CSET419 – Introduction to Generative AI

Lab – 3

Objective

The objective of this lab is to understand and implement a **Variational Autoencoder (VAE)** to:

- Learn latent representations of data
- Generate **diverse and novel samples** from a learned probability distribution
- Understand the role of **encoder, decoder, latent space, and KL-divergence**

Learning Outcomes

After completing this lab, students will be able to:

1. Explain the difference between Autoencoders and Variational Autoencoders
2. Implement a VAE using a deep learning framework
3. Train a VAE on image data
4. Generate new samples from the latent space
5. Analyze reconstruction quality and sample diversity

Theory Overview

What is a Variational Autoencoder (VAE)?

A Variational Autoencoder is a generative model that learns a probability distribution over the data. Unlike standard autoencoders, VAEs learn a latent distribution rather than fixed latent vectors.

Experiment:

Task 1: Dataset Preparation

- Load the dataset
- Normalize the input data
- Split into training and testing sets

Task 2: Build the VAE Architecture

- Design the Encoder network

- Compute latent mean (μ) and log-variance ($\log \sigma^2$)
- Apply the reparameterization trick
- Design the Decoder network

Task 3: Define the Loss Function

- Implement reconstruction loss (Binary Cross-Entropy or MSE)
- Implement KL divergence loss
- Combine both losses

Task 4: Train the VAE

- Train the model for sufficient epochs
- Monitor training and validation loss

Task 5: Sample Generation

- Sample random vectors from standard normal distribution
- Pass them through the decoder
- Visualize generated samples

Task 6: Latent Space Visualization (Optional)

- Reduce latent dimension to 2
- Plot latent representations

Dataset

Use any one of the following datasets:

- MNIST (recommended)
- Fashion-MNIST

Expected Output:

1. Trained VAE model

2. Reconstructed images

3. Newly generated samples
4. Loss curves