Advanced Machine Learning (CS60073)

Assignment 2: 100 Marks

Instructions:

- Submit a single jupyter notebook for the entire assignment
- Annotate your solution using markdown before every code cell

---

Part A — Gaussian Process Regression

1. **(10 pts)** Implement Gaussian Process (GP) regression with an RBF kernel from scratch. Show predictive mean and 95% credible intervals on noisy samples from

$$f(x) = sin(3x) + 0.5cos(5x), x \in [0,2\pi].$$

2. **(10 pts)** Implement marginal log-likelihood optimization to tune kernel hyperparameters. Report the learned values and plot GP predictions before and after optimization.

Part B — Expectation-Maximization: Gaussian Mixture Model (20 pts)

1. **(10 pts)** Implement the EM algorithm for a Gaussian Mixture Model (GMM) with full covariance matrices.
2. **(10 pts)** Apply your GMM to the `Iris` dataset (https://scikit-learn.org/1.5/auto_examples/datasets/plot_iris_dataset.html). Compare clustering results to k-means. Plot log-likelihood across iterations and analyze convergence.

Part C — EM for Bayesian Linear Regression

1. **(10 pts)** Implement EM updates to estimate the prior precision λ and noise precision β for Bayesian linear regression.
2. **(10 pts)** Apply your method to a regression dataset, for example generated with `sklearn.datasets.make_regression` with transformed features into polynomial terms (e.g., degree 3 or 4 using `PolynomialFeatures` from scikit-learn). Plot the evidence (log-marginal likelihood) across iterations and discuss how λ, β evolve.

Part D — Variational Inference

1. **(20 pts)** Implement Mean-Field VI with the Reparameterization Trick
   - Assume a mean-field Gaussian approximation:

     $$q(w) = \mathcal{N}(w \mid \mu, diag(\sigma^2))$$

   - Use the **reparameterization trick**:

     $$w = \mu + \sigma \odot \epsilon, \epsilon \sim N(0, I)$$

   - Optimize the Evidence Lower Bound (ELBO) using gradient ascent with PyTorch.

2. **(20 pts)** Run experiments
   - Dataset: use `sklearn.datasets.make_classification` on a binary subset of the digits dataset.
   - Plot the ELBO as a function of iteration.
   - Evaluate predictive accuracy and calibration (Brier score) on a held-out test set.
   - Compare to standard MAP logistic regression (e.g., scikit-learn's `LogisticRegression`).
