THE UNIVERSITY OF HONG KONG
DEPARTMENT OF PHYSICS

Assignment 5

Course: *Machine Learning in Physics (PHYS3151)* – Prof. *Zi Yang Meng*
Tutor: *Mr. Tim Lok Chau*
Due date: ***29th April, 2025***

This assignment is a project. You have to deliver a presentation on how you solve the questions and your results. Efforts beyond the answer to these questions are welcome.

2D Ising model phases classification

In this assignment, we will use a neural network to classify phases (ordered or disordered) for a 12×24 2D honeycomb lattice Ising model.

In all questions, you need to use data files Ising_config_L12.csv and T_data_L12.csv, which you can download from the Moodle page. The Ising_config_L12.csv contains 6400 configurations, obtained by Monte Carlo simulation, and the 288 columns are spin orientations on every lattice site. T_data_L12.csv contains the corresponding temperature of each configuration, going from 0.5 to 2.5 with a step of 0.1.

For the 2D honeycomb lattice Ising model, the critical temperature $T_c \approx 1.51$, below which the system is in an ordered state, and disordered above, which is demonstrated in assignment 4.

1. **Data Processing**

   (a) Load the data files, use the configurations as features $\mathbf{X}$. Perform standardisation on $\mathbf{X}$.

   (b) Construct the expected output list $\mathbf{Y}$. Label the sample with $T > T_c$ as 1 and 0 otherwise, i.e. it tells if the system is ordered or disordered.

2. **Model Construction** Construct two Neural Networks with the following parameters.

   (a) 288 input neurons, 1 hidden layer with 3 neurons, and 1 output neuron.

   (b) 288 input neurons, 2 hidden layers with 3 neurons on each layer, and 1 output neuron.

   In both cases, the output neuron is a number from 0 to 1. If its value is above 0.5, the model predicts the input configuration to be in the ordered state (y=1).

3. **Model Training** Randomly divide the 6400 configurations into training set with **5000 data** and a testing set with **1400 data**. The testing set should only be used to evaluate the models, but not for updating the weights. Train the models using back propagation discussed in the lectures. Store lists for the following variables obtained during training the neural network, and plot the following against the number of iterations,

(a) Cost J calculated from the training set.

(b) Accuracy calculated from the training set.

(c) Accuracy calculated from the testing set.

4. **Model Evaluation** For each fully trained network, perform forward propagation using all 6400 configurations, log the value of the output neuron.

   (a) Make a scatter plot of the final output value against temperature.

   (b) At each temperature, find the mean value of all points with the same temperature. Plot the mean value on the scatter plot.

   Sample plots for the two models are as follows.

5. **Discussion** With the above variables and your own experience, compare the performance of the 3 models.