THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY
Department of Computer Science and Engineering
COMP4211: Machine Learning
Fall 2020 Assignment 1
Due time and date: 11:59pm, Oct 11 (Sun), 2020.

IMPORTANT NOTES

- **Your grade will be based on the correctness and clarity.**

- **Late submission: 25 marks will be deducted for every 24 hours after the deadline.**

- **ZERO-Tolerance on Plagiarism: All involved parties will get zero mark.**

Q1. Given boolean inputs A,B,C (which take either 0 or 1), show that the function $A \vee B \vee C$ can be represented by a single perceptron.

Q2. Given boolean inputs A,B,C (which take either 0 or 1), show that the function $\neg A \wedge (B \vee C)$ can be represented by a set of perceptrons.

Q3. The following shows the computation graph of the function $f(w, x) = 1/(1 + e^{-(w_0 x_0 + w_1 x_1 + w_2)})$. Show in the figure the gradients of f with respect to x_0 and x_1 as in the lecture notes.
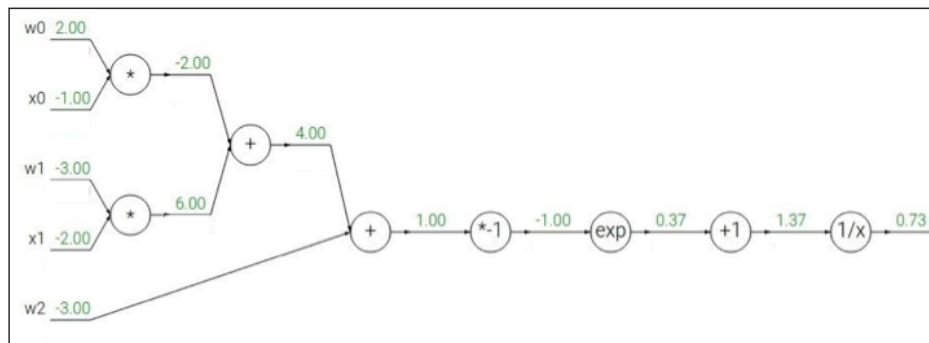
Q4. Use `sklearn` to implement the following steps. Write your answers in `A2.ipynb`, which should include codes, results and your conclusions. You can plot figures by `seaborn` or `matplotlib` in Python.

1. Load the MNIST dataset from http://yann.lecun.com/exdb/mnist/. Randomly use 80% of the data for training, and the remaining 20% for testing.

2. Display (i) the first sample; (ii) number of training samples; and (iii) number of testing samples.

3. Design an appropriate MLP such that the training accuracy is at least 95%. You can use the MLPClassifier in https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html. Show the (i) training accuracy; and (ii) testing accuracy that you obtain.

4. Double the number of hidden units in all hidden layers, show the (i) training accuracy; and (ii) testing accuracy of this model. Similarly, try reducing the number of hidden units in all hidden layers, and again show the (i) training accuracy; and (ii) testing accuracy of this model. What observation do you have?

Submission Guidelines

Put your answers to **Q1**, **Q2** and **Q3** in the file A1.pdf, put answers of **Q4** in A2.ipynb. Your answers in the A2.ipynb file should include source codes, results, and your conclusions. Submit the files via canvas.