CS480/580 Introduction to Artificial Intelligence

## Assignment 5

Due date: 12/09/2022

Total Points: 100

1. Pattern recognition

In this assignment, you are asked to design a neural network from scratch to recognize handwriting digits (0-9). For example, the following sample represents a handwriting 9.

00000000000000000000000000000000

00000000000001111100000000000000

00000000000001111100000000000000

00000000001111111110000000000000

00000000011111111111111000000000

00000000011111111111111100000000

00000000111111111111111100000000

00000000111111111111111110000000

00000001111111100001111110000000

00000001111111000001111111000000

00000001111110000001111111000000

00000001111110000000111111000000

00000011111100000001111110000000

00000001111111000001111110000000

00000001111111110001111110000000

00000001111111111111111110000000

00000000001111111111111110000000

00000000001111111111111110000000

00000000000111111111111110000000

00000000000000111001111110000000

00000000000000000000111111000000

00000000000000000000111100000000

00000000000000000000111100000000

00000000000000000000111100000000

00000000000000000001111110000000

00000000000000000001111110000000

00000000001111100001111100000000

00000000001111111111111100000000

00000000001111111111111100000000

00000000011111111111111000000000

00000000000011111111111000000000

00000000000000001111110000000000

The data file can be downloaded from

<https://archive.ics.uci.edu/ml/machine-learning-databases/optdigits/optdigits-orig.windep.Z>

The data file contains 1797 instances from 43 writers. The data is prepared by NIST to extract normalized bitmaps of handwritten digits from a preprinted form.

Task 1: Encoding (50 pts)

Unzip the data file from the link. Encode the data into input and target files for neural network training.

Task 2: Neural Network Training (50 pts)

* Option 1: Build your own Neural Network from scratch. (+10 pts)
* Option 2: Use Matlab Neural Network Toolbox.

Train the neural network (pattern net) based on your input/output files. Report your testing, training, and validation accuracy and provide analysis.

Bonus: (20 pts for CS480 students and 10 pts for CS580 students)  
Using 1024 features for neural network is very cumbersome. Is there a way to reduce the number of features? If you have an idea, show me how effective it is.

Please submit your well-documented programs, readme file, and analysis report via blackboard before the assignment due date.