Homework 1
CSCE 790 - Section 007: Neural Networks and Their Applications
Fall 2024
Due Date- September 28, End of day (Blackboard)

Instructions:

- Submit a single PDF with solutions. Use the following format to name the PDF file - "CSCE790-HW-1-Lastname."
- The solutions should be very clear and should follow all the instructions below.
- If the solutions are not readable, they will not be graded
- If you refer to any resource to get your solutions, add an acknowledgement and list the references (details of the source, e.g., book, website, etc.).
- Include codes to the problems (as a link to GitHub or Notebook) and codes should be clearly commented
- Do not attach colab printed page or .ipynb file in your solution as an additional document. Add their links in the PDF and the codes should be accessible with the link without needing additional permissions
- Add good figures; the figures should be generated with pdf, eps, or svg file extension and added to your solution, and they should have detailed caption and label

Problems:

(20 Points) Read the article in the link below (a) and/or a paper from the reference (b). Use the codes in the article and re-run the code to generate the results. Submit the following (c-e).

(a) Article: Classification of handwritten digits

(b) Reference: Compilation of prominent results with MNIST

(c) Generate any figure that you think would convey if your neural network model is trained properly (submit captioned image with good resolution).

(d) Copy of the code with your understanding of how the code works (a pseudo-code with comments would do).

(e) With the same neural network architecture with the same number of hidden layers, if you increase the number of neurons in the hidden layers, does the network performance improve, stay the same, or degrade? Generate a figure (or figures) that conveys the reason for your conclusion.

(20 Points) (a) A perceptron with two inputs and one output is given by

$$y = \sigma(-4.79x_1 + 5.90x_2 - 0.93). \tag{1}$$

For the inputs defined in the domain $[-2, 2] \times [-2, 2]$, plot the output surface y as a function of x_1, x_2, when the activation function is (a) sigmoid (b) Hard limit and (c) Radial basis function. Plot the function with with 100, 5000, 10000, sample points from the domain.

(b) A two-layer NN with two inputs and one output is given by

$$y = W^T \sigma(V^T x + b_v) + b_w, \tag{2}$$

with weight matrices and thresholds/biases given by $V^T = \begin{bmatrix} -2.69 & -2.80 \\ -3.39 & -4.56 \end{bmatrix}$, $b_v = \begin{bmatrix} -2.21 \\ 4.76 \end{bmatrix}$, $W = \begin{bmatrix} -4.91 \\ 4.95 \end{bmatrix}$, and $b_w = -2.28$. For the inputs defined in the domain $[-2, 2] \times [-2, 2]$, plot the output surface y as a function of x_1, x_2, when the activation function is (a) sigmoid (b) Hard limit and (c) Radial basis function. Plot the function with with 100, 5000, 10000, sample points from the domain.

(10 Points) Read the article in the link given. Use the codes in the article to reproduce the results given in it. Submit the results, your code, and a detailed step-by-step description of the algorithm. Link

(10 Points) Read the article in the link given. Try to summarize the different learning rules described in the article using your understanding of how they work. Link

(10 Points) Read the article in the link given. Use the codes in the article to reproduce the results. Submit the results, your code, and a detailed step-by-step description of the algorithm. Link

(20 Points) Repeat Examples A and B from this paper: Link. Use first order Euler approximation to simulate (11)-(13) and (17) and generate training and test data. Reproduce all the relevant results, and provide a short summary of the examples and your implementation.

(10 Points) Are these functions linear or nonlinear? Add a formal justification to your answer. (a) $f(x) = |x|$, (b) $f(x) = x^2 + 2x + 2$, and (c) $f(x) = \frac{1}{x}$.