`

RV UNIVERSITY Go, change the world an initiative of RV EDUCATIONAL INSTITUTIONS	USN **School of Computer Science and Engineering** **B.Tech (Hons.)** **Midterm Question Paper (SET 1 – Answer Scheme)** **Academic Year 2024-2025**

Course: Foundations of Generative AI		**Course Code: CS3234**	**Semester: V**
Time:	**Duration: 90 minutes**	**Date :**	**Max Marks: 25**

Notes/ Instructions:

a) Answer all questions

Sl. No.	PART A – (MCQs) Max Marks(5)	Marks	L1-L6	CO
1.	Mark the correct statement regarding feedforward neural networks a) Feedforward neural networks do not use backpropagation for training b) Feedforward neural networks only use backpropagation for testing c) Feedforward neural networks use backpropagation to adjust weights during training d) Feedforward neural networks do not require any training mechanism **Answer: C**	**1**	**L2**	**CO1**
2.	The technique that allows control over the style and tone of generated text is? a) Chain of Thought b) Prompt Persona Pattern c) Denoising Techniques d) LSTM **Answer: B**	**1**	**L3**	**CO2**
3.	The role of the "Attention Mechanism" in transformer architectures is? a) To filter noise from input data b) To focus on specific parts of input data c) To store memory for long sequences d) To handle data embeddings **Answer: B**	**1**	**L3**	**CO3**
4.	The following is NOT a characteristic of vector databases? a) Support for multi-dimensional indexing	**1**	**L2**	**CO1**

	b) High computational efficiency in querying vectors c) Relational joins across tables d) Approximate search techniques for large-scale data **Answer: C**			
5.	The main challenge in training the reverse process of diffusion models is? a) Predicting the noise correctly at each step b) Generating the final output in one step c) Reducing the model's complexity d) Controlling the learning rate **Answer: A**	**1**	**L2**	**CO3**

Sl. No.	**PART B – Max Marks (20)**	**Marks**	**L1-L6**	**CO**
6.	a. Calculate the output size of a convolutional layer if the input image is of size 32x32, and the filter size is 5x5 with a stride of 1 and padding of 2. (2 Marks) **Answer: Output size:** 32x32 The formula for calculating output size of a convolutional layer is: Output size=(W−F+2P)/S + 1 Where: W=32 (input size), F=5 (filter size), P=2 (padding), S=1(stride) Substituting values: you get 32 **Output size:** 32x32 **b.** Describe the key principles of Prompt Engineering, mention any 3 prompt engineering techniques and their respective applied use case scenarios (3 Marks) **Answer:** Key Principles of Prompt Engineering: • **Clarity:** Make your prompt as clear as possible. Ambiguity in the prompt can lead to varied responses, potentially derailing the intended context or output • **Specificity:** General prompts can lead to general outputs. The more specific your prompt, the more likely you'll get a specific, useful response	**5**	**L3**	**CO1 (6a)** **CO2 (6b)**

	• **Contextual Consistency:** Your prompt should be contextually aligned with the desired output. If the context is not maintained, the model may produce irrelevant or inconsistent results • **Instructiveness:** Include explicit instructions within your prompt for better control over the output. You can instruct GPT-4 about the format, tone, or structure of the expected response • **Concision:** Brevity is essential. Although GPT-4 can handle large prompts, concise prompts generally make for more effective interactions • **Iterative Refinement:** Experimentation is key. Outputs can be improved by iteratively refining the prompt based on feedback Techniques and Use Cases: 1. Zero-Shot Prompting: o Use Case: Asking a model to perform a task without providing examples, such as asking a model to summarize text without prior training data. 2. Few-Shot Prompting: o Use Case: Providing a few examples of desired input-output pairs to guide the model, like showing examples of sentiment analysis for specific sentences. 3. Chain-of-Thought Prompting: o Use Case: Encouraging a model to reason step-by-step, useful in mathematical problem-solving tasks where the process matters.			
7.	a. Explain the Attention Mechanism used in Transformer architectures. How does it improve upon traditional RNNs? (2 Marks) **Answer:** This is brief answer for your reference **Attention Mechanism:** In Transformer models, attention allows the model to focus on specific parts of the input sequence while processing each token, by calculating weighted relevance scores between tokens. **Improvement Over RNNs:** Traditional RNNs suffer from long-range dependency issues, as they process input sequentially. Attention mechanisms enable parallel processing and can directly access relevant information from any part of the sequence, overcoming the limitation of sequential dependence. b. How would you adapt a transformer architecture to handle multimodal inputs like both text and image? (3 Marks) **Answer:** This is brief answer for your reference 1. **Separate encoders** for text and image inputs (e.g., transformer for text, CNN for images) 2. **Fusion mechanisms** (e.g., concatenation or cross-attention) to combine text and image representations	**5**	**L3**	**CO3**

	3. Use a **shared decoder** or task-specific heads based on the final task 4. Incorporate **position embeddings** to preserve order and spatial information 5. Pretrain on multimodal tasks and fine-tune for specific use cases This architecture allows the transformer model to effectively process and learn from multimodal data, making it suitable for tasks like **image captioning**, **visual question answering**, and **multimodal classification**.			
8.	a. Explain how Variational Autoencoders (VAEs) can be used to perform data augmentation for image datasets. Provide a scenario where this could be useful (3 Marks) **Answer:** This is brief answer for your reference, detailed answer should use diagram, other details • **VAEs for Data Augmentation:** VAEs generate new samples by encoding input data into a latent space and adding random noise during the sampling process. This allows for the creation of new, realistic images that can expand a dataset. • **Scenario:** In medical imaging, VAEs can generate new, synthetic scans of rare diseases to augment training data for machine learning models, helping to improve diagnosis performance. b. For a Diffusion model, given the following: At step t=1, βt (at t=1) =0.2, the initial image pixel value is x0 (at t=0) =250, the noise added follows $\epsilon \sim N(0, 1)$, where the sampled noise ϵ=0.7. Calculate the noisy pixel value x1 step t=1. **(2 Marks)** **Final Answer: 223.92 (refer notes or slides for formula)**	**5**	**L3**	**CO3**
9.	a. Given the query Q=[1, 2], the key K=[2, 1], and d(model) =4, calculate the scaled dot-product attention score (2 Marks) **Answer:** • **Dot product:** $Q \cdot K = (1*2) + (2*1) = 4$ • **Scaled dot-product attention score:** Dot Product/sqrt(d (model)) = 4/2= 2 **Final Answer:** 2 b. In a transformer model, if d (model) =256 and the model uses 4 heads, what is the dimension d (k) for each head? (1 Mark)	**5**	**L3**	**CO3**

	Answer:
	d(k)=d(model)/number of heads=256/4=64
	
	Final Answer: 64
	
	c. Compare architecture of GPT from other transformer models like BERT, identify the differences and their impact? (2 Marks)
	
	Answer: This is brief answer for your reference, detailed answer should use diagram, other details
	
	• **GPT (Generative Pretrained Transformer):**
	  ○ **Architecture:** Uses a unidirectional (left-to-right) architecture, focusing on predicting the next token in a sequence.
	  ○ **Impact:** Excellent for generative tasks like text generation but limited in capturing context from the right of the token.
	• **BERT (Bidirectional Encoder Representations from Transformers):**
	  ○ **Architecture:** Uses a bidirectional attention mechanism, where context from both the left and right of each token is considered.
	  ○ **Impact:** More suited for understanding and classification tasks, such as question answering and sentiment analysis.			

Course Outcomes

CO 1 Explore fundamental concepts of Generative AI and its diverse applications in various industries
CO 2 Apply prompt engineering principles when working with LLMs such as ChatGPT
CO 3 Utilize multimodal LLMs to generate text, audio, image, and video content
CO 4 Apply generative AI skills to develop a project that addresses a real-world use case

Marks Distribution									
L1	**L2**	**L3**	**L4**	**L5**	**L6**	**CO1**	**CO2**	**CO3**	**CO4**
0	3	22	0	0	0	5	4	16	0