Lesson 3: Variational Autoencoders and Generative Adversarial Network

Overview:

This assignment is designed to enhance understanding and skills in advanced AI technologies through engaging tasks. It focuses on developing practical solutions using AI in various creative scenarios. The objective is to deepen knowledge in the application of AI for innovative and generative tasks, emphasizing hands-on experience in utilizing AI for diverse creative outputs.

Instructions:

1. Read the problem carefully
2. Identify the concepts that are relevant to the scenarios
3. Apply the concepts learned to solve the problem
4. Write your solution and explain why you chose it

Task:

Select between VAE and GAN to determine which is best suited for the given scenarios:

Scenario 1: You want to create an app that can generate realistic images of animals based on text descriptions, such as **a blue cat with green eyes** or **a unicorn with rainbow wings**.

Scenario 2: You want to create a music generator to compose original melodies and harmonies based on a genre, mood, or theme.

Discussion Questions (Optional)

If time permits, discuss the below question:

a) Discuss the architecture you have chosen, explaining why and how it fits that scenario.

Answer Key

Problem 1:

Generative model type: Generative Adversarial Networks (GANs)

Explanation: GANs are well-suited for generating realistic images. In this scenario, the GAN can be trained on a dataset of animal images and text descriptions. The generator network of the GAN can take the text description as input and generate an image that matches the description. The discriminator network can then evaluate the generated image's realism and provide feedback for further improvement.

Problem 2:

Generative model type: Variational Autoencoders (VAEs)

Explanation: VAEs are suitable for this scenario because they can learn the underlying distribution of music data and generate new samples based on that distribution. The VAE can be trained on a dataset of music compositions from different genres, moods, or themes. By sampling from the latent space of the VAE, it can generate new melodies and harmonies that align with the given genre, mood, or theme.