Assignment Code: DS-AG-031

Generative AI - Text Generation and Machine Translation | **Assignment**

Instructions: Carefully read each question. Use Google Docs, Microsoft Word, or a similar tool to create a document where you type out each question along with its answer. Save the document as a PDF, and then upload it to the LMS. Please do not zip or archive the files before uploading them. Each question carries 20 marks.

Total Marks: 200

Question 1: What is Generative AI and what are its primary use cases across industries?

Answer:

Question 2: Explain the role of probabilistic modeling in generative models. How do these models differ from discriminative models?

Answer:

Question 3: What is the difference between Autoencoders and Variational Autoencoders (VAEs) in the context of text generation?

Answer:

Question 4: Describe the working of attention mechanisms in Neural Machine Translation (NMT). Why are they critical?

Answer:

Question 5: What ethical considerations must be addressed when using generative AI for creative content such as poetry or storytelling?

Answer:

Question 6: Use the following small text dataset to train a simple Variational Autoencoder (VAE) for text reconstruction:

["The sky is blue", "The sun is bright", "The grass is green",
"The night is dark", "The stars are shining"]

1. Preprocess the data (tokenize and pad the sequences).
2. Build a basic VAE model for text reconstruction.
3. Train the model and show how it reconstructs or generates similar sentences.

Include your code, explanation, and sample outputs.

(*Include your Python code and output in the code box below.*)

Answer:

Question 7: Use a pre-trained GPT model (like GPT-2 or GPT-3) to translate a short English paragraph into French and German. Provide the original and translated text.

(*Include your Python code and output in the code box below.*)

Answer:

Question 8: Implement a simple attention-based encoder-decoder model for English-to-Spanish translation using Tensorflow or PyTorch.

(*Include your Python code and output in the code box below.*)

Answer:

Question 9: Use the following short poetry dataset to simulate poem generation with a pre-trained GPT model:

["Roses are red, violets are blue,",
"Sugar is sweet, and so are you.",
"The moon glows bright in silent skies,",
"A bird sings where the soft wind sighs."]

Using this dataset as a reference for poetic structure and language, generate a new 2-4 line poem using a pre-trained GPT model (such as GPT-2). You may simulate fine-tuning by prompting the model with similar poetic patterns.

Include your code, the prompt used, and the generated poem in your answer.

(*Include your Python code and output in the code box below.*)

Answer:

Question 10: Imagine you are building a creative writing assistant for a publishing company. The assistant should generate story plots and character descriptions using Generative AI. Describe how you would design the system, including model selection, training data, bias mitigation, and evaluation methods. Explain the real-world challenges you might face.

(*Include your Python code and output in the code box below.*)

Answer: