AI3603: Computer Vision, Spring 2025

Indian Institute of Technology Hyderabad
Homework 2, Generative Modelling
25 points. Assigned 16.03.2025, Due **11:59 pm on 22.03.2025**

The essence of the scientific spirit is to realize what a wonderful world it is that we live in. – Sir C. V. Raman

Instructions:

- It is **strongly recommended** that you work on your homework on an *individual* basis. If you have any questions or concerns, feel free to talk to the instructor or the TAs.
- Use the MNIST dataset.
- You are free to use Copilot. Please turn in your prompts.
- Please turn in Python Notebooks with the following notation for the file name: `your-roll-number-hw2.ipynb`.

Questions:

1. **Variational Autoencoder (VAE) (10)**
   (a) Train a VAE on the MNIST dataset. Describe your architecture. Use the reparameterization trick to ensure gradient propagation.
   (b) We discussed in class that $q(z|x)$ is a Gaussian distribution that whose parameters are learned to be "close" to the standard normal distribution. Show that the latent distribution $q(z|x)$ does indeed vary with the conditioning variable x. Why is this important?
   (c) Visualise the generative process of the VAE by sampling from the latent space. Do "nearby" samples in the latent space correspond to "similar" samples in the image space?
   (d) Verify qualitatively and quantitatively that the model has indeed learned to generate *diverse samples of good quality*.
2. **Denoising Diffusion Probabilistic Model (DDPM) (10)**
   (a) Train a DDPM on the MNIST dataset. Use a U-Net model and Gaussian noise for diffusion.
   (b) Experiment with different time steps for the diffusion process. How does the number of steps affect the quality and diversity of the generated samples?
   (c) Visualise the generative process of the DDPM by starting from a noise image. Use an appropriate number of steps for the diffusion process.
   (d) Do "nearby" samples in the latent space lead to "similar" samples in the image space?
   (e) Verify qualitatively and quantitatively that the model has indeed learned to generate *diverse samples of good quality*.
3. **Comparing DDPM and VAE (5)**
   (a) Fundamentally, how does a DDPM differ from a VAE?
   (b) How does a DDPM compare with a VAE in terms of sample quality and diversity?
   (c) How does a DDPM compare with a VAE in terms of training time?