# AIML For beginners

## Overview

- GenAI applications like ChatGPT, Gemini, and Claude function atop foundation models, crucial for understanding AI capabilities.
- GenAI mastery offers potential earnings between ₹150,000 to ₹200,000, emphasizing the significance of expertise in this field.
- Prompts are vital instructions for GenAI; without them, no output is generated, highlighting importance in effective communication with AI.
- Prompt engineering methods outlined: Zero-shot, One-shot, Few-shot, and Chain-of-thought prompting, each enhancing AI performance in distinct ways.
- Foundation models require extensive datasets and investment, demonstrating the complexities of developing high-quality AI applications.
- Companies prefer API-based access to AI models over developing in-house due to cost-effectiveness and resource constraints.
- Tokenization is essential for model training, converting text into numerical tokens to facilitate language understanding.
- The transformer architecture enables coherent output generation and word relationships through advanced attention mechanisms.
- Technical demos showcased ChatGPT's multimodal capabilities, including text, image, and voice, enhancing practical understanding of AI applications.
- Agent AI Mastery program recommended version 0.25.09 with 18 three-hour classes, starting topics in AI and Python for forthcoming enrollment milestones.

## Notes

### Session Introduction and Attendance (05:29 - 08:10)

- Shahid K21 welcomed participants to the AI/ML/DL and GenAI for beginners session.
- Vinod Nakka confirmed previous attendance at 2nd and 3rd gen HDMK mastery sessions and scheduled a separate technical roadmap meeting for 10:00.
- Deepika attended for first time but had reviewed materials previously.

### Core GenAI Concepts (12:29 - 20:57)

- Distinguished between GenAI applications (ChatGPT, Gemini, Claude) and actual GenAI models - applications are built on top of foundation models.
- Explained Agentic AI as subset of GenAI, emphasizing GenAI mastery as pathway to earning ₹150-200k based on experience.
- Defined prompts as instructions given to GenAI applications that trigger content generation - no instruction means no output.

### Prompt Engineering Deep Dive (20:57 - 35:58)

- Covered prompt engineering as methodology for crafting effective instructions to reduce AI hallucination and improve accuracy.
- Demonstrated Zero-shot prompting: asking model to perform tasks without examples, relying on pre-existing knowledge.
- Explained One-shot prompting: providing single example to guide AI understanding and reduce errors.
- Introduced Few-shot prompting: giving multiple examples for better context and classification accuracy.
- Detailed Chain-of-thought prompting: improving reasoning ability through step-by-step instructions for complex tasks.

### Model Architecture and Processing (35:58 - 01:38:57)

- Clarified foundation models as brain behind GenAI applications, trained on massive internet-scale datasets requiring significant investment.
- Explained API-based approach: companies purchase API keys rather than building models from scratch due to cost considerations.
- Demonstrated tokenization process using OpenAI tokenizer - breaking text into smaller units (tokens) assigned numerical IDs during training.
- Covered word embeddings as numerical representations placing similar words close together in vector space for meaning relationships.
- Introduced transformer architecture for understanding word relationships and generating coherent outputs through attention mechanisms.

### Technical Demonstrations (44:12 - 54:37)

- Showed live ChatGPT interaction demonstrating text, image, and voice capabilities of multimodal foundation models.
- Explored model types: LLMs (text-only), diffusion models (multiple output types), multimodal models (multiple input/output types).
- Used transformer visualization tool to demonstrate probability calculations and word prediction processes.

### Student Support and Program Information (41:07 - 01:53:24)

- Manju clarified Agent AI Mastery program recordings - recommended version 0.25.09 with 18 three-hour classes over 0.25.08 with 26 two-hour classes.
- MeAreG (new January enrollee) received program roadmap: Cloud basics starting January 10th, AI basics from January 24th, Python for AI from February 1st, followed by AWS AI.
- Shared 1:1 consultation calendar link for personalized technical and career guidance sessions.

# Summary

## Generative AI for Beginners Overview

Shahid conducted a session on Generative AI for beginners, focusing on its components and practical applications. He explained that while ChatGPT is an AI application, it is not itself a generative AI, but rather a subset of it. Shahid emphasized the importance of prompt engineering in crafting effective instructions for generative AI applications, and demonstrated some examples of prompt engineering through a lab session. He also highlighted the potential career opportunities in generative AI, noting that a strong understanding of the subject could lead to high-paying jobs in the field.

## Understanding Prompt Engineering Techniques

Sahid explained the concepts of prompt engineering, including zero-shot, one-shot, few-shot, and chain of thought prompting, demonstrating how these techniques can be used to improve AI model performance and reduce hallucination. He also discussed the significance of fine-tuning models and the role of prompt engineering in this process. Additionally, Sahid introduced the concepts of Large Language Models (LLMs) and foundation models, explaining their capabilities and how they can handle various types of data inputs.

## Generative AI: Foundations Explained

The meeting covered foundational concepts of generative AI models, focusing on tokenization, embeddings, and transformers. Shahid explained how AI models process and understand human language by breaking down text into tokens, assigning numerical IDs, and using embeddings to represent word meanings in vector spaces. He emphasized the importance of these concepts for generating text, images, and other outputs. The session also touched on transformer models, which use attention mechanisms to understand relationships between words and generate coherent outputs. Participants were encouraged to explore these topics further and ask questions for the next session.