R.M.K
GROUP OF ENGINEERING INSTITUTIONS

R.M.K
GROUP OF
INSTITUTIONS

R.M.K
GROUP OF
INSTITUTIONS
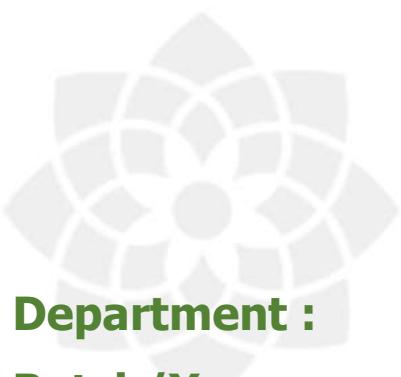
Please read this disclaimer before proceeding:

Digital Course Material
22AI005
Introduction to Generative AI

Department :		**Information Technology**
Batch/Year	**:**	**2022-2026/IV**
Created by	**:**	**Ms.G.K.Monica**
Date	**:**	**20-08-2025**

TABLE OF CONTENTS

RMK GROUP OF INSTITUTIONS

COURSE OBJECTIVES

The Course will enable learners to:

- To understand the basic concepts of Generative AI.

- To build Generative AI systems to generate images.

- To understand the concept used in Generative AI Models.

- To use various Generative AI models.

- To compare and use the various Large Language Models.

- To understand the basics of Prompt Engineering.

PRE REQUISITES

❀ PRE-REQUISITE CHART

**22AI005
INTRODUCTION TO
GENERATIVE AI**

INTRODUCTION TO GENERATIVE AI

L T P C
3 0 0 3

OBJECTIVES:

* To understand the basic concepts of Generative AI.
* To build Generative AI systems to generate images.
* To understand the concept used in Generative AI Models.
* To use various Generative AI models.
* To compare and use the various Large Language Models.
* To understand the basics of Prompt Engineering.

UNIT I INTRODUCTION   9

Generative Models – Image transformation – Challenges - Deep Neural Networks – Perceptron – back propagation – CNN – RNN – Optimizer.

UNIT II IMAGE GENERATION   9

Creating encodings of images – variational objective – Inverse Autoregressive flow – Importing CIFAR – Creating the network from TensorFlow 2.

UNIT III GENERATIVE ADVERSARIAL NETWORKS   9

Generative Adversarial Networks – Vanilla GAN – Improved GANs – Progressive GAN – Challenges – Paired style transfer – Unpaired style transfer – Deepfakes – Modes of operation – key feature set – High level flow – Replacement – Re-enactment.

UNIT IV LARGE LANGUAGE MODELS   9

Overview of LLMs - Transformers – GPT – Types of LLMs – Key concepts – other Transformers – T5 – Generative Pre-Training Models – Multi-modal Models – DALL.E 2

UNIT V PROMPT ENGINEERING   9

Basics – In-Context Learning – In-Context Prompting – Techniques – Image Prompting – Prompt Hijacking – Challenges.

TOTAL: 45 PERIODS

COURSE OUTCOME

Course Code	Course Outcome Statement	Cognitive/ Affective Level of the Course Outcome	Expected Level of Attainment
Course Outcome Statements in Cognitive Domain			
C305.1	Elaborate the basic concepts of Generative AI	Understand K2	60%
C305.2	Build Generative AI systems to generate images	Analyse K4	60%
C305.3	Apply the concepts used in Generative AI Models	Apply K3	60%
C305.4	Use various Generative AI models.	Apply K3	60%
C305.5	Compare and use the various Large Language Models	Analyse K4	60%
C305.6	Analyze the basics of Prompt Engineering.	Apply K3	60%
Course Outcome Statements in Affective domain			
C305.7	Attend the classes regularly	Respond (A2)	95%
C305.8	Submit the Assignments regularly.	Respond (A2)	95%
C305.9	Participation in Seminar/Quiz/ Group Discussion/ Collaborative learning and content beyond syllabus	Valuing (A3)	95%

RMK
GROUP OF
INSTITUTIONS

CO-PO/PSO MAPPING

Correlation Matrix of the Course Outcomes to Programme Outcomes and Programme Specific Outcomes Including Course Enrichment Activities

Course Outcomes (Cos)		PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
		K3	K4	K5	K5	K3/K5	A2	A3	A3	A3	A3	A3	A2	K3	K3	K3
C305.1	K3	3	2	1	1	3						3		3	3	3
C305.2	K4	3	3	2	2	3						3		3	3	3
C305.3	K2	2	1									2		3	3	3
C305.4	K3	3	2	1	1	3						3		3	3	3
C305.5	K4	3	3	2	2	3						3		3	3	3
C305.6	K3	3	2	1	1	3						3		2	2	2
C305.7	A2												3			
C305.8	A2								2	2	2		3			
C305.9	A3						3	3		3	3		3			
C305		3	3	2	2	3	1	1	1	3	3	3	3	3	3	3

UNIT IV

LARGE LANGUAGE MODELS

LECTURE PLAN – UNIT IV

Sl. No	TOPIC	NO OF PERIODS	PROPOSED LECTURE	ACTUAL LECTURE	PERTAINING CO(s)	TAXONOMY LEVEL	MODE OF DELIVERY
			UNIT II IMAGE GENERATION				
			PERIOD	PERIOD			
1	Creating encodings of images	1			CO2	K2	PPT
2	variational objective	1			CO2	K2	PPT
3	Inverse Autoregressive flow	1			CO2	K2	PPT
4	Importing CIFAR	1			CO2	K3	PPT
5	Creating the network from TensorFlow 2	1			CO2	K3	PPT

R.M.K
GROUP OF
INSTITUTIONS

LECTURE PLAN – UNIT IV

ASSESSMENT COMPONENTS

- AC 1. Unit Test
- AC 2. Assignment
- AC 3. Course Seminar
- AC 4. Course Quiz
- AC 5. Case Study
- AC 6. Record Work
- AC 7. Lab / Mini Project
- AC 8. Lab Model Exam
- AC 9. Project Review

MODE OF DELEIVERY

MD 1. Oral presentation
MD 2. Tutorial
MD 3. Seminar
MD 4 Hands On
MD 5. Videos
MD 6. Field Visit

ACTIVITY BASED LEARNING – UNIT IV
QUIZ- LINK

LLM:

https://datasciencedojo.com/blog/learn-llm-with-our-quiz/

GPT

https://www.jagranjosh.com/general-knowledge/chatgpt-quiz-1685343406-1

T5

https://wandb.ai/mukilan/T5_transformer/reports/Exploring-Google-s-T5-Text-To-Text-Transformer-Model--VmlldzoyNjkzOTE2

Multi-modal Models

https://forumias.com/blog/question/with-reference-to-multimodal-artificial-intelligence-ai-consider-the-following-statements-1-it-is-a-type-of-ai-that-can-process-and-understand-multiple-types-of-data-or-modalities-such-as-text-i/

DALL.E 2

https://realpython.com/quizzes/generate-images-with-dalle-and-openai-api/

C. Image Classification Model
D. Multi-modal Model

2. What is the main objective of the "Self-Attention" mechanism in Transformer models?

A. To segment input data
B. To capture relationships between all words in a sentence regardless of their distance
C. To reduce over fitting
D. To increase the model's computational complexity

3. What makes multi-modal models like DALL.E 2 unique compared to traditional text-only models?

A. They focus solely on generating text
B. They can process and generate outputs in more than one modality (e.g., text and images)
C. They require less data for training
D. They are only used for text-to-speech tasks

4. The primary function of T5 in NLP tasks is to:

A. Classify text into categories
B. Treat all NLP tasks as a text generation problem
C. Translate text between languages
D. Summarize text

5. Which of the following tasks can DALL.E 2 perform?

A. Generating images from text descriptions
B. Text classification
C. Image captioning
D. Both A and C

4.1 Overview of LLMs

Definition:

A LLM(Large Language Model) is AI(Artificial Intelligence) system which apply neural network techniques , to process and understand the human languages using self-supervised learning and to perform tasks like translation, summarization, and content generation.

There are many techniques that were tried to perform natural language-related tasks but the LLM is purely based on the deep learning methodologies. LLM are highly efficient in capturing the complex entity relationships in the text at hand and can generate the text using the semantic and syntactic of that particular language in which we wish to do so.

How do Large Language Models work?

❀ Large Language Models (LLMs) operate on the principles of deep learning, leveraging neural network architectures to process and understand human languages.

❀ These models, are trained on vast datasets using self-supervised learning techniques. The core of their functionality lies in the intricate patterns and relationships they learn from diverse language data during training.

❀ LLMs consist of multiple layers, including feed forward layers, embedding layers, and attention layers. They employ attention mechanisms, like self-attention, to weigh the importance of different tokens in a sequence, allowing the model to capture dependencies and relationships.

Architecture of LLM:

▪ A Large Language Model's (LLM) architecture is determined by a number of factors, like the objective of the specific model design, the available computational resources, and the kind of language processing tasks that are to be carried out by the LLM.

▪ The general architecture of LLM consists of many layers such as the feed forward layers, embedding layers, attention layers. A text which is embedded inside is collaborated together to generate predictions.

Important components to influence LLM architecture :

• Model Size and Parameter Count

• input representations

• Self-Attention Mechanisms

• Training Objectives

• Computational Efficiency

• Decoding and Output Generation

Types of LLM:

1. Autoregressive language models

❀ Autoregressive models generate text by predicting the next word given the preceding words in a sequence. Models such as GPT-3 fall into this category. Autoregressive models are trained to maximize the likelihood of generating the correct next word, conditioned by context.

❀ **Example**: GPT-3

2. Transformer-based models

❀ Transformers are a type of deep learning architecture used in large language models. This transformer architecture allows the model to process and generate text effectively, capturing long-range dependencies and contextual information.

❀ **Example**: Roberta (Robustly Optimized BERT Retraining Approach) by Face book AI

3. Encoder-decoder models

❀ Encoder-decoder models are commonly used for machine translation, summarization, and question-answering tasks.

❀ These models consist of two main components: an encoder that reads and processes the input sequence and a decoder that generates the output sequence. The encoder learns to encode the input information into a fixed-length representation, which the decoder uses to generate the output sequence. The transformer-based model known as the 'Transformer' is an example of an encoder-decoder architecture.

❀ **Example:** Marian (Marian Neural Machine Translation) by the University of Edinburgh

4. Pre-trained and fine-tuned models

* Many large language models are pre-trained on large-scale datasets, enabling them to understand language patterns and semantics broadly.

* These pre-trained models can then be fine-tuned on specific tasks or domains using smaller task-specific datasets. Fine-tuning allows the model to specialize in a particular task, such as sentiment analysis or named entity recognition.

* **Example**: ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately)

5. Multilingual models

* Multilingual models are trained on text from multiple languages and can process and generate text in several languages. They can be useful for tasks such as cross-lingual information retrieval, machine translation, or multilingual chat bots.

* By leveraging shared representations across languages, multilingual models can transfer knowledge from one language to another.

* **Example**: XLM (Cross-lingual Language Model) developed by Face book AI Research

6. Hybrid models

* Hybrid models combine the strengths of different architectures to achieve improved performance.

* For example, some models may incorporate both transformer-based architectures and recurrent neural networks (RNNs).

* **Example**: Until (Unified Language Model) is a hybrid LLM that integrates both autoregressive and sequence-to-sequence modeling approaches

Advantages of Large Language Models (LLMs):

- **Zero-shot learning :** LLMs can generalize to tasks without explicit training, enabling adaptability to new applications.

- **Data handling :** They process vast data efficiently, making them ideal for tasks like translation and summarization.

- **Fine-tuning :** LLMs can be fine-tuned for specific domains, allowing for continuous learning and customization.

- **Task automation :** LLMs automate language-related tasks like code generation and content creation, freeing human resources for more complex work.

Applications of Large Language Models (LLMs):

1. **Natural Language Processing (NLP) :** LLMs power tasks like sentiment analysis, translation, text summarization, and chatbots.

2. **Content Creation:** Social media posts, automating blog writing, product descriptions, and creative writing.

3. **Code Generation:** Assisting developers by generating code, providing debugging support, and automating repetitive coding tasks.

4. **Healthcare:** Analyzing medical literature, aiding diagnostics, and providing recommendations based on patient data.

5. **Customer Support:** Automating and improving customer interactions with virtual assistants and AI-driven chatbots.

6. **Education:** Personalized tutoring, language learning, and content generation for instructional materials.

7. **Business Automation:** Streamlining workflows through automated report generation, data analysis, and email responses.

4.2 Transformer

Transformer is a neural network architecture used for performing machine learning tasks. Transformer Architecture is a model that uses self-attention that transforms one whole sentence into a single sentence. Transformer based LLM models architectures, which have revolutionized natural language processing tasks that includes,

1. Input Embeddings: Text is tokenized and embedded into vectors that capture semantic and syntactic information.

2. Positional Encoding: Adds token order information to embeddings, allowing the model to consider token sequence.

3. Encoder: Uses self-attention and feed-forward neural networks to analyze text, preserving context and meaning.

➢**Self-Attention:** Weighs token importance by computing attention scores, capturing relationships between tokens.

➢**Feed-Forward Neural Network:** Applies non-linear transformations to tokens independently to capture complex interactions.
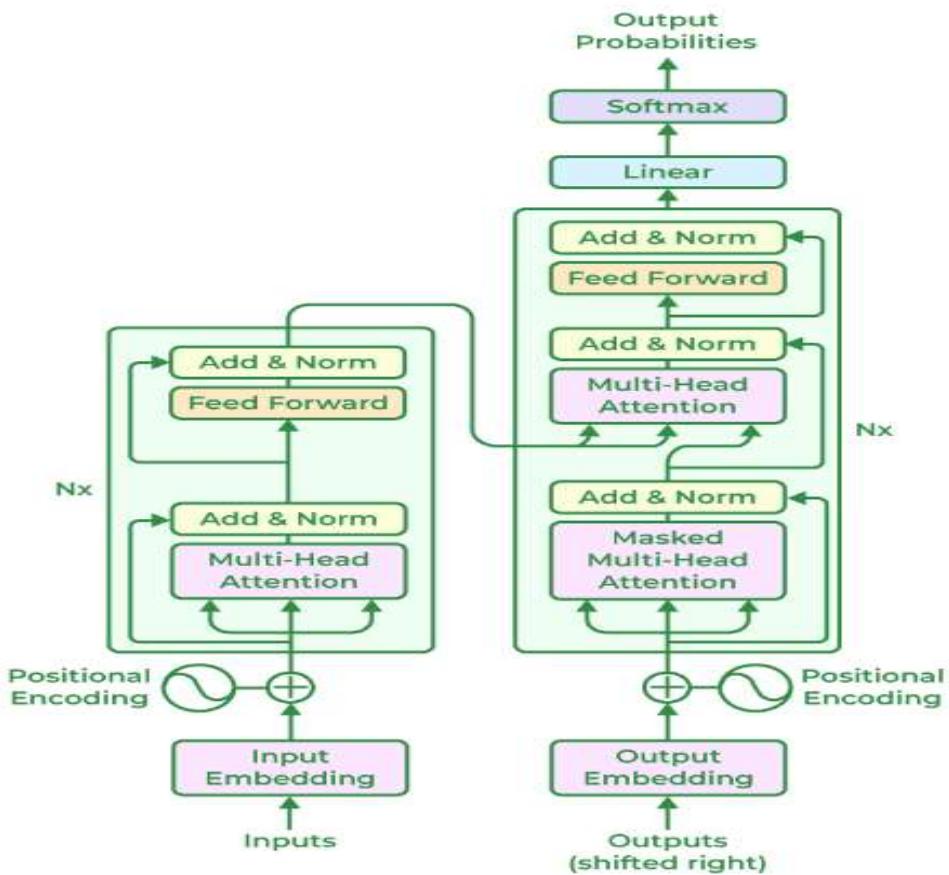
4.Decoder Layers: Autoregressive generation of sequential outputs by attending to previously generated tokens.

5.Multi-Head Attention: Performs self-attention with different weights to capture diverse relationships within the input.

6. Layer Normalization: Applied after each layer to stabilize learning and improve generalization.

7. Output Layers: Task-specific, typically involving linear projection and SoftMax for tasks like language modeling.

Transformer

- The actual architecture of transformer-based models can change and be enhanced based on particular research and model creations. To fulfill different tasks and objectives, several models like GPT, BERT, and T5 may integrate more components or modifications.

4.3 GPT:

o GPT is based on the transformer architecture, which was introduced in the paper "**Attention is All You Need**" by Vaswani et al. in 2017.

o The core idea behind the transformer is the use of self-attention mechanisms that process words in relation to all other words in a sentence, contrary to traditional methods that process words in sequential order.

o This allows the model to weigh the importance of each word no matter its position in the sentence, leading to a more nuanced understanding of language.

o As a generative model, GPT can produce new content. When provided with a prompt or a part of a sentence, GPT can generate coherent and contextually relevant continuations.

o This makes it extremely useful for applications like creating written content, generating creative writing, or even simulating dialogue.

Definition:

❖Generative Pre-trained Transformers, commonly known as GPT, are a family of neural network models that uses the transformer architecture and is a key advancement in artificial intelligence (AI) powering generative AI applications such as ChatGPT.

❖GPT models give applications the ability to create human-like text and content (images, music, and more), and answer questions in a conversational manner. Organizations across industries are using GPT models and generative AI for Q&A bots, text summarization, content generation, and search.

4.4 Types of LLM

1. Autoregressive language models

* Autoregressive models generate text by predicting the next word given the preceding words in a sequence.

* Models such as GPT-3 fall into this category. Autoregressive models are trained to maximize the likelihood of generating the correct next word, conditioned by context.

* While they excel at generating coherent and contextually relevant text, they can be computationally expensive and may suffer from generating repetitive or irrelevant responses.

* **Example**: GPT-3

2. Transformer-based models

* Transformers are a type of deep learning architecture used in large language models. The transformer model, introduced by Vaswani et al. in 2017 is a key component of many LLMs.

* This transformer architecture allows the model to process and generate text effectively, capturing long-range dependencies and contextual information.

* **Example**: RoBERTa (Robustly Optimized BERT Pre-training Approach) by Facebook AI

3. Encoder-decoder models

* Encoder-decoder models are commonly used for machine translation, summarization, and question-answering tasks.

* These models consist of two main components: an encoder that reads and processes the input sequence and a decoder that generates the output sequence.

- The encoder learns to encode the input information into a fixed-length representation, which the decoder uses to generate the output sequence.

- The transformer-based model known as the 'Transformer' is an example of an encoder-decoder architecture.

- **Example:** MarianMT (Marian Neural Machine Translation) by the University of Edinburgh

4. Pre-trained and fine-tuned models

- Many large language models are pre-trained on large-scale datasets, enabling them to understand language patterns and semantics broadly.

- These pre-trained models can then be fine-tuned on specific tasks or domains using smaller task-specific datasets.

- Fine-tuning allows the model to specialize in a particular task, such as sentiment analysis or named entity recognition.

- This approach saves computational resources and time compared to training a large model from scratch for each task.

- **Example**: ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately)

5. Multilingual models

- Multilingual models are trained on text from multiple languages and can process and generate text in several languages.

- They can be useful for tasks such as cross-lingual information retrieval, machine translation, or multilingual chatbots.

- By leveraging shared representations across languages, multilingual models can transfer knowledge from one language to another.

- **Example**: XLM (Cross-lingual Language Model) developed by Facebook AI Research

6. Hybrid models

❈ Hybrid models combine the strengths of different architectures to achieve improved performance.

❈ For example, some models may incorporate both transformer-based architectures and recurrent neural networks (RNNs). RNNs are another type of neural network commonly used for sequential data processing. They can be integrated into LLMs to capture sequential dependencies in addition to the self-attention mechanisms of transformers.

❈ **Example**: UniLM (Unified Language Model) is a hybrid LLM that integrates both autoregressive and sequence-to-sequence modeling approaches

❈ These are just a few examples of the different types of large language models developed. Researchers and engineers continue to explore new architectures, techniques, and applications to advance the capabilities of these models further and address the challenges of natural language understanding and generation.

4.5 Key concepts:

1. **Text Understanding and Generation:** Large Language Models (LLMs) understand and generate text using transformer architectures that leverage self-attention mechanisms for context and language coherence.

2. **Pre-training and Fine-Tuning:** LLMs like GPT are pre-trained on vast datasets, then fine-tuned on specific tasks to improve performance across a range of NLP applications, from text summarization to question answering.

3. **Multi-modal Extensions:** Some LLMs, like DALL·E 2, extend beyond text to handle multi-modal data, generating images from text descriptions, enhancing their versatility in real-world applications.

4.6 Other Transformers:

Bidirectional Transformers: Unlike autoregressive models, which process text sequentially from left to right, bidirectional transformers (like BERT) consider both the left and right context, leading to better understanding of sentence structure and meaning.

❀ **Example:** BERT (Bidirectional Encoder Representations from Transformers)

Masked Language Models (MLMs): These models mask certain tokens in a sentence and predict them based on the surrounding context. This forces the model to understand the context deeply to predict the missing words.

❀ **Example:** BERT

Causal Transformers: In causal language models, each token can only attend to previous tokens, which is useful for text generation tasks like auto-complete or dialogue systems.

❀ **Example:** GPT-2

Sparse Transformers: These optimize the transformer model by limiting attention to a subset of tokens, reducing the computational cost without sacrificing much accuracy in sequence processing.

❀ **Example:** Reformer

Longformer: Longformer modifies the self-attention mechanism of transformers to handle longer documents or sequences more efficiently, using a combination of global and local attention patterns.

❀ **Example:** Longformer by Allen Institute for AI

4.7 T5

❈ **Text to Text Transfer Transformer:**

❈ Data augmentation using Text to Text Transfer Transformer (T5) is a large transformer model trained on the **Colossal Clean Crawled Corpus (C4)** dataset. Google open-sourced a pre-trained T5 model that is capable of doing multiple tasks like translation, summarization, question answering, and classification.

❈ Transfer learning, where a model is first pre-trained on a data-rich task before being fine-tuned on a downstream task, has emerged as a powerful technique in natural language processing (NLP).

❈ The effectiveness of transfer learning has given rise to a diversity of approaches, methodology, and practice.

Techniques:

There are five data augmentation techniques:

1. Word Embeddings
2. BERT
3. Back Translation
4. Text to Text Transfer Transformer
5. Ensemble Approach.

Application:

• Text Classification

• Text Summarization

• Machine Translation

• Question Answering

• Text-to-SQL

• Text Generation

• Sentence Paraphrasing

4.8 Architecture of Generative Pre-trained Transformer

The transformer architecture, which is the foundation of GPT models, is made up of feed forward neural networks and layers of self-attention processes.

1. **Input Embedding**
   - **Input**: The raw text input is tokenized into individual tokens (words or subwords).
   - **Embedding**: Each token is converted into a dense vector representation using an embedding layer.

2. **Positional Encoding**: Since transformers do not inherently understand the order of tokens, positional encodings are added to the input embeddings to retain the sequence information.

3. **Dropout Layer**: A dropout layer is applied to the embeddings to prevent overfitting during training.

4. **Transformer Blocks**
   - **LayerNorm**: Each transformer block starts with a layer normalization.
   - **Multi-Head Self-Attention**: The core component, where the input passes through multiple attention heads.
   - **Add & Norm**: The output of the attention mechanism is added back to the input (residual connection) and normalized again.
   - **Feed-Forward Network**: A position-wise feed-forward network is applied, typically consisting of two linear transformations with a GeLU activation in between.
   - **Dropout**: Dropout is applied to the feed-forward network output.
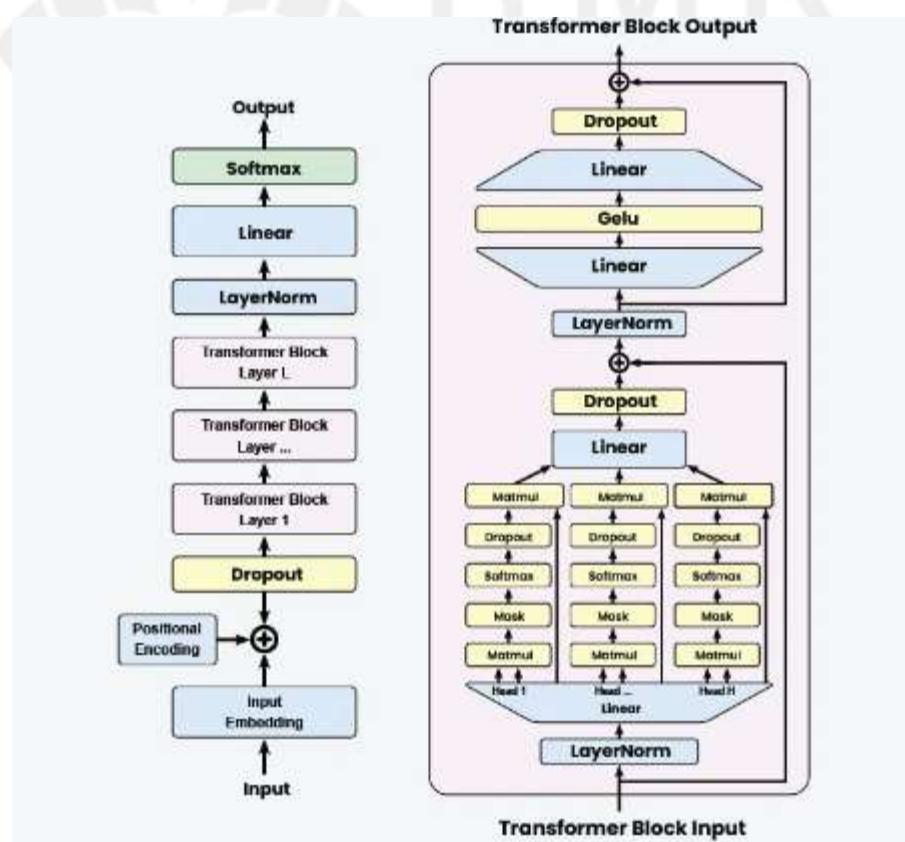
5. **Layer Stack:** The transformer blocks are stacked to form a deeper model, allowing the network to capture more complex patterns and dependencies in the input.

6. Layers

- **LayerNorm**: A final layer normalization is applied.
- **Linear**: The output is passed through a linear layer to map it to the vocabulary size.
- **Softmax**: A softmax layer is applied to produce the final probabilities for each token in the vocabulary.

Advantages of GPT:

- Flexibility
- Scalability
- Constant Availability
- Cost effective
- Contextual Understanding

Applications of Generative Pre-trained Transformer

❊ The versatility of GPT models allows for a wide range of applications, including but not limited to:

❊ **Content Creation:** GPT can generate articles, stories, and poetry, assisting writers with creative tasks.

❊ **Customer Support:** Automated chatbots and virtual assistants powered by GPT provide efficient and human-like customer service interactions.

❊ **Education:** GPT models can create personalized tutoring systems, generate educational content, and assist with language learning.

❊ **Programming:** GPT-3's ability to generate code from natural language descriptions aids developers in software development and debugging.

❊ **Healthcare:** Applications include generating medical reports, assisting in research by summarizing scientific literature, and providing conversational agents for patient support.

Conclusion:

➢ Artificial intelligence has advanced significantly with the Generative Pre-trained Transformer models, especially in natural language processing.

➢ Every version of GPT, from GPT-1 to GPT-4, has increased the capabilities of AI in terms of comprehending and producing human language.

➢ Although GPT models' capabilities present a plethora of prospects in a variety of sectors, it is imperative to tackle the ethical issues that come with them in order to guarantee their responsible and advantageous application.

➢ GPT models are expected to stay at the vanguard of AI technology evolution, propelling innovation and industry revolution.
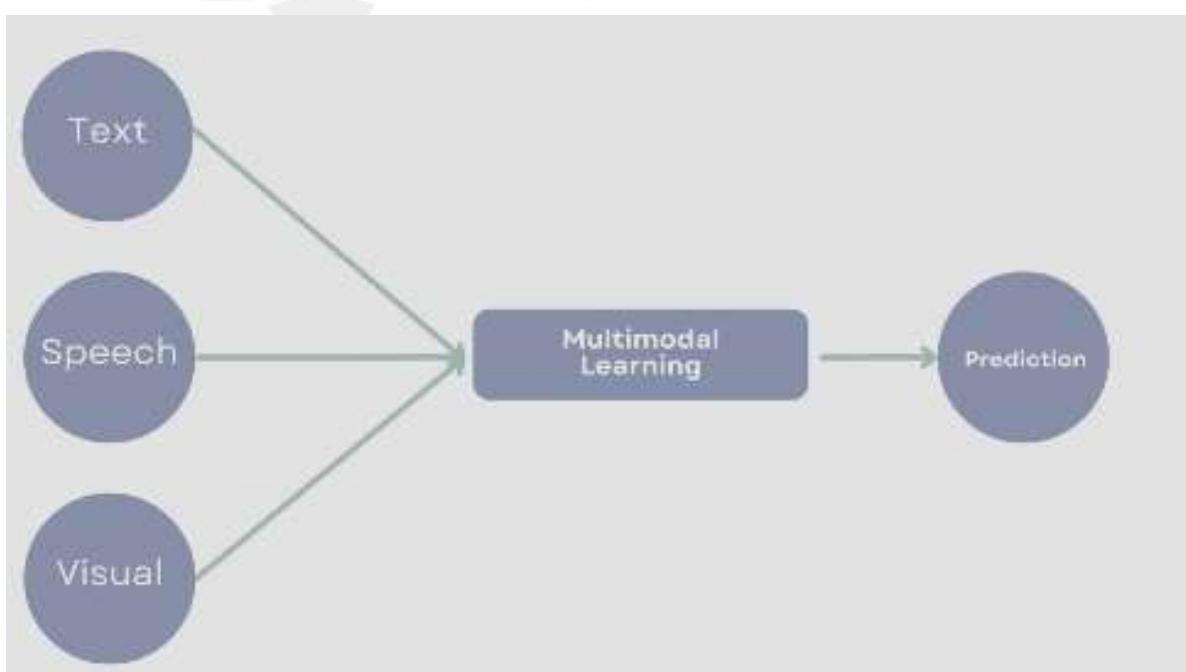
4.9 Multi-Modal Models

Definition:

•A model is considered multimodal if it can handle and integrate information from different modalities.

•For instance, an MLLM can interpret a text description, analyze a corresponding image, and generate a response that encompasses both forms of input.

•This capability allows MLLMs to perform tasks that require a nuanced understanding of various types of data, making them more versatile and powerful.

Challenges:

➢Representation

➢Fusion

➢Alignment

➢Translation

➢Co-learning

4.10 DALL.E 2

✻ **DALL.E 2 Image Generator is an AI software by OpenAI that can create realistic images and art from descriptions provided to it by the users in their own words.**

✻ DALL.E 2 is capable of creating original images and combining various concepts, properties, attributes, and styles into one.

✻ Through DALL.E 2 Image Generator, you can also make edits to already existing images, based on just a natural language caption. These edits feel completely realistic.

✻ Moreover, it can also generate different variations inspired by the original image provided to it.

How does DALL.E 2 Image Generator work?

➢ The working of DALL.E 2 is a three-step process which has been described below:

➢ First, a **text prompt is encoded** which trains to map the prompt provided to a representation space.

➢ Then, a model image is encoded from the encoded text which captures the semantic information of the prompt containing the text encoding.

➢ Finally, an image **DECODER** generates the full image, which is a visual manifestation of this semantic information.

Advantages:

▪ High-Quality Image Generation

▪ Creative Flexibility

▪ Contextual Understanding

▪ Customization

▪ Time-Saving for Design and Prototyping

ASSIGNMENT – UNIT IV

1. What is the primary purpose of GPT in natural language processing tasks? (K6)

2. Describe the core concept behind the Transformer model in LLMs. (K5)

3. What is the role of DALL.E 2 in multi-modal AI systems? (K4)

4. How do you fine-tune T5 for a specific NLP task like summarization? (K6)

5. Outline the steps to train a multi-modal model using text and images in deep learning.

PART A- UNIT-IV

1. What is the primary characteristic that defines Large Language Models (LLMs)?

LLMs are defined by their ability to process and generate human-like text based on large-scale datasets. They use deep learning architectures like Transformers to understand language context and perform tasks such as text generation, translation, and summarization.

2. How does the self-attention mechanism in Transformer models work?

The self-attention mechanism allows the model to assign different importance (weights) to various words in the input sequence. It helps capture relationships between words regardless of their position, which is crucial for understanding long-range dependencies in text.

3. What is the key difference between GPT and other Transformer models like BERT?

GPT is primarily a generative model, designed for text generation and autoregressive tasks, predicting the next word in a sequence. BERT, on the other hand, is designed for text understanding tasks like classification and is trained bidirectionally.

4. What distinguishes different types of LLMs like GPT, BERT, and T5?

The key difference lies in the type of task they are optimized for:

❀ GPT focuses on text generation.

❀ BERT is optimized for understanding and classification tasks.

❀ T5 is a versatile model that converts all tasks into a text-to-text format.

5. What is a significant advantage of using a Transformer model over traditional RNNs or LSTMs?

Transformers avoid sequential data processing, making them faster and more efficient, especially for long sequences. The self-attention mechanism allows Transformers to process the entire input sequence simultaneously rather than word by word.

6. What are the two key components of the pre-training and fine-tuning process in Generative Pre-Training models like GPT?

The two key components are:

- ❋ **Pre-training**: The model learns from large amounts of unlabeled data by predicting the next word in a sentence.

- ❋ **Fine-tuning**: The pre-trained model is further trained on a specific dataset for a particular task to adapt its learned knowledge to that task.

7. What makes T5 (Text-to-Text Transfer Transformer) unique in handling NLP tasks?

T5 treats all NLP tasks as text generation tasks. Whether it's translation, summarization, or classification, every input-output pair is framed as a text-to-text problem, making it highly versatile across different NLP applications.

8. How do multi-modal models differ from traditional NLP models?

- Multi-modal models, like DALL.E 2, can process and generate outputs across multiple modalities (e.g., text and images).

- Traditional NLP models, in contrast, are limited to processing only text data.

9. What is the role of DALL.E 2 in multi-modal learning?
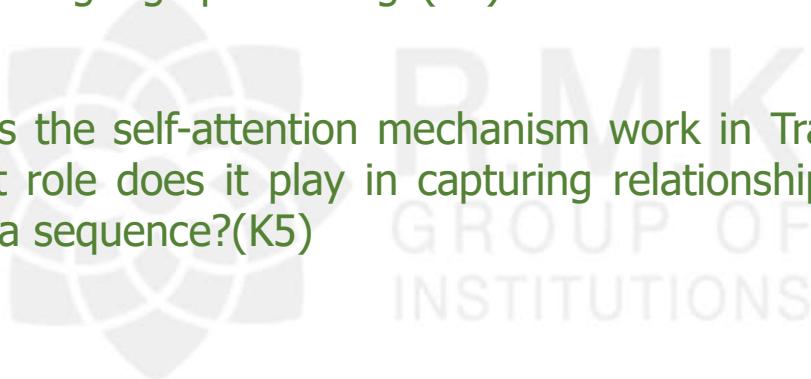
DALL.E 2 is a multi-modal model that generates images from textual descriptions. It combines the understanding of language with the ability to create visual content, advancing applications in creative industries, content generation, and design.

10. What is a common application of LLMs like GPT in real-world tasks?

LLMs like GPT are commonly used in applications such as chatbots, content generation, translation, summarization, and automated code generation, where coherent and contextually appropriate text is required.

PART B- UNIT II

1. Explain in detail about the Large Language Model(K3)

2. Compare the computational advantages and disadvantages of T5 model using its techniques.

3. Describe in detail about how GPT works.(K2)

4. Explain the process of fine-tuning a pre-trained language model for a specific downstream task. Why is this approach effective in natural language processing?(K3)

5. How does the self-attention mechanism work in Transformers, and what role does it play in capturing relationships between words in a sequence?(K5)

SUPPORTIVE ONLINE COURSES – UNIT IV

✹ https://aws.amazon.com/what-is/large-language-model/#:~:text=Large%20language%20models%2C%20also%20known,decoder%20with%20self%2Dattention%20capabilities

✹ https://aws.amazon.com/what-is/gpt/

✹ https://medium.com/fenwicks/tutorial-2-94-accuracy-on-cifar10-in-2-minutes-7b5aaecd9cdd

✹ https://www.kdnuggets.com/2023/03/multimodal-models-explained.html

Real Time Applications in Day to Day life and to Industry

1.Personal Assistants:

🏵**Smart Home Automation:** LLM-powered assistants like Siri, Alexa, or Google Assistant help control smart devices, answer questions, and provide reminders.

🏵**Chatbots:** LLMs are embedded in customer service bots to provide instant responses to queries, saving time and providing accurate information.

2.Education & Learning:

🏵**Tutoring & Homework Help:** LLMs can explain complex topics, assist with problem-solving, and provide step-by-step solutions in subjects like math, science, or language learning.

🏵**Language Translation:** LLMs power real-time language translation apps, making communication across languages easier for travelers and professionals.

3.Content Creation:

🏵**Writing Assistance:** Tools like Grammarly and AI-based content generators assist in drafting emails, reports, articles, or creative writing with suggestions on tone, grammar, and style.

🏵**Social Media Management:** LLMs generate social media posts and content ideas, optimizing engagement based on trending topics.

PRESCRIBED TEXT BOOKS AND REFERENCE BOOKS

TEXT BOOKS:

1. CLA'S SELECTED QUESTIONS & ANSWERS ON Transfer of property act,1882, & easement act,1882 2023 in english

2. A Text Book on Chat GPT and Artificial Intelligence (Paperback, Dr. ZAREENA SULTANA)

REFERENCES:

1. L Pineda, G Garza - Computational Linguistics, 2000 - direct.mit.edu

2. G Zames - IEEE Transactions on automatic control, 1981 - ieeexplore.ieee.org

3. O Katar, D Özkan, Ö Yıldırım… - Turkish Journal of …, 2023 - dergipark.org.tr

MINI PROJECT SUGGESTIONS

❋**TEAM 1:** Detecting Large Language Model

❋**TEAM 2:** Sudoku Solver using GPT

❋**TEAM3:** Experiencing T5 model

❋**TEAM 4:** Sudoku Solver Multi Modal Model

THANK YOU

R.M.K
GROUP OF
INSTITUTIONS