**Testing(120 mins)**

1. Describe Gradient Descent algorithm. How does the algorithm work? Why does the Artificial Neural Network(**ANN**) use the Gradient Descent algorithm?

2. What is the difference between algorithm Stochastic Gradient Descent(**SGD**), Mini-batch Gradient Descent and Batch Gradient Descent?

3. What is the learning rate? Why does ANN must tune this parameter?

4. List 16 most use the hyperparameters in Deep Learning that you know. Explain the meaning of each hyperparameter?

5. What is the purpose of activation functions in Deep Learning? List 3 of them and explain how one of them works?

6. Comparison between Machine Learning and Deep Learning?

7. Describe forward propagation and backward propagation? Why must it use backward propagation? (hint: chain rules)

8. How many parameters of this ANN?

Input vector 100 -> Neural Network 512 node with bias activation relu -> Neural Network 512 node without bias activation softmax.

9. In the project for Agriculture, we need to give advice on watering level on that day for all the plants. Our input is weather forecast on that day, weather forecast on the next two day, soil moisture at current time, plant growth time. Our data has one thousand five hundred days of that plant. Our desired output is how much water we need watering on that day.

* Use linear regression with ANN to solve the problem above.
* How many layers did you use?
* Describe how many nodes did you use in that mode?
* Which activation did you use? Why?
* Which loss function did you use? Why?