

Artificial Intelligence GenAI

1. Introduction

Generative AI (GenAI) is a branch of Artificial Intelligence that enables machines to **create new content** — such as text, images, code, music, audio, and videos — similar to what humans can create.

Before GenAI, AI systems could only classify or predict.

Now AI can **generate**.

This shift happened mainly due to **Transformers + Large Language Models (LLMs)**.

2. What is Generative AI?

Definition:

Generative AI refers to AI systems capable of producing **original, high-quality content** by learning patterns from large datasets.

Simple meaning:

GenAI = Machine learns → Machine creates

Examples of things GenAI can generate:

- Text (ChatGPT)
- Code (Copilot)
- Images (Midjourney, DALL·E)
- Video (Sora, Runway)
- Music (Suno, Udio)
- 3D models
- Voice & speech

3. Need of Generative AI

1 Automate Creative Work

- Write articles, emails, scripts
- Generate designs, logos, posters
- Create product descriptions automatically

2 Reduce Development Time

- AI writes boilerplate code
- AI creates test cases, documentation
- Fast prototyping for apps & websites

3 Personalization at Scale

- Personalized ads
- Personalized learning
- Tailored recommendations

4 Make complex tasks easier

- Summarizing long research papers
- Turning ideas → working code
- Converting text → videos

5 Enhance productivity & creativity

GenAI is like a **super assistant** for every domain:

- Developers
- Writers
- Designers
- Marketers

- Students
-

4. Real-World Applications of GenAI

A. Text Generation

- ChatGPT, Claude, Gemini
- Customer support chatbots
- Email drafting
- Legal document generation
- Content writing

B. Coding & Development

- GitHub Copilot
- Cursor IDE
- Auto-code generation & debugging
- Test case creation
- Code review automation

C. Image Generation

- Midjourney
- DALL·E
- Stable Diffusion
Uses:
 - Ads
 - Thumbnails
 - E-commerce photo creation

- Artistic visuals

D. Video Generation

- Sora AI
- Runway ML
- Pika Labs
Useful for:
 - Marketing videos
 - Movie scenes
 - Animation
 - Educational content

E. Music & Audio Generation

- Suno
- Udio
- AI voice cloning
Examples:
 - Background music
 - Podcast voices
 - Audio books

F. Business Use Cases

- Product recommendation chatbots
- Automatic slide deck creation
- Financial report summarization
- HR resume screening
- Sales pitch generation

G. Medical & Healthcare

- AI doctor notes
- Diagnosing assistance
- Medical research summarization

H. Gaming

- AI NPC behavior
 - AI-generated levels, assets, textures
 - Dynamic story creation
-

5. How Does Generative AI Work?

1 Large-Scale Training

Models are trained on:

- Books
- Websites
- Code repositories
- Images
- Videos
- Scientific papers

2 Transformer Architecture (Core Engine)

Introduced by Google (Attention is All You Need).

Transformers understand context using **self-attention**.

3 LLMs (Large Language Models)

Examples:

- GPT
- LLaMA
- Claude
- Gemini
- Mistral

They learn:

- Language patterns
- Grammar
- Knowledge
- Reasoning
- Problem-solving

4 Diffusion Models (for Images & Videos)

Used in:

- Stable Diffusion
- DALL·E 3
- Midjourney
- Sora

Process:

- Start with random noise
- Step-by-step remove noise to form an image/video

5 Reinforcement Learning (for better behavior)

RLHF — Reinforcement Learning from Human Feedback

Improves:

- Safety
 - Tone
 - Accuracy
-

6. Types of Generative AI Models

1. Text-Based Models

- GPT
- Claude
- LLaMA
- Mistral

Uses:

- Text generation
- QA
- Coding
- Summarization

2. Image Models

- Stable Diffusion
- DALL·E
- Midjourney

Uses:

- Art
- Product design

- Posters & branding

3. Video Models

- Sora
- Runway
- Pika Labs

Uses:

- Movies
- Ads
- Animation

4. Audio/Music Models

- Suno
- Udio
- ElevenLabs

Uses:

- Songs
- Dubbing
- Voice generation

5. Multimodal Models

Understand: **text + image + video + audio + code**

Examples:

- GPT-4o
- Gemini 2.0
- Claude 3.5 Sonnet

7. Challenges in Generative AI

1 Hallucinations

AI may produce incorrect information confidently.

2 Copyright Issues

Training data may include copyrighted material.

3 Data Privacy

Sensitive user info must not leak.

4 Bias

AI may reflect societal or dataset biases.

5 Computational Cost

Training LLMs requires:

- Massive GPUs (H100, MI300)
- Huge datasets
- Billions of parameters

6 Security Risks

- Deepfake generation
- Fake voices
- Misleading content

7 Ethical Use

Need guidelines and regulations.

8. Comparison: Traditional AI vs Generative AI

Feature	Traditional AI	Generative AI
---------	----------------	---------------

Goal	Predict/ classify	Create new content
Output	Labels, numbers	Text, images, code, video
Examples	Spam filter, OCR	ChatGPT, Midjourney
Techniques	ML/ DL	Transformers, LLMs, Diffusion

9. GenAI Workflow (Architecture)

1. **Input** (Prompt)
 2. **Tokenization**
 3. **Embedding generation**
 4. **Model processing** (Transformer layers)
 5. **Output generation**
 6. **Post-processing**
 7. **Response delivered**
-

10. Industry Use Case Examples (Real World)

E-commerce

- Auto product descriptions
- AI product photos
- AI chat assistant
- Dynamic pricing models

FinTech

- Fraud pattern generation
- Customer query assistants

- Financial report summarization

EdTech

- AI tutors
- Personalized learning paths
- Auto-graded assignments

Healthcare

- Radiology image generation
- Diagnostic assistance
- Doctor note generation

Media & Entertainment

- Movie scene generation
 - Script writing
 - AI voice actors
-

11. Assignment (Hands-on: 45 mins)

Part A — Text Generation

Use HuggingFace or any LLM and generate:

- A 200-word article
 - A 10-line poem
 - Summary of a paragraph
-

Part B — Image Generation

Use Stable Diffusion or DALL·E and generate:

- 3 product images
 - 1 logo concept
 - 1 poster design
-

Part C — Build a Simple Chatbot

Using:

- GPT API (OpenAI)
OR
- LLaMA 3 locally

Features:

- User asks a question
 - AI returns meaningful answer
 - Add memory (optional)
-

Part D — Short Theory

Explain in 4–5 lines each:

1. What is Generative AI?
2. Difference between Discriminative & Generative models
3. What are embeddings?
4. What is a diffusion model?
5. What is prompt engineering?