U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

June 2025 Semester End Main Examinations

Programme: B.E. **Semester: 07**

Branch: Artificial Intelligence & Machine Learning **Duration: 3 hrs.**

Course Code: 24AM7PCGAL **Max Marks: 100**

Course: Generative AI with Large Language Models

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

		UNIT - I	***CO***	***PO***	**Marks**
1	a)	Derive the Evidence Lower Bound (ELBO) used in variational autoencoders, given the following relationships: $\log p(x)=E_{q(z\|x)}$ [$\log p(x\|z)$]$-D_{KL}$ ($q(z\|x)$ || $p(z)$)	*CO1*	*PO3*	**07**
	b)	Distinguish between Generative and Discriminative models.	*CO2*	*PO2*	**06**
	c)	Given mean(x)=4.8, mean(y)=5.9, stddev(x)=1.1, stddev(y)=0.8, random sample for x: ϵ_x=-0.4, random sample for y: ϵ_y=0.6 Calculate latent space variable (z) and the reconstructed data points ($\hat{x}$, $\hat{y}$) of a VAE model.	*CO3*	*PO3*	**07**
		OR			
2	a)	Explain how the Convolutional Neural Networks (CNNs) can be applied for Generative AI tasks.	*CO1*	*PO2*	**06**
	b)	For a VAE with latent variable $z \sim N(\mu,\sigma^2)$, derive the gradient of the reparameterization trick $z = \mu + \sigma \cdot \epsilon$, where $\epsilon \sim N(0,1)$. Explain its significance in training.	*CO2*	*PO3*	**08**
	c)	Given: Input (X): [0.5, 0.6, 0.7, 0.8] Encoder Weights (W_e): [[0.3, 0.2], [0.5, 0.4], [0.2, 0.3], [0.4, 0.5]] Encoder Biases (b_e): [0.1, 0.2] Decoder Weights (W_d): [[0.4, 0.3, 0.2, 0.5], [0.5, 0.4, 0.3, 0.2]] Decoder Biases (b_d): [0.15, 0.1, 0.05, 0.2] Illustrate the encoder-decoder process of an Autoencoder by i. Constructing a latent space representation Z of the given input vector X. ii. Reconstructing the input vector 'X' using results of (i).	*CO3*	*PO3*	**06**

			UNIT - II			
	3	a)	In a transformer model using the dot-product attention mechanism ,given the input vectors Q, K, and V with dimensions d_q, d_k, and d_v respectively, derive the computational complexity of computing the attention mechanism for a sequence of length N.	*CO1*	*PO2*	**07**
		b)	Elaborate on Multi GPU compute strategies with respect to parallelism.	*CO2*	*PO1*	**06**
		c)	Apply tokenization, embedding (d_{model} = 4, row major embedding from 0.1 to 2.0) and positional encoding processes of a Transformer Model on the sentence "AI is amazing.(fullstop)" and Compute the Final Embedded Tokens matrix.	*CO3*	*PO1*	**07**
			OR			
	4	a)	In what ways does the BloombergGPT model outperform general-purpose models in financial tasks? Explain using two specific financial applications highlighting its advantages.	*CO2*	*PO2*	**05**
		b)	Differentiate between one-shot prompting and few-shot prompting.	*CO1*	*PO2*	**05**
		c)	Using the Layer Normalization parameters: Scaling Factor(γ)=2, shifting Factor(β)=1, apply Layer Normalization to the input vector [1,3,5,7] and determine the transformed output vector.	*CO3*	*PO3*	**10**
			UNIT - III			
	5	a)	Write architecture of T5 model and explain its working.	*CO1*	*PO2*	**07**
		b)	In Reinforcement Learning with Human Feedback (RLHF), given a sequence of 3 actions with rewards [8, 12, 20] and a discount factor of 0.85, calculate the discounted cumulative reward for the sequence.	*CO1*	*PO3*	**05**
		c)	Derive the mathematical formulation of LoRA.	*CO3*	*PO3*	**08**
			OR			
	6	a)	Elaborate on Bilingual Evaluation Understudy (BLEU) metric.	*CO2*	*PO2*	**07**
		b)	Calculate the total reward received over 5 iterations in RLHF, given a 20% reward decay per iteration and an initial reward of 100.	*CO1*	*PO3*	**05**
		c)	Explain parameter-efficient fine-tuning (PEFT) techniques and how LoRA differs from traditional fine-tuning.	*CO3*	*PO3*	**08**
			UNIT - IV			
	7	a)	Explain the retrieval module and the generator module of a RAG along with required mathematical equations.	*CO1*	*PO2*	**10**

		b)	The reward model $R_{chatbot}$ is defined as a weighted combination of response relevance (R_{rel}) and user engagement (R_{eng}). Given: Weight W_{rel} = 0.6, W_{eng} = 0.4 Relevance Scores (R_{rel}) = [0.8, 0.9, 0.7] Engagement Scores (R_{eng}) = [0.6, 0.8, 0.9] i. Derive the $R_{chatbot}$ relation. ii. Compute $R_{chatbot}$ scores with respect to the R_{rel} and R_{eng} values recorded. iii. Determine the best response and justify the selection.	*CO3*	*PO3*	**10**
			OR			
	8	a)	Using the respective mathematical representations, explain the construction process of a reward model.	*CO1*	*PO2*	**10**
		b)	Derive the clipped surrogate objective of Proximal Policy Optimization (PPO) technique and determine the $L_{CLIP}(\theta)$ value using the provided data: Probability ratio: $r_t(\theta)$=1.2 Advantage estimate of time: A_t=0.5 Clipping parameter: ϵ=0.1	*CO3*	*PO3*	**10**
			UNIT - V			
	9	a)	Elaborate on the generator and the discriminator objective function of a Style Transfer Generative Adversarial Network (GAN).	*CO2*	*PO1*	**06**
		b)	Given a GAN with a discriminator loss $L_D = -E\,[\log(D(x))] - E[\log(1-D(G(_z)))]$ derive the updates for the discriminator weights.	*CO1*	*PO2*	**06**
		c)	In a Deep Convolutional Generative AI Network (DCGAN) i. If the Discriminator (D) classifies the generated image as real with the probability of 0.75, Compute the Generator Loss. ii. If the Discriminator classifies real images and fake images correctly with probabilities of 0.92, and 0.85. Compute the Discriminator Loss.	*CO3*	*PO3*	**08**
			OR			
	10	a)	Explain the step-by-step procedure to be followed while training a GAN model.	*CO2*	*PO2*	**06**
		b)	In WGAN, the Wasserstein loss is given by $L_D = E[D(x)] - E[D(G(z))]$ Explain how the gradient penalty term is added to enforce the Lipschitz constraint.	*CO1*	*PO2*	**06**
		c)	Given the probabilities: Conditional GAN (CGAN): • Discriminator Loss (L_D) for real images: 0.4. • Generator Loss (L_G) for the generated images: 0.85. Wasserstein GAN (WGAN):	*CO3*	*PO3*	**08**

			• Discriminator Loss (L_D) for the real images: 0.22. • Generator Loss (L_G) for the generated images: 0.69. i. Interpret the Discriminator Loss and Generator Loss scores of both the models in terms of training behavior for both models. ii. Given that both models have 200 neurons in the first layer, and the generator in both cases has a latent vector of size 200 and a class label vector of size 4, with a total bias of 200. Calculate the number of parameters in the conditional input layer for both CGAN and WGAN.			

******