developing foundational machine learning concepts to realising transformational applications:

Andrew Ng

Andrew Ng popularised and democratised machine learning through education initiatives. Prolific founder and leader spreading AI globally.

Yoshua Bengio

Yoshua Bengio is a pioneer of deep learning techniques that sparked the AI revolution. Continues pushing boundaries of AI safety and reasoning.

Yann LeCun

Yann LeCun is the architect of pioneering CNNs powering modern computer vision. Drives Facebook (now Meta) AI research and REALM language models.

Geoffrey Hinton

Geoffrey Hinton is the neural network innovator who helped ignite machine learning mania through backpropagation and capsules.

Fei-Fei Li

Fei-Fei Li is the instrumental figure in large-scale image datasets, AI ethics and education. Leads Stanford HAI bridging AI and humanities.

Current and future applications

AI and machine learning are now embedded in many common consumer products and services:

- Smartphone assistants like Siri, Alexa and Google Assistant
- Image recognition and organisation in apps like Google Photos
- Predictive text typing suggestions
- Recommender systems providing personalised suggestions
- Self-driving and driver assistance car technologies
- Disease risk screening and medical diagnosis support

In the coming decades, AI could replicate more complex creative and contextual human skills previously thought impossible:

- Natural conversational ability
- Producing original art, music, literature
- Advanced strategic planning and decision making

- High-level physical dexterity and movement

Ethical considerations around data, bias, transparency, and impact on jobs and society will be crucial in guiding the continued advancement of AI technologies.

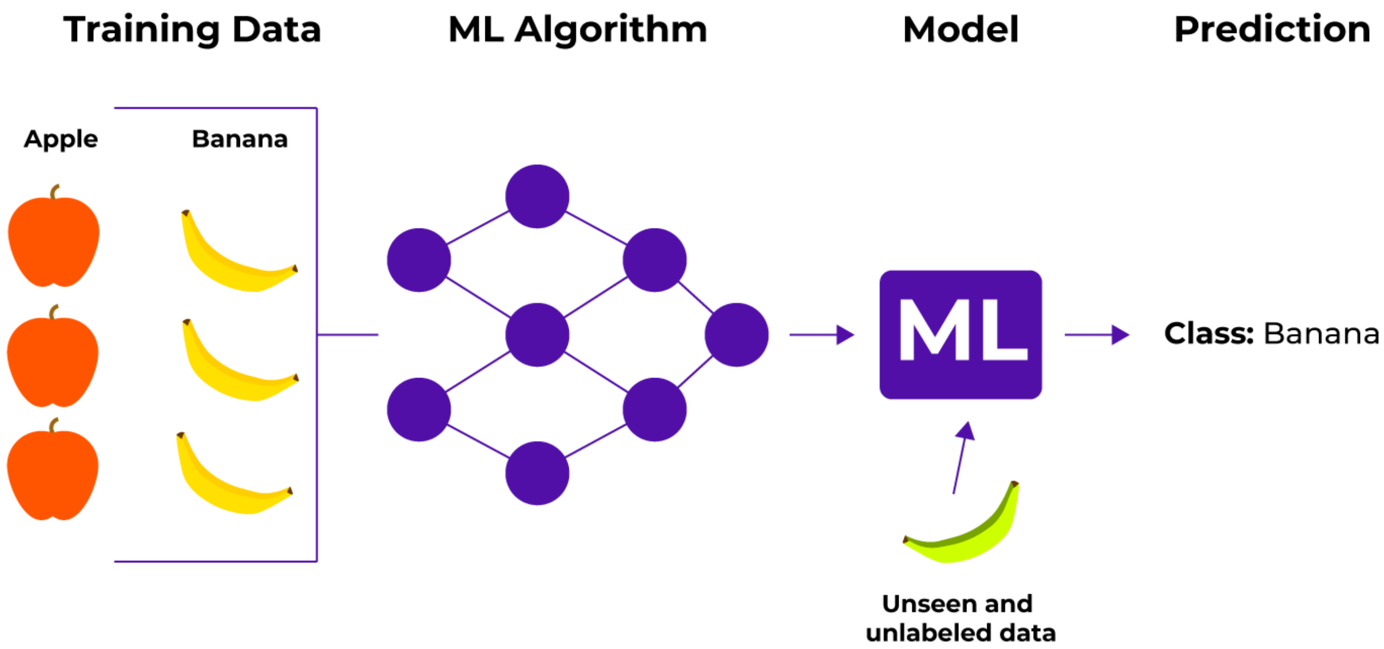
Machine Learning and AI Models

There are three main types of machine learning models:

Supervised learning

In supervised learning, models are trained on labelled datasets that contain inputs and desired outputs. For example:

- Image classification models are trained on images labelled with the object classes they depict (e.g. apples and bananas). Given new images, the models can predict these labels.

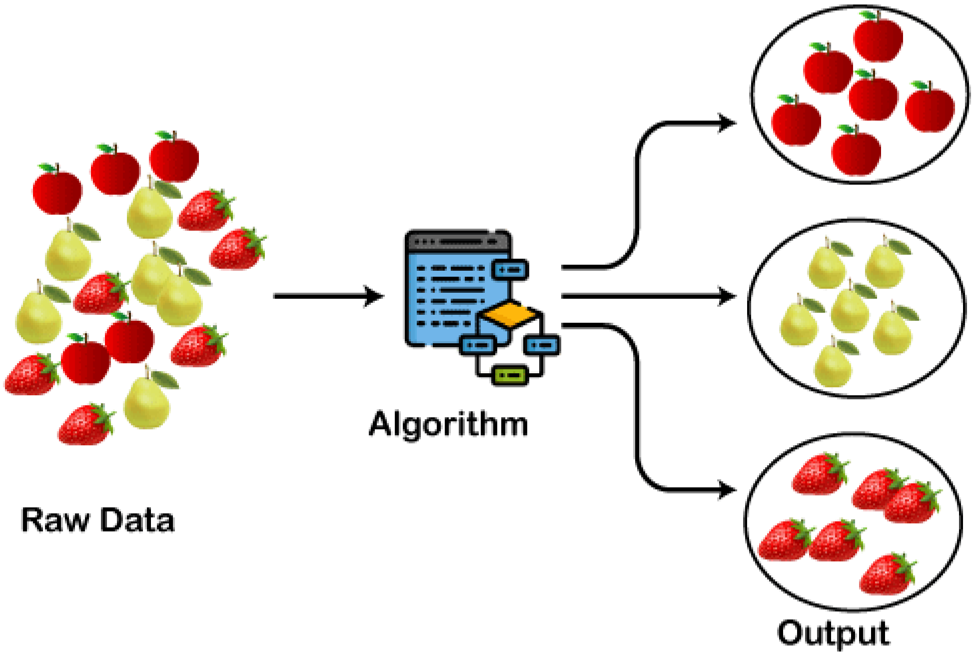


Supervised learning requires collecting and labelling large training datasets, which can be expensive and time-consuming. However, performance is often strong after sufficient training data is provided.

Unsupervised learning

Unsupervised learning models analyse unlabelled input data to find patterns and structure. Some examples include:

- Clustering algorithms group data points (e.g., apples, pears and strawberries) based on detected similarities. This can reveal categories within datasets.



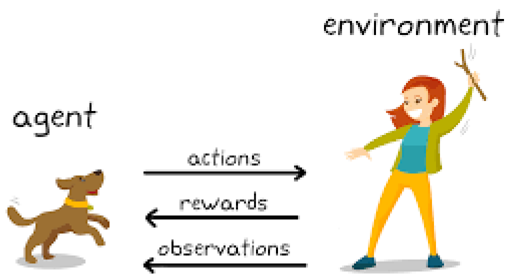
- Identifies outliers that are significantly different from the norm.
- Association rule learning finds interesting relationships between variables.

Unsupervised learning derives insights directly from data distributions without needing labelling, reducing data demands. But performance metrics can be harder to define.

Reinforcement learning

In reinforcement learning, agents interact dynamically with environments, receiving rewards or penalties for actions to learn behaviours that maximise cumulative reward. For instance:

- Robot dogs taught to walk properly through trial-and-error reinforcement of stable gaits.



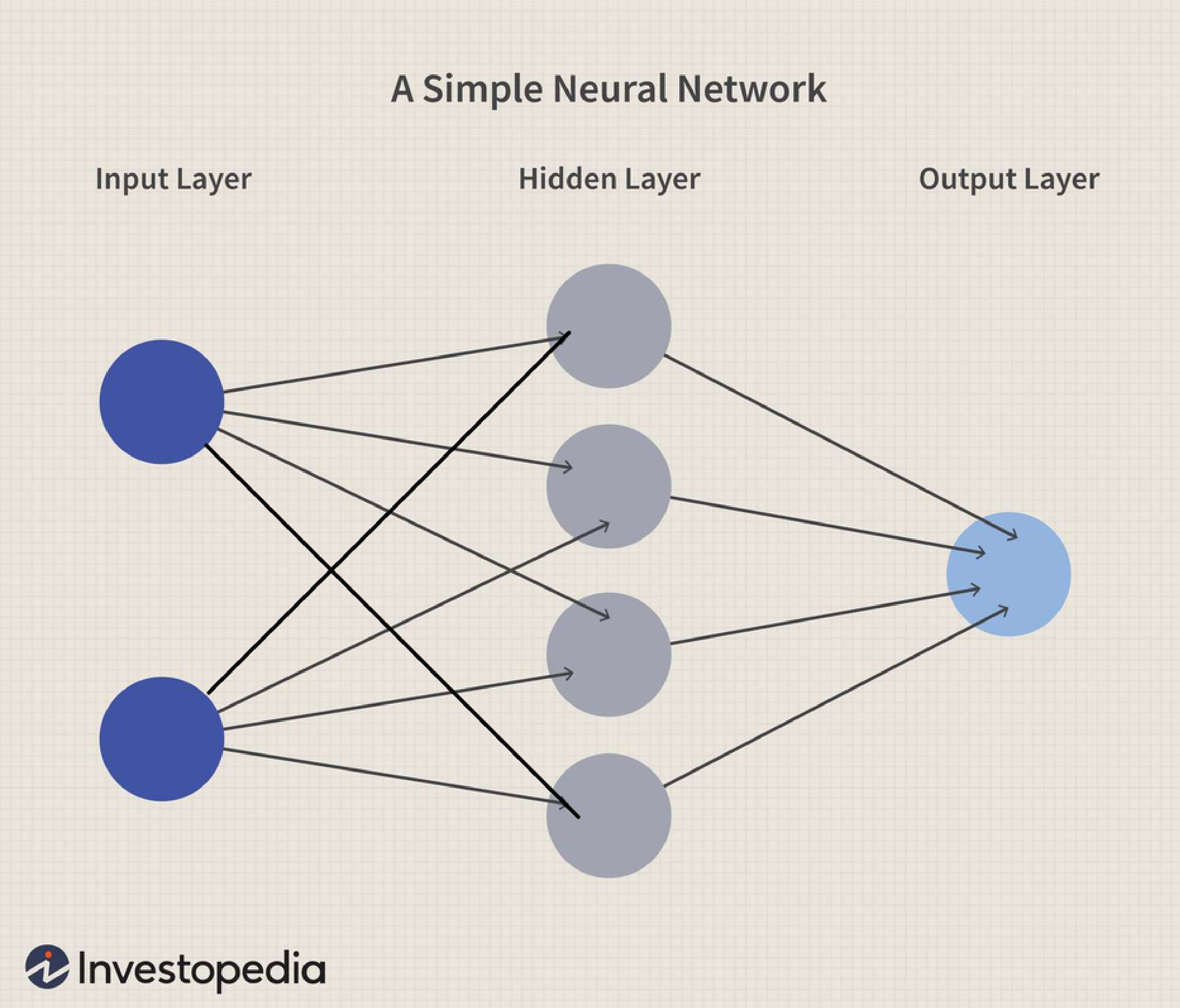
- Game-playing algorithms that learn winning strategies based on scores.

Reinforcement learning enables very responsive adaptation but requires carefully engineering reward functions.

These three model types are integral to machine learning, with a vast array of models designed for diverse tasks. While exploring various models is beyond this text's scope, the focus here is on a unique machine learning architecture: neural networks.

Neural networks

A neural network is a type of machine learning model loosely inspired by biological neural networks in the human brain. They are designed to recognise patterns and features in data.



At a very basic level, they consist of the following components:

Inputs

- Data such as images, text or sound samples get fed into the model
- Each input gets assigned to one of the initial **nodes**

Nodes

- Also called neurons or units
- Organised in layers - usually an input layer, one or more hidden layers, and an output layer
- Each node assigns a weight to its input, performs a calculation on it and passes the value to connected nodes

- Think of a node receiving multiple inputs, assessing their importance, and communicating its assessment to other neurons deeper in the network

Connections

- Direct connections between the output of nodes in one layer to the input of nodes in the next layer
- Enable information and weighted values to be passed along for further processing steps

Outputs

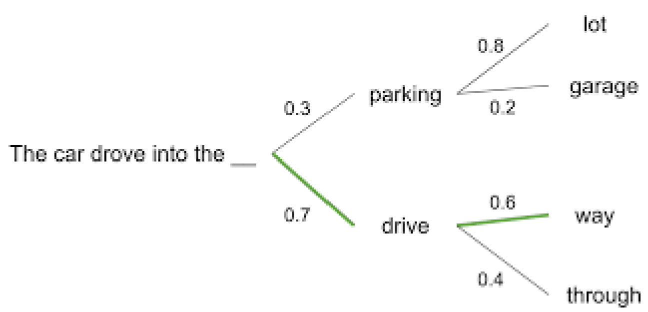
- The end result showing what the neural network has predicted based on all the calculations stepped through each layer of connections
- Output could be a classification, a prediction, recommendation or other inferred pattern in the data

The strengths of connections between nodes and the node weight values are optimised during model training to improve accuracy on sample data. The trained model can then be applied to new unseen data.

Various neural network types cater to specific tasks, such as vision, numerical processing, and language understanding. In this context, we will provide a brief overview of a large language model (LLM).

Large language model (LLM)

Large language models are a type of AI system that are trained on massive amounts of text data to predict probable next words when given a few words to start with. For example, if you type "The car drove into the..." the model will generate plausible continuations like "drive way" or "parking lot".



These models consist of billions of interlinked neural network parameters optimised through exposure to diverse examples of text structures. Thanks to their huge scale and smart architectures capturing complex language patterns, they can now write surprisingly human-like text given a prompt, power new applications like chatbots, and even reason about concepts they were never explicitly trained on. While impressive, they do have limitations in fully understanding context or factual knowledge without further progress. But large language models represent a major recent advance in AI's ability to not just process but dynamically generate natural language.

Popular llms

There are proprietary and open-sourced LLMs. Here is a list of popular one:

1. ChatGPT (https://chat.openai.com/)
2. Google Bard (https://bard.google.com/chat)
3. Claude AI (https://claude.ai/chat/)
4. LLAMA2 (https://ai.meta.com/llama/)

AI Classroom Activities

There are a number of activities you can implement in your classroom using both traditional AI as well as generative AI. This section will walk through two traditional AI activities and one generative AI activity. Form small groups of students to run these activities.

Image classification using Teachable Machine

Teachable Machine is a free, easy-to-use web-based tool from Google for building basic image classification models without needing to code.

- Website link: Teachable Machine
- Account needed: No
- Platform: Fully broswer-based

Activity: Classifying happy and sad faces

This activity guides students through using Teachable Machine to create a two-class image classification model for detecting happy and sad facial expressions.

Steps:

- Collect a dataset of happy face photos and sad face photos using a webcam.
- Label the photos in Teachable Machine.
- Train a machine learning model on this dataset.
- Evaluate model accuracy by testing the output of the model.
- Use the model to predict if new facial images are happy or sad.

Key learning outcomes:

- Data collection and labelling
- Training machine learning models
- Model evaluation

For full activity instructions, see: Teachable Machine Activity Guide

Students can also integrate their Teachable Machine model into Scratch coding projects for interactive apps. However, this currently has limited support on the main Scratch website. See Teachable Machine with Scratch activity guide for more details on how to use a different Scratch version to achieve it.

Image classification with Landing AI

Landing AI is a more advanced platform for training custom visual AI models.

- Website link: Landing AI
- Account needed: Yes, free trial available
- Platform: Browser-based

Activity: Building an image classifier

Follow the guide to take students through gathering images, labelling datasets, training models, and deploying models with Landing AI:

Landing AI Activity Guide

Key learning outcomes:

1. Data collection and labelling
2. Training machine learning models
3. Model performance evaluation
4. Model deployment

Generative AI children's book creation

This activity introduces students to generative AI by having them use text and image generation models to create personalised children's books.

Groups will be provided example user profiles to inspire ideas. Students will use AI tools to source content while applying their own creativity to construct an original story.

Instructions

1. Choose a user profile and brainstorm initial book ideas
2. Map out core storyline, characters, and assets list
3. Use AI tools (Google BARD and Playground AI) to generate text and images
4. Iteratively improve and structure content in PowerPoint or Google Slides
5. Construct at least one full chapter

Sample user prompt

> Within your team, use Google BARD and PlaygroundAI to create a personalised children's book for your assigned user profile. Apply creativity in directing the AI and designing an engaging storyline. Construct at least one full chapter structured in PowerPoint.

For Playground AI instructions, see: Playground AI Guide

Key learning outcomes:

- Generative AI capabilities
- AI content filtering and improvement
- Creative direction of AI output

AI impact and critical thinking activity

If you want to bring ethical side and the social impacts of AI in your classroom, here are some sample questions to brainstrom.

Activity: AI impact brainstrom

1. Have students brainstorm current or future AI applications in 3 categories: transportation, medicine, entertainment. Examples: self-driving cars, robotic surgery tools, movie recommendation algorithms.
2. For each category, split into small groups to analyse positive and negative potential societal impacts. Consider things like jobs, access, privacy.
3. Groups present out analyses. Discuss tradeoffs identified and possible solutions.

Activity: critical thinking questions

Transportation:

- How might self-driving vehicle AI impact jobs like truck driving, taxis, public transit? Are job losses inevitable or preventable?
- Could AI transportation developments restrict mobility for groups like the elderly or disabled? How could they be designed inclusively?

Medicine:

- AI systems ingest huge amounts of patient health data. How does this impact privacy? Should anonymisation happen?
- If an AI diagnosis system makes an error, who should be liable? The developer, hospital, or someone else?

Entertainment:

- If AI generated songs or films become indistinguishable from human-made, will it diminish human creativity? Or enhance it?
- Should algorithms recommending news, movies, or products be transparent how they profile and target users? Why or why not?

These sample questions aim to get students investigating nuanced AI impact dilemmas through different lenses while thinking critically.

References

- What is Artificial Intellience? Click here to access. (Accessed: 23rd Nov 2023)
- What Is Artificial Intelligence? Definition, Uses, and Types. Click here to access it. (Accessed: 23rd Nov 2023)
- AI is closer than ever to passing the Turing test for 'intelligence'. What happens when it does? Click here to access it. (Accessed: 23rd Nov 2023)
- History of Artificial Intelligence. Click here to access. (Accessed: 23rd Nov 2023)
- Turing, A. M. (1950). Computing Machinery and Intelligence. Mind, 49, 433-460. https://doi.org/10.1093/mind/LIX.236.433
- McCorduck, P. (2004). Machines Who Think (2nd ed.). Natick, MA: A. K. Peters, Ltd.