# XJTLU Entrepreneur College (Taicang) Cover Sheet

| | |
|---|---|
| Module code and Title | Introduction to Neural Networks |
| School Title | School of Artificial Intelligence and Advanced Computing |
| Assignment Title | Coursework |
| Submission Deadline | 5 pm China time (UTC+8 Beijing) on Monday 11th April 2022 |
| Final Word Count | NA |
| If you agree to let the university use your work anonymously for teaching and learning purposes, please type **"yes"** here. | |

I certify that I have read and understood the University's Policy for dealing with Plagiarism, Collusion and the Fabrication of Data (available on Learning Mall Online). With reference to this policy I certify that:

- My work does not contain any instances of plagiarism and/or collusion.
  My work does not contain any fabricated data.

**By uploading my assignment onto Learning Mall Online, I formally declare that all of the above information is true to the best of my knowledge and belief.**

| **Scoring – For Tutor Use** | |
|---|---|
| **Student ID** | |

| **Stage of Marking** | **Marker Code** | **Learning Outcomes Achieved (F/P/M/D) (please modify as appropriate)** | | | **Final Score** |
|---|---|---|---|---|---|
| | | **A** | **B** | **C** | |
| 1st Marker – red pen | | | | | |
| Moderation<br>– green pen | **IM Initials** | The original mark has been accepted by the moderator (please circle as appropriate): | | | Y / N |
| | | Data entry and score calculation have been checked by another tutor (please circle): | | | Y |
| 2nd Marker if needed – green pen | | | | | |

| **For Academic Office Use** | | | **Possible Academic Infringement (please tick as appropriate)** | |
|---|---|---|---|---|
| **Date Received** | **Days late** | **Late Penalty** | ☐ **Category A** | Total Academic Infringement Penalty (A,B, C, D, E, Please modify where necessary)<br>____________________ |
| | | | ☐ **Category B** | |
| | | | ☐ **Category C** | |
| | | | ☐ **Category D** | |
| | | | ☐ **Category E** | |

## Students

The assignment must be typed in an MS Word or a PDF document and submitted via Learning Mall Online to the correct drop box. Only electronic submission is accepted and no hard copy submission.

All students must download their file and check that it is viewable after submission. Documents may become corrupted during the uploading process (e.g. due to slow internet connections). However, students themselves are responsible for submitting a functional and correct file for assessments.

The submission needs to include code, test cases, results, and a brief analysis. If there are some errors in the program, debugging errors should be submitted.

# Q1. Short Answer Questions

1. Identify with a graph whether each of the following two input logic operations are linear separable and can therefore be represented by a simple perceptron. If not, can it be represented by a 3-layer network? Provide an appropriate weight vector with activation function that outputs the correct symbol 0 or 1. **[10 marks]**

   (a.) $x_1$ AND $x_2$ (3)

   (b.) NOT($x_1$ OR $x_2$) (3)

   (c.) ($x_1$ AND NOT $x_2$) OR ($x_2$ AND NOT $x_1$) (4)

2. Consider a 3-layer backpropagation network with $d$ input units, $n$ hidden units, $c$ output units, and bias. The network has activation function $f(.)$, with input to hidden weights labelled $\omega_{ji}$ and hidden to output weights labelled $\omega_{kj}$. The output activation of a hidden unit labelled $y_j$ and the output activation of an output unit labelled $z_k$. The network is going to be trained using a squared error $J$.

   (a) How many weights are in the network? **[2 marks]**

   (b) For a single input vector $\mathbf{x} = (x_1, ..., x_d)$ with associated target vector $\mathbf{t} = (t_1, ..., t_c)$, using the chain rule, calculate expressions for the following derivatives: **[13 marks]**

      i. $\partial J / \partial z_k$ (3)

      ii. $\partial J / \partial \omega_{kj}$ (3)

      iii. $\partial J / \partial y_j$ (3)

      iv. $\partial J / \partial \omega_{ji}$ (4)

   Hence derive formulas for the weight updates required to perform steepest descent in the squared error $J$.

# Q2. CNN-based Image Classification Programming

For this question, you will experiment with convolutional neural networks. **[45 marks]**

- **Language and Platform** Python (version 3.5 or above) and Tensorflow with Keras or Pytorch (newest version).You may use any libraries available on Python platform, such as numpy, scipy, matplotlib, etc.
- **Dataset** CIFAR10. Link: https://www.cs.toronto.edu/~kriz/cifar.html
- **CNN Model** AlexNet. ( Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. ”Imagenet classification with deep convolutional neural networks.” Advances in neural information processing systems 25 (2012).)
- **Code Submission** Your code submission should be a package file containing your source code (with the instruction on how to run them).

This question is split in a number of steps. Answer the questions for each step in the document and attach the corresponding code.

1. **Data preparation** Load the test set of CIFAR10 (10000 images), randomly split the set into training set (7000 images), validation set (2000 images) and test set (1000 images). Use python command to display the number of data entries, the number of classes, the number of data entries for each classes, the shape of the image size, etc. Report the results and attach the screenshot in the document. What else should be done to the data before feed into the model? Explain the process and add them to your program. (10)

2. **Model building** Write the AlexNet model. Print the model using python commands. Explain the model what the inputs and outputs are, how many covolutional layers and fully connection layers. What is the intuition for dropout used in convolutional neural networks? (5)

3. **Training** Train the model at least 50 epochs. Explain what the lost function is, which optimizer do you use, and other training parameters, e.g., learning rate, epoch number, batch size, etc. Plot the training history, e.g., produce two graphs (one for training and validation losses, one for training and validation accuracy) that each contains 2 curves.

   Re-train the model by decreasing the training set to 5000 images by random sampling, replacing with a new optimizer, and keep the training epoch unchanged. Explain where you have modified your code. Compare the results on validation set, explain why the new trained model perform better or worse than the previous one, and prove it using ablation experiment.

   Has your first model converged? If not, train a convergent model. (15)

4. **Test** Test your converged network on the test set. Show the accuracy and confusion matrix using python commands. (5)

5. **Model Modification** Add two layers, e.g., one convolutionl layer and one fully connection layer into the original AlexNet, explain where you add them and the corresponding parameters. Re-train and re-test the modified model, show the related intermediate (i.e. graphs of loss and accuracy in the training process) and final results (i.e. accuracy and confusion matrix). (10)

## Q3. Real-world Application Questions

Give **ONE** specific real-world problem that can be solved by neural networks. Answer the questions below (answer to each question should not exceed 150 words). **[30 marks]**

1. Detail the issues raised by this real-world problem, and explain how neural networks maybe used to address these issues. (5)
2. Design an algorithm using a exiting published neural network to solve this problem. Which network do you choose? Explain. (5)
3. Draw a diagram of the workflow for the algorithm you proposed, and describe in details. (5)
4. Do you need labeled data to train the network? If your answer is yes, specify what kind of label you need. If your answer is no, indicate how you train the network with unlabeled data. (5)
5. How to collect your training data? What pre-processing works should be done on the raw data before training? (5)
6. If you find the performance of your proposed algorithm is unsatisfied. What would be the reason for that? How can you solve it? (5)

**The End**

# Marking Criteria

(1). The marks for each step in Q2 are divided into two parts

| Rubrics | Marking Scheme | Marks |
|---|---|---|
| Program [60%] | The program runs correctly, the solution is well thought out. The code is well commented, and has a clear layout. | 60% |
| | The program runs correctly, program outputs are mostly accurate and informative. The code has clear layout. | 40% |
| | Some of the component parts of the problem can be seen in the solution, but the program wont produce any outcome. The code is difficult to read in places. | 20% |
| | The component parts of the program are incorrect or incomplete, providing a program of limited functionality that meets some of the given requirements. The code is difficult to read. | 0% |
| Question Answer [40%] | All question are answered correctly, plentiful evidence of clear understanding of the CNN | 40% |
| | Some of the answers not correct, convincing evidence of understanding of the CNN | 20% |
| | Answers are incorrect, very little evidence of understanding of the CNN | 0% |

(2). Marking scheme for each sub-question in Q3

| **Marks** | **Scope, quantity and relevance of studied material** | **Evidence of understanding (through critical analysis)** |
|---|---|---|
| **5** | High quality of originality. Extensive and relevant literature has been creatively chosen, and outlined and located in an appropriate context. | There is plentiful evidence of clear understanding of the topic. |
| **4** | Shows originality. The major key points and literature have been outlined and put in an adequate context. The major points of those sources are reasonably brought out and related in a way which reveals some grasp of the topic in question. | There is convincing evidence of understanding of the topic. |
| **3** | Effort has gone into developing a set of original ideas. Some relevant key points and literature are outlined, but this outline is patchy, unclear and/or not located in an adequate context. | There is some evidence of understanding of the topic. |
| **2** | May demonstrate an incomplete grasp of the task and will show only intermittent signs of originality. There are some mention of relevant key points, but this outline is very patchy, unclear, and/or very inadequately placed in context. | There is limited evidence of understanding of the topic. |
| **1** | Shows very limited ability to recognise the issues represented by the brief. There is little mention of relevant key points. | There is very little evidence of understanding of the topic. |