Instead of midterm quiz Project PhD

Students will implement a minimal **Convolutional VAE** and a **DCGAN** on **CIFAR-10** (or CelebA-64 if GPU allows), train both under a small compute budget, and produce a short comparative analysis that demonstrates understanding of objectives, optimization behaviour, and representation vs sample quality.

Primary goals (what students must *show* in 10 days)

- Working, reproducible code for a conv-VAE and a DCGAN.
- Quantitative evaluation: reconstruction loss (VAE), linear-probe accuracy on frozen VAE latents, and FID for GAN samples (use small sample counts if needed).
- Short analysis (≤4 pages) explaining differences, two failure modes observed, and one justified mitigation each.
- Reproducible reproduce.sh that runs the main experiment end-to-end.

Minimal scope (what to implement)

1. **Dataset:** CIFAR-10 (32×32 color). Use standard train/val split. (If GPU scarce, use 10k subsample.)
2. **VAE:** Conv encoder → μ, logσ; latent dim = 64. Decoder mirrors encoder. Use BCE or MSE as appropriate.
    - Baseline ELBO (β = 1). Optionally run β = 4 for 1 ablation.
3. **GAN:** DCGAN (standard conv-transpose generator + conv discriminator). z dim = 100.
    - Train with standard GAN loss or hinge loss. WGAN-GP optional if time permits.
4. **Evaluations:**
    - VAE: recon grid, latent interpolation, log ELBO components.

    - Representation: linear probe (logistic regression) on frozen VAE z (report accuracy).
    - GAN: sample grid and FID (use pytorch-fid or similar; if compute limited, compute FID on 2k samples).
5. **Deliverables:** code repo, reproduce.sh, notebook with images & metrics, short report (≤4 pages), 5-slide presentation.

Minimal hyperparameters (fast, practical)

- Optimizer: Adam. VAE lr = 1e-3, betas=(0.9,0.999). GAN lr = 2e-4, betas=(0.5,0.999).
- Batch size: 128 (or 64 if GPU mem limited).
- VAE latent dim: 64.
- GAN z dim: 100.
- VAE epochs: 15–30 (practical tradeoff). GAN epochs: 30–80 (shorter if necessary).
- Random seed fixed and logged.

Deliverables checklist (submit exactly these)

- code/ with training & eval scripts, clear README.
- reproduce.sh (single command runs main experiments).
- notebook.ipynb with sample grids + metric tables.
- report.pdf (≤4 pages) and slides.pdf (≤5 slides).
- results/ folder with: vae_recons.png, gan_samples.png, fid.txt, linear_probe.csv.
- **Reproducibility — 30 pts**

Single command (reproduce.sh) runs and reproduces main figures/metrics; environment & seeds documented.

- **Core experiments & metrics — 30 pts**

  VAE reconstructions + latent interpolations; frozen-latent linear probe; GAN samples + FID (report sample count).

- **Analysis & insight — 25 pts**

  Explains *why* each model behaved as observed, lists two failure modes, and proposes one justified mitigation per model.

- **Code quality & documentation — 10 pts**

  Clean, modular code, clear README, sensible defaults, and saved checkpoints.

- **Report & presentation — 5 pts**

  Concise report (≤4 pages) with figures/tables and a 5-slide deck summarizing key findings.