AMATH 445/645 Assignment 2 Winter 2026

Due date: Friday, February 13 at 11:59 pm

Note: Undergraduate students are not required to do Question 2, but they are encouraged to attempt it.

Question 1: Cell Viability Prediction via Neural Networks

In this question, we continue creating machine learning algorithms which can predict cell viability from bioprinting parameters, similar to [1]. However, this time we are using neural networks.

You will train in PyTorch:

- a **classification** neural network to predict acceptable/unacceptable cell viability,
- a **regression** neural network to predict real-valued cell viability.

Preprocess the dataset according to the provided code in this thesis.
For classification, *do not* use the "Acceptable Viability?" column. Instead, label a sample as acceptable if viability $\geq 70\%$, and unacceptable otherwise. Split both datasets 85/15 into training and test sets.

Validation protocol. On the training set, perform **5-fold cross validation**. Report the mean performance across folds and the performance when training on the full training set.
Metrics.

- Regression: MSE
- Classification: Accuracy, Precision, Recall

Important: Only evaluate the default model and your final best model on the test set.

1. **Default models.** Train regression and classification networks in PyTorch using the default optimizer settings from Tutorial 2. Report training (5-fold CV + full training) and test performance.
2. **Activation functions.** Test ReLU and tanh, plus one mixed-activation configuration. State which performs best.
3. **Learning rate and momentum.** Try two learning rates from different orders of magnitude and train with/without momentum. Plot and discuss the effect on the 5 training loss curves.
4. **Best model vs default.** Combine your best hyperparameters, train on the full training set, and evaluate on the test set. Compare to the default model and discuss.

Question 2: PINNs for Damped Oscillators

The damped oscillator model is

$$x''(t) + 2\gamma\, x'(t) + \omega^2 x(t) = 0, \quad t \in [0, 10].$$

1. **Data generation.** Solve with $\gamma = 0.1$, $\omega = 1$, $x(0) = 1$, $x'(0) = 0$ (you may use `solve_ivp`).
2. **PINN setup.** Build a continuous PINN. Treat γ and ω as trainable. Use a loss combining:
   - data MSE,
   - physics residual MSE on collocation points,
   - initial-condition penalty.

State how many data and collocation points you use.

3. **Training.** Train the PINN. Plot total loss and the convergence of γ, ω.
4. **Evaluation.** Compare predicted $x(t)$ and inferred parameters with ground truth; discuss discrepancies.
5. **Visualization.** Plot ground truth vs PINN prediction for $x(t)$.

Question 3: Backpropagation in Feedforward Neural Networks

For a single training example $\mathbf{x}$, define

$$\delta_j^{(\ell)} = \frac{\partial c}{\partial z_j^{(\ell)}}.$$

1. Show that

$$\delta_j^{(L)} = \frac{\partial c}{\partial a_j^{(L)}} g_L'(z_j^{(L)}), \qquad \delta_j^{(\ell)} = \Big(\sum_k \delta_k^{(\ell+1)} W_{kj}^{(\ell+1)}\Big) g_\ell'(z_j^{(\ell)}).$$

2. Show that

$$\frac{\partial c}{\partial W_{ij}^{(\ell)}} = a_i^{(\ell-1)}\, \delta_j^{(\ell)}, \qquad \frac{\partial c}{\partial b_j^{(\ell)}} = \delta_j^{(\ell)}.$$

Question 4: Theoretical Properties of k-Means

Given data $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{R}^d$, the k-means objective is

$$J = \sum_{r=1}^{k} \sum_{\mathbf{x}_i \in \mathcal{C}_r} \|\mathbf{x}_i - \boldsymbol{\mu}_r\|^2.$$

1. Show that for fixed assignments, the optimal centroid is

$$\boldsymbol{\mu}_r = \frac{1}{|\mathcal{C}_r|} \sum_{\mathbf{x}_i \in \mathcal{C}_r} \mathbf{x}_i.$$

2. Explain why k-means can converge to a local (not global) minimum and why initialization matters.

References

[1] Shuyu Tian et al. "Machine assisted experimentation of extrusion-based bioprinting systems". In: *Micromachines* 12.7 (2021), p. 780.