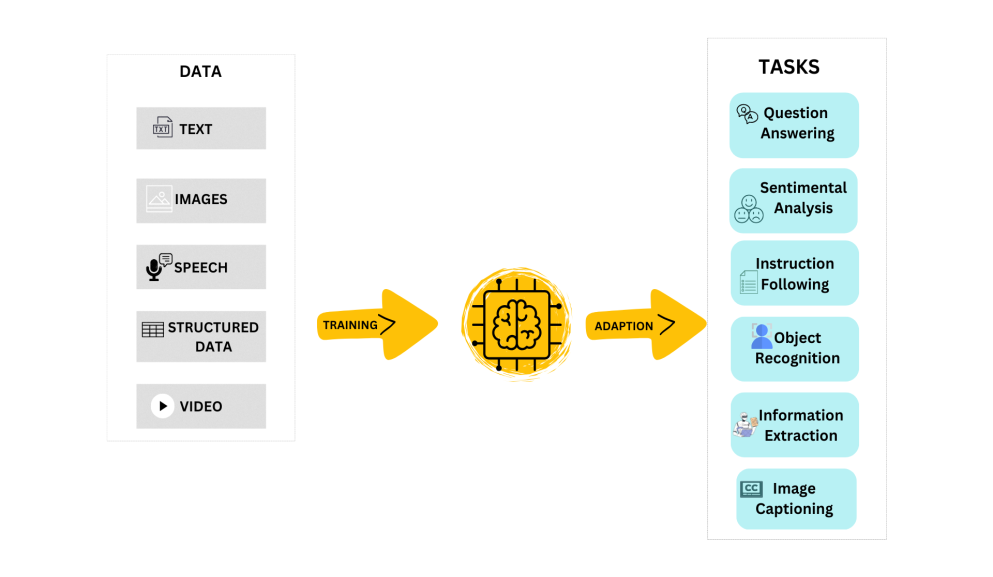
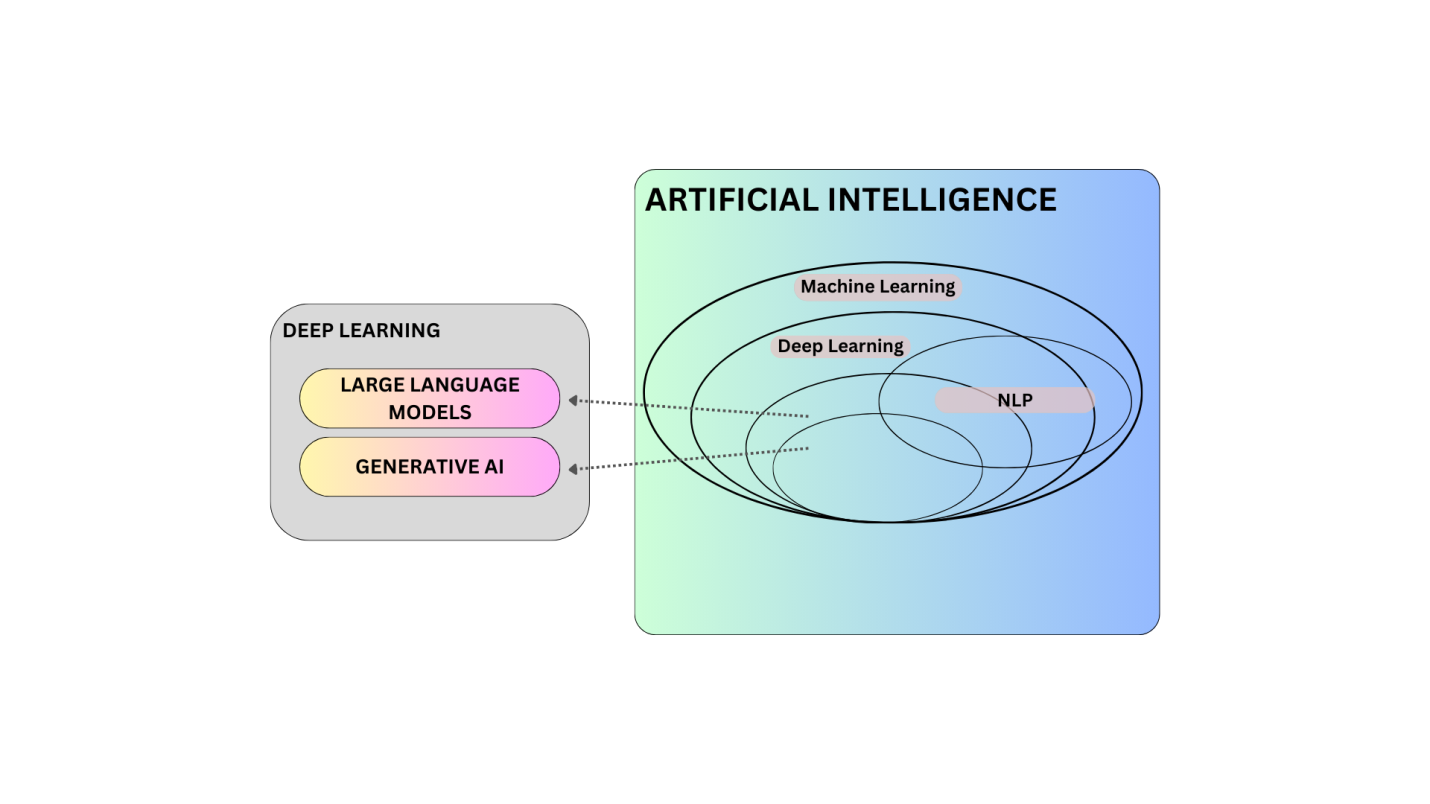
Introduction to Generative AI

Generative AI (Gen AI)

Generative AI is a type of artificial intelligence that can create new content, like writing, images, music, or even code, instead of just recognizing patterns or making predictions. It learns from a lot of data and uses that knowledge to produce something new and original, which can look or sound like the data it was trained on.



How Gen AI is differing from ML, DL, NLP and LLM?



LLM (Large Language Model)

A **Large Language Model (LLM)** is a type of artificial intelligence model designed to understand and generate human language. It is built using a neural network that is trained on massive amounts of text data, allowing it to perform a wide variety of language-related tasks, such as answering questions, writing essays, translating languages, summarizing information, and even generating creative content like stories or poems.

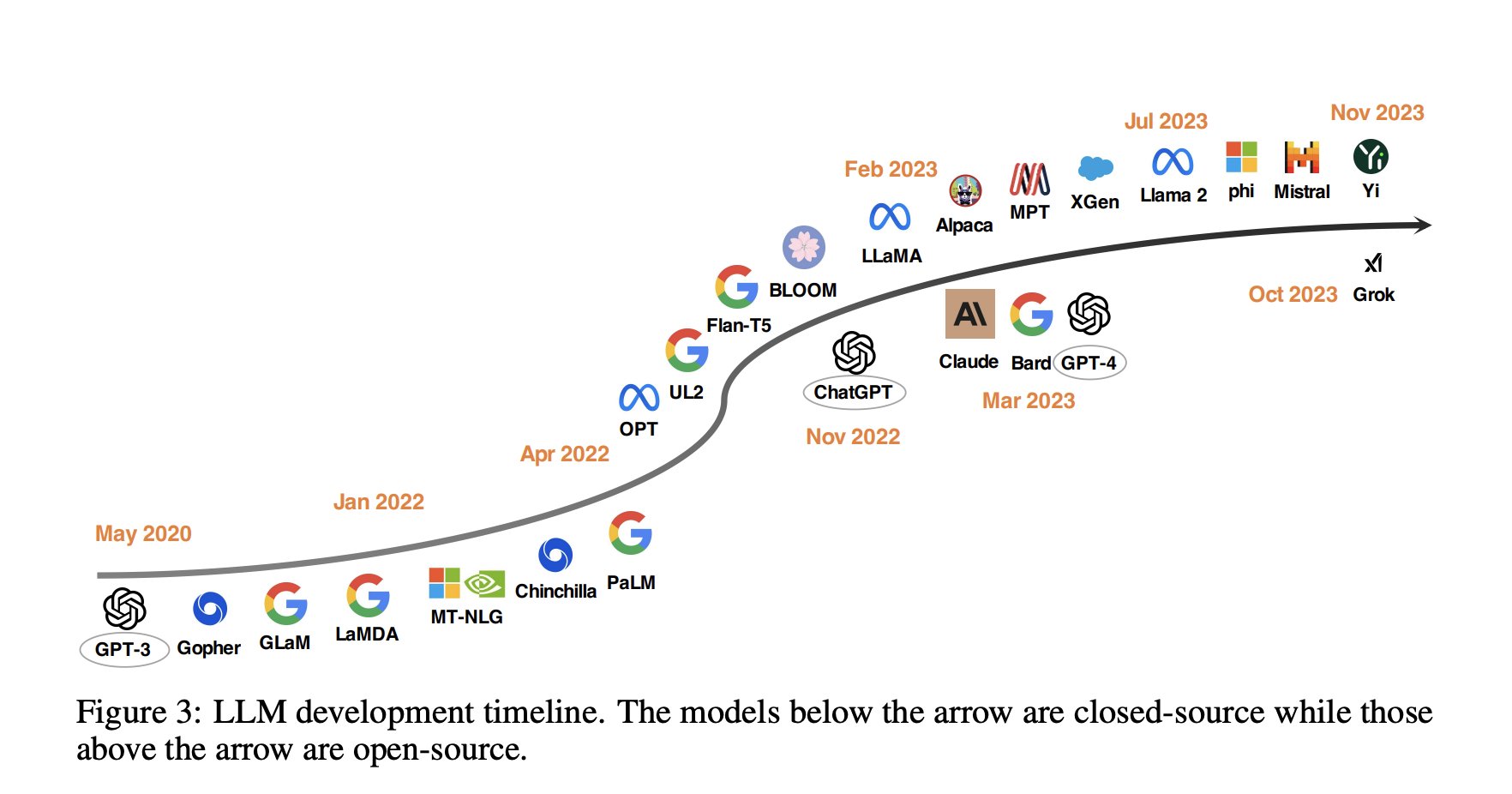
Comparison of LLM and Gen AI:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Aspect** | | |  | | --- | | **Generative AI (Gen AI)** | | |  | | --- | | **Large Language Models (LLMs)** | |
| |  | | --- | | Definition | | |  | | --- | | AI models that can create new content, such as text, images, audio, or video |  |  | | --- | |  | | A specific type of Gen AI focused on understanding and generating text |
| |  | | --- | | Focus | | |  | | --- | | Can generate various forms of content (text, images, audio, video, etc.) | | Primarily focuses on text-based tasks like writing, answering questions, or summarizing |
| |  | | --- | | Examples | | |  | | --- | | - DALL·E (image generation)  - GANs (image/video generation)  - Jukebox (music generation) |  |  | | --- | |  | | - GPT-3, GPT-4 (text generation)  - BERT (text understanding) |
| |  | | --- | | Output Types | | |  | | --- | | - Text  - Images  - Audio  - Video | | - Text only |
| |  | | --- | | Purpose | | |  | | --- | | Broader purpose, used in creativity, design, entertainment, and more | | |  | | --- | | Specialized in language-related tasks like chatting, translation, summarization | |
| |  | | --- | | Key Techniques | | |  | | --- | | Uses various models like GANs, VAEs, Transformers |  |  | | --- | |  | | Primarily based on Transformer architecture |
| |  | | --- | | Domain | | |  | | --- | | Multi-domain (art, design, text, music, etc.) |  |  | | --- | |  | | Focused on natural language processing (NLP) |
| |  | | --- | | Example Use Cases | | |  | | --- | | - Image synthesis  - Music composition  - Video generation |  |  | | --- | |  | | - Chatbots  - Text generation  - Document summarization |
| |  | | --- | | Subset Relationship | | |  | | --- | | Broad category, includes LLMs as a subset | | |  | | --- | | Subset of Generative AI, specifically dealing with text | |

**Traditional Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP)** versus **Generative AI (Gen AI)**:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Aspect** | | |  | | --- | | **Traditional ML/DL/NLP** | | |  | | --- | | **Generative AI (Gen AI)** | |
| |  | | --- | | **Definition** | | |  | | --- | | Traditional approaches for training models to recognize patterns and make predictions based on existing data. | | |  | | --- | | AI models that generate new content (text, images, audio, video, etc.) from learned patterns. | |
| |  | | --- | | **Focus** | | |  | | --- | | Focuses on tasks like classification, prediction, and pattern recognition. | | |  | | --- | | Focuses on creating new data, such as generating realistic images, writing text, composing music, etc. | |
| |  | | --- | | **Examples** | | |  | | --- | | - ML: Linear Regression, Decision Trees  - DL: CNNs, RNNs  - NLP: Named Entity Recognition (NER), Machine Translation | | |  | | --- | | - DALL·E (image generation)  - GPT-3/GPT-4 (text generation)  - GANs (image/video generation) | |
| |  | | --- | | **Output Types** | | |  | | --- | | Predictions or classifications (e.g., spam detection, image recognition, text translation). | | |  | | --- | | New, creative outputs like text, images, audio, or video. | |
| |  | | --- | | **Data Dependency** | | |  | | --- | | Requires structured, labeled data for training and typically focuses on pre-defined outputs. | | |  | | --- | | Can work with both structured and unstructured data to create novel outputs that resemble the training data. | |
| |  | | --- | | **Purpose** | | |  | | --- | | To make predictions, classifications, or decisions based on input data. | | |  | | --- | | To generate new content and simulate creative processes similar to human creativity. | |
| |  | | --- | | **Training Data** | | |  | | --- | | Needs specific input-output pairs for supervised learning tasks. | | |  | | --- | | Trained on large datasets (e.g., text, images) to learn the patterns and then generate new data. | |
| |  | | --- | | **Complexity** | | |  | | --- | | Traditional ML is simpler (requires less data and computation); DL models (like CNNs and RNNs) need more data and are more complex. | | |  | | --- | | Generally more complex due to larger models and more diverse outputs, often using advanced architectures like GANs or Transformers. | |
| |  | | --- | | **Application Domain** | | |  | | --- | | Used in tasks like fraud detection, stock market prediction, image classification, language translation. | | |  | | --- | | Used in creative tasks like text writing, image creation, music composition, and video generation. | |
| |  | | --- | | **Architecture** | | |  | | --- | | - ML: Decision Trees, SVMs, Logistic Regression  - DL: Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) | | |  | | --- | | Generative models like Transformer-based LLMs (GPT), GANs, VAEs. | |
| |  | | --- | | **Model Behavior** | | |  | | --- | | Recognizes patterns and makes decisions based on those patterns (e.g., "What is this image?"). | | |  | | --- | | Generates new data by learning from patterns (e.g., "Create an image based on this text prompt"). | |
| |  | | --- | | **Example Use Cases** | | |  | | --- | | - Fraud detection in banking  - Predicting house prices  - Recognizing objects in images  - Translating languages | | |  | | --- | | - Text generation for writing assistance  - Creating art or images  - Generating music or audio  - Video creation based on descriptions | |
| |  | | --- | | **User Interaction** | | |  | | --- | | Users typically input data and get predictions or classifications. | | |  | | --- | | Users provide prompts or partial input (e.g., text description), and the model generates new content. | |
| |  | | --- | | **Subset Relationship** | | |  | | --- | | Traditional ML/DL/NLP are broader in their focus on structured prediction tasks. | | |  | | --- | | Gen AI is a subset of AI that specializes in creating new, creative outputs. | |

Milestone in LLM



|  |  |  |
| --- | --- | --- |
| Aspect | Gen AI | LLM |
| Definition | AI Model that can create new content such as text, images, audio, or video. | A specific type of Gen AI focused on understanding and generating text. |
| Focus | Can generate various forms of context (text, images, audio, video, etc.) | Primarily focuses on text-based tasks like writing, answering questions, or summarizing. |
| Examples | -DALL·E (image generation)  -GANs (image/video generation)  -Jukebox (music generation) | -GPT-3, GPT-4 (text generation)  -BERT (text understanding) |
| Output Types | - Text  - Images  - Audio  - Video | Text only |
| Purpose | Broader purpose, used in creativity, design, entertainment, and more | Specialized in language-related tasks like chatting, translation, summarization |
| Key Techniques | Uses various models like GANs, VAEs, Transformers | Primarily based on Transformer architecture |
| Domain | Multi-domain (art, design, text, music, etc.) | Focused on natural language processing (NLP) |
| Example Use Cases | -Image synthesis  -Music composition  -Video generation | - Chatbots  -Text generation  -Document summarization |
| Subset Relationship | Broad category, includes LLMs as a subset | Subset of Generative AI, specifically dealing with text |