



Lecture 01: Introduction to GenAI



What is ChatGPT & How Does It Work?

- ChatGPT कोई भी जवाब **direct memory** से **fetch** नहीं करता।
- यह model **हर सवाल का जवाब स्टोर नहीं करता**, बल्कि **question के base पर reasoning** करता है।
- ChatGPT **user behavior और context** को पढ़कर जवाब देता है।
- ❌ हर question-answer को पहले से store कर पाना **possible नहीं है** क्योंकि combinations infinite हैं।



Example:

Hi, how ru? → Hi, how r u?



लेकिन हर variation का मतलब अलग हो सकता है, इसलिए fixed memory answer approach काम नहीं करता।



How ChatGPT Handles Unseen Questions

- जब आप कोई **new या modified question** पूछते हैं — जो training data में कभी नहीं देखा गया हो — तब भी:
 - Model **pattern को पहचान कर answer generate** करता है।



Example Pattern:

Input:

10000

11000

12000

13000

User asks → 14000

Model sees pattern → +1000

→ Predicts: 15000



ChatGPT का logic है: **"Understand the pattern → then predict"**



यही GPT की training का core idea है: **Pattern Recognition & Logical Generalization**

Tokenization: The First Step

- ChatGPT input को **tokens** में **break** करता है ताकि model उसे समझ सके।
- यह हर sentence को छोटे-छोटे parts (tokens) में बदलता है।

 **Tokenization = Words को model-understandable numbers में बदलना**

Tokenization Example:

Sentence: Hi, how are you?

Tokens:

hi → 36

how → 29

are → 1231


you → 820

→ [36, 29, 1231, 820]

Model adds predicted tokens:

→ [36, 29, 1231, 820, 280, 70]

- इन numbers का मतलब actual words से होता है:
 - 280 → am
 - 70 → fine

 Model इसी तरह हर जवाब को पीछे से जोड़ता जाता है।

What is GPT? And What is LLM?

GPT: *Generative Pre-Trained Transformer*

- GPT एक ऐसा AI model है जो खुद answer **generate** करता है।
- यह पुराने examples या stored answers पर निर्भर नहीं रहता।
- Model training के दौरान patterns और logic सीखता है।

LLM: *Large Language Model*

- LLM एक ऐसा model है जो बहुत massive data पर trained होता है।
- ये natural language को deeply समझने में capable होता है।

Transformer:

- एक प्रकार का deep learning architecture है।
 - Input को एक format से दूसरे format में **transform** करता है।
 - GPT इसी पर based है।
-

Same Question, Different Answers — Why?

- ChatGPT कभी-कभी **same question पर अलग answers** दे सकता है।
- यह उसके randomness, creativity और context समझने की ability के कारण होता है।

Example:

Question: Hi, how are you?

Answer 1: Hi, I'm doing well. And you?

Answer 2: Hello! I'm great, how about you?

💡 Model हर बार **slightly अलग pattern pickup** करता है → इसलिए अलग answers possible होते हैं।

Summary: Key Takeaways

- ✓ ChatGPT हर सवाल का जवाब memory से नहीं देता।
 - ✓ Model unseen inputs पर भी pattern पहचानकर जवाब देता है।
 - ✓ हर input को पहले tokenize किया जाता है (words → numbers)।
 - ✓ GPT = *Generative Pre-Trained Transformer* → जवाब खुद बनाता है।
 - ✓ LLM = *Large Language Model* → language को deeply समझता है।
 - ✓ एक ही सवाल पर multiple answers possible हैं — यह GPT की flexibility है।
 - ✓ Transformer architecture GPT का foundation है — powerful, scalable & intelligent.
-

Pro Tip for Interviews / Exams:

अगर आपसे पूछा जाए "*ChatGPT कैसे काम करता है?*", तो ये 3 words याद रखो:

Tokenization – Pattern Recognition – Generation