



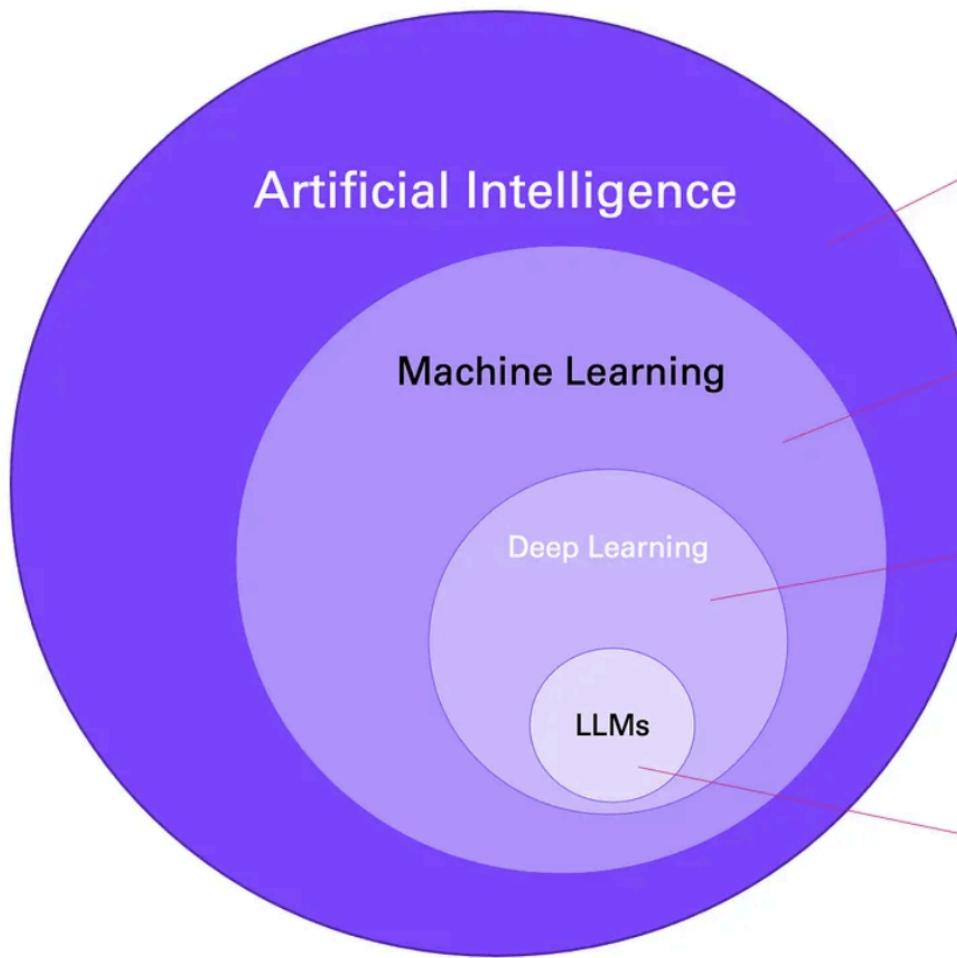
AUFITAL INTELLIGENCE

INTELIGATION

RESION

CHIUTTUNP FSRLIN





**Intelligent Machines**  
Broadly defined

**Pattern Recognition**  
Learning general patterns from data

**Neural Networks**  
Learning general patterns in unstructured data (i.e. images, text, audio, etc.)

**Large Language Models**  
Learning to understand natural language (i.e. text)

LLM are machine learning models that are really good  
at understanding and generating human language.  
Based on Transformers a type of neural network

Large language models have proven instrumental in addressing a wide array of challenges, from automating complex tasks to enhancing decision-making processes. However, scaling these models has also introduced considerable complexities, such as high computational costs, reduced accessibility, and the environmental impact of extensive resource requirements.

One Model can be used for a whole variety of tasks,

One Model can be used for a whole variety of tasks,

Chat

One Model can be used for a whole variety of tasks,

Chat

Copy-writing

One Model can be used for a whole variety of tasks,

Chat

Copy-writing

Translation

One Model can be used for a whole variety of tasks,

Chat

Copy-writing

Translation

Summarization

One Model can be used for a whole variety of tasks,

Chat

Copy-writing

Translation

Summarization

Brainstorming

One Model can be used for a whole variety of tasks,

Chat

Copy-writing

Translation

Summarization

Brainstorming

Code Generation

LLM are instance of Foundation Models.

Foundation models are pre-trained on large amounts of unlabelled and self-supervised data. The Model learns from patterns in the data in a way that produces generalizable and adaptable output. Applied specifically to text and text-like things. Trained on petabytes of data.

1 Gigabyte in size that can store about 178 million words. In a petabyte there are about 1 million gigabyte.

## Data

Text

Images

Speech

Structured Data

3D Signals

Training

Foundation Model

Adaptation

Question Answering

Sentiment Analysis

Information Extraction

Image Captioning

Object Recognition

Instruction Following



Few examples of LLM models that can be downloaded  
from Huggingface

Model	Parameters	Size	Download
Llama 3.2	3B	2.0GB	<code>ollama run llama3.2</code>
Llama 3.2	1B	1.3GB	<code>ollama run llama3.2:1b</code>
Llama 3.1	8B	4.7GB	<code>ollama run llama3.1</code>
Llama 3.1	70B	40GB	<code>ollama run llama3.1:70b</code>
Llama 3.1	405B	231GB	<code>ollama run llama3.1:405b</code>
Phi 3 Mini	3.8B	2.3GB	<code>ollama run phi3</code>
Phi 3 Medium	14B	7.9GB	<code>ollama run phi3:medium</code>
Gemma 2	2B	1.6GB	<code>ollama run gemma2:2b</code>
Gemma 2	9B	5.5GB	<code>ollama run gemma2</code>
Gemma 2	27B	16GB	<code>ollama run gemma2:27b</code>
Mistral	7B	4.1GB	<code>ollama run mistral</code>
Moondream 2	1.4B	829MB	<code>ollama run moondream</code>
Neural Chat	7B	4.1GB	<code>ollama run neural-chat</code>
Starling	7B	4.1GB	<code>ollama run starling-lm</code>
Code Llama	7B	3.8GB	<code>ollama run codellama</code>
Llama 2 Uncensored	7B	3.8GB	<code>ollama run llama2-uncensored</code>
LLaVA	7B	4.5GB	<code>ollama run llava</code>
Solar	10.7B	6.1GB	<code>ollama run solar</code>