CSE655 – Deep Learning (Spring 2025)

Homework #4

Hand-in Policy: Via Teams. Late submission rules are given in the course syllabus.
Collaboration Policy: You are expected to do your own work. No collaboration is permitted.
Grading: This homework will be graded on the scale 100.

Description: KANs vs MLPs.

Part 1: Theoretical Foundations (Hand Written – 30 points)

1. **Kolmogorov–Arnold Representation Theorem (10 points)**
    - State the theorem and explain its significance in the context of function approximation.
    - Describe how it differs from the universal approximation theorem typically associated with neural networks.
2. **KAN Architecture (10 points)**
    - Explain how KANs use learnable univariate functions instead of weights.
    - Discuss how compositions of these univariate functions can approximate multivariate functions.
3. **Comparison with DNNs (10 points)**
    - Discuss potential advantages and limitations of KANs in terms of model interpretability, computational cost, generalization, suitability for high-dimensional inputs

---

Part 2: Implementation (Coding – 40 points)

1. **KAN Implementation (20 points)**
    - Implement a simplified version of a KAN using any deep learning using PyTorch.
    - Use piecewise linear functions to model univariate transformations.
2. **Training on Toy Dataset (10 points)**
    - Train both your KAN and a standard fully connected neural network to regress the function $f(x, y) = sin(x) + cos(y) + x^2 - y^2$.
3. **Analysis (10 points)**
    - Plot the training and validation loss curves.
    - Compare training time, number of parameters, and accuracy of the two models.

Part 3: Critical Discussion (Essay – 30 points)

Based on your experiments, what are the key takeaways regarding when KANs might be preferable over DNNs?

What to hand in:

- Hand-written PDF report with answers to Part 1 and Part 3
- Jupyter notebook or Python script for Part 2
- All plots and figures embedded in the report