AI and Human Labour

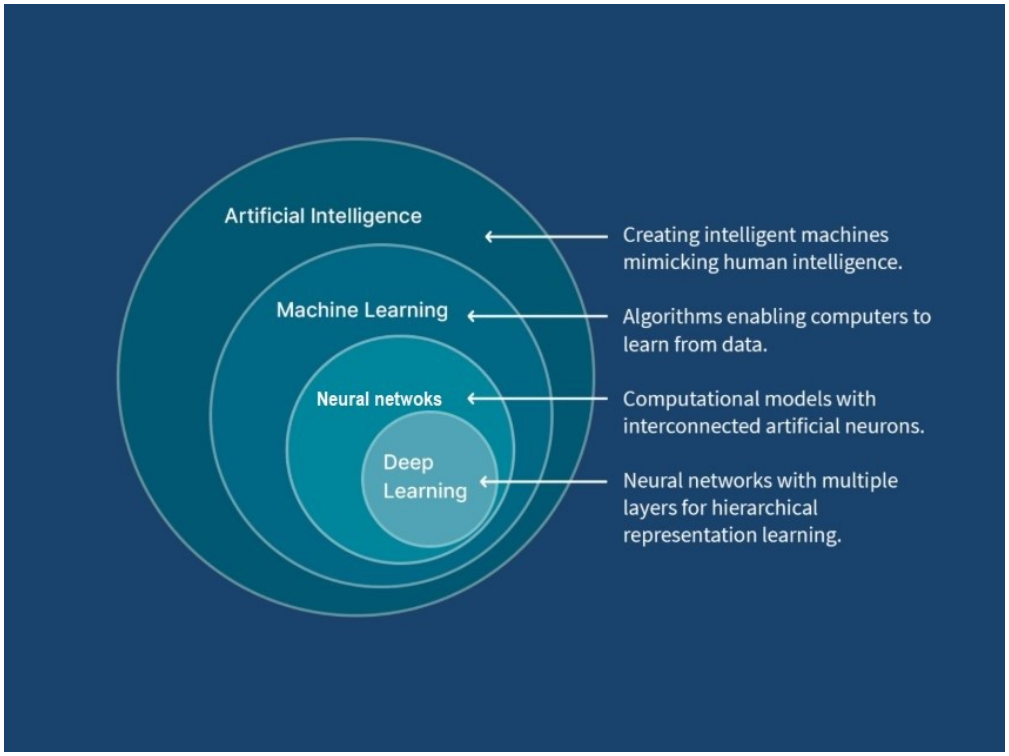
Mains: *GS III – Science and Technology- Developments and their Applications and Effects in Everyday Life*

Why in News?

Recently, there has been a lot of challenges faced by the AI data annotators and workers in various countries.

What are the important components of AI?

- **Artificial Intelligence (AI) –** It is a technology that empowers machines to perform tasks traditionally requiring human intelligence, such as learning, problem-solving, decision-making, and understanding language.
- *It works by using algorithms* to analyze vast amounts of data, identify patterns, adapt to new information, and ultimately act and make decisions with minimal human input
- **Algorithms** – In AI, an algorithm is a step-by-step set of instructions that enables a machine to process data, learn from it, identify patterns, and make decisions or predictions autonomously.
- **Machine learning (ML)** – It is a type of artificial intelligence that enables computers to learn from data and improve their performance over time without being explicitly programmed.
- **Neural network** – It is a machine learning system, inspired by the human brain that uses layers of interconnected nodes (neurons) to process data, recognize patterns, and make predictions.

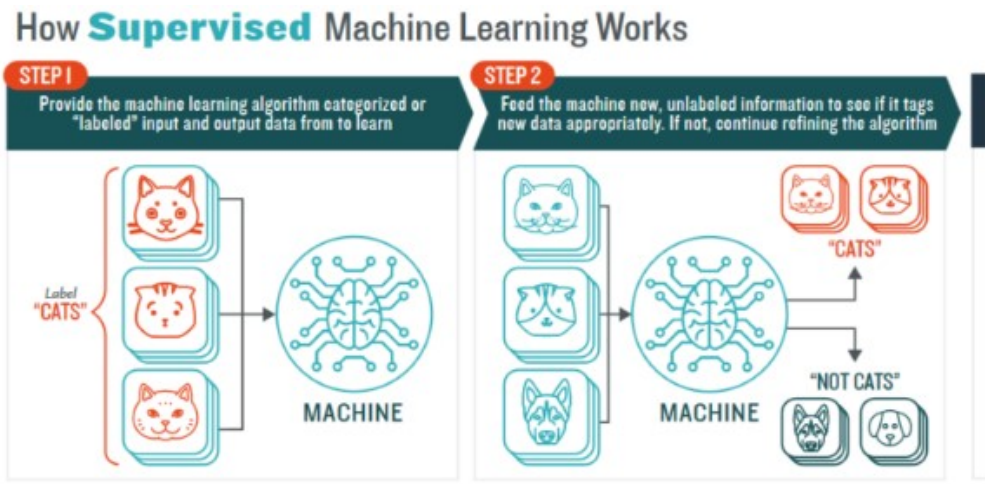


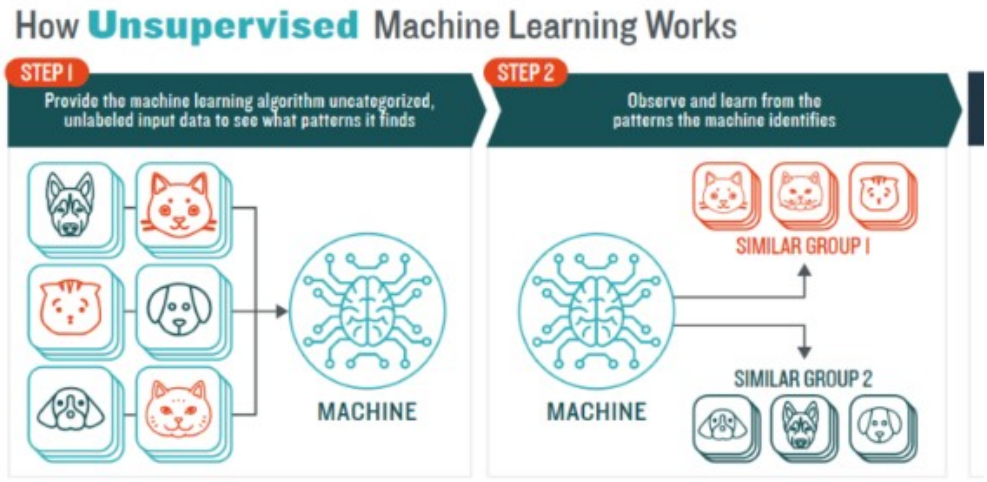
- **Deep learning** – It is a powerful subfield of artificial intelligence (AI) that uses "deep" artificial neural networks with multiple layers to learn complex patterns from vast amounts of data.
- It enables machines to perform tasks like image recognition, language translation, and autonomous driving without explicit human instruction.
- **Data annotation** – It is the process of labeling or tagging data to make it understandable for machine learning (ML) models.
- This involves adding descriptive information, such as labels, tags, or bounding boxes, to data like images, text, audio, and video.
- Annotated data helps ML models learn to recognize patterns, make predictions, and perform complex tasks.

What are the roles of human in AI and machine learning?

- **Role of human** – AI is not getting smarter on its own, it has been built on and continues to rely on human labour and energy resources.
- These systems are fed information and trained by workers who are invisibilised by large tech companies, and mainly located in developing countries.
- A machine cannot process the meaning behind raw data.
- **Data annotation** – Data annotators label raw images, audio, video, and text with information that trains AI and Machine Learning (ML) models.
- This, then, becomes the training set for AI and Machine Learning (ML) models.
    - **For example**, a large-language models (LLM) cannot recognise the colour ‘yellow’ unless the data has been labelled as such.
    - Similarly, self-driving cars rely on information from video footage that has been labelled to distinguish between a traffic sign and humans on the road.
- The higher the quality of the dataset, the better the output and the more human labour is involved in creating it.
- Data annotators play a major role in training LLMs like ChatGPT, Gemini, etc.

- An LLM is trained in three steps:
    - Self-supervised learning
    - Supervised learning
    - Reinforcement learning.
- **Self-supervised learning** – In this step the machine picks up information from large datasets on the Internet.
- **Supervised learning and Reinforcement learning** – The data labellers or annotators enter in this stage.





- Here the information is fine-tuned for the LLM to give the most accurate response.
- **Automated features requiring humans** – Even features marketed as 'fully automated' are often underpinned by invisible human work.
    - **For example**, our social media feeds are 'automatically' filtered to censor sensitive and graphic content.
- This is only possible because human moderators labelled such content as harmful by going through thousands of uncensored images, texts and audio.
- The exposure to such content daily has also been reported to cause severe mental health issues like post-traumatic stress disorder, anxiety and depression in the workers.
- **Voice actors** – There are voice actors and actors behind AI-generated audios and videos.
- Actors may be required to film themselves dancing or singing for these machines to

recognize human movements and sounds.
- Children have also been reportedly engaged to perform such tasks.

- **Feedback and response** – Humans give feedback on the output the AI produces for better responses to be generated over time, as well as remove errors and jailbreaks.

What are the issues involved?

- **Labour outsourcing** – This meticulous annotating work is outsourced by tech companies in Silicon Valley to mainly workers in countries like Kenya, India, Pakistan, China and the Philippines for low wages and long working hours.
- **Errors in data** – Several tech companies have been accused of employing non-experts for technical subjects that require prior knowledge.
- This is a contributing factor in the errors found in the output produced by AI.
    - **For instance**, A data labeller from Kenya revealed that they were tasked with labelling medical scans for an AI system intended for use in healthcare services elsewhere, despite lacking relevant expertise
- Data labelling can be of two types
    - Those which do not require subject expertise
    - Those which are more niche and require subject expertise.

- However, due to errors resulting from this, companies are starting to ensure experts for such information being fed into the system.
- **Poor working conditions** – In 2024, AI tech workers from Kenya sent a letter to former U.S. President Joe Biden talking about the poor working conditions they are subjected to.
- Our working conditions amount to modern-day slavery," the letter read.
- **Long working hours** – They said the content they have to annotate can range from pornography and beheadings to bestiality for more than eight hours a day, and for less than $2 an hour, which is very low in comparison to industry standards.
- **Strict deadlines** – There are also strict deadlines to complete a task within a few seconds or minutes.
- They are constantly surveilled, and if they fall short of the targeted output, they are fired.
- **Undermining labour laws** – In Kenya, these US companies are undermining the local labor laws, the country's justice system and violating international labor standards.
- When workers raised their concerns to the companies, they were sacked and their unions dismantled.
- **Engaged as gig workers** – Most AI tech workers are unaware of the large tech company they are working for and are engaged in online gig work.
- This is because, to minimise costs, AI companies outsource the work through intermediary digital platforms.
- The advancement of AI is powered by such "ghost workers."
- **Fragmented labour network** – There are subcontract workers in these digital platforms who are paid per "microtask" they perform.
- Hence, the labour network becomes fragmented and lacking transparency.
- **Lack of recognition** – The lack of recognition and informalisation of their work helps tech companies to perpetuate this system of labour exploitation.

What lies ahead?

- There is a need to bring in stricter laws and regulations on AI companies and digital platforms.
- There should be regulations for the labour supply chains powering AI, ensuring transparency, fair pay, and dignity at work.

Reference

The Hindu| The Hidden Human Cost of AI