Title: Learning Probability Density Functions using data only

Dataset: Consider NO_2 concentration as the feature (x).

Dataset link: https://www.kaggle.com/datasets/shrutibhargava94/india-air-quality-data

Objective

To learn an unknown probability density function of a transformed random variable using a Generative Adversarial Network (GAN).

Tasks to be Performed

Step-1: Transform each value of x into z using the transformation function given below.

$$z = Tr(x) = x + a_r * \sin(b_r * x) \quad \textbf{(1)}$$

where $a_r = 0.5 * (r \bmod 7)$, $b_r = 0.3 * (r \bmod 5 + 1)$

where, mod returns remainder and r is your UNIVERSITY ROLL NUMBER.

Step-2: PDF Estimation using GAN (Core Task)

In this assignment, no analytical form of the probability density function is given.

You are required to:

1. Assume that the transformed variable (z) is sampled from an unknown distribution
2. Design and train a Generative Adversarial Network (GAN) to learn this distribution
3. Use the generator network to implicitly model the probability density of (z)

Important Notes

- The GAN must learn the distribution only from samples of (z)
- No parametric PDF (Gaussian, exponential, etc.) should be assumed
- The discriminator should distinguish between:
    - Real samples: (z)
    - Fake samples: (z_f = G(error)), where (error follows N(0,1))

Step-3: PDF Approximation from Generator Samples

After training the GAN:

1. Generate a large number of samples (z_f) from the generator
2. Estimate the probability density (p_h (z)) using:
    - Histogram density estimation **or** Kernel Density Estimation (KDE)

What to Submit

- Transformation parameters (a_r, b_r)
- GAN architecture description
- PDF plot obtained from GAN samples
- Observations on:
    - Mode coverage
    - Training stability
    - Quality of generated distribution

Important Instructions

- **Deadline:** 29 Jan 2026, 11:59 PM
- Late submissions will receive **zero marks**
- Technical issues must be reported **before the deadline**
- **Use of AI-based tools is strictly prohibited**
- Submissions exhibiting AI-generated patterns will be treated as **academic misconduct**
- If AI usage is detected → **Marks = 0**