CS3950 Assessed Coursework

This assignment must be submitted by **March 28 2025, 2pm**. Feedback will be provided by 01 May 2025. Extensions can only be granted by the office or academic advisor (not the lecturer).

Learning outcomes assessed

- Be able to perform backpropagation on simple neural network graphs.
- Understand gradient-based optimization algorithms.
- Understand strategies for key learning parameter settings.
- Understand differences between several common activation functions.
- Understand recurrent neural networks and LSTMs.
- Apply deep learning algorithms to real-world data, fine-tune hyper-parameters and evaluate model performance.

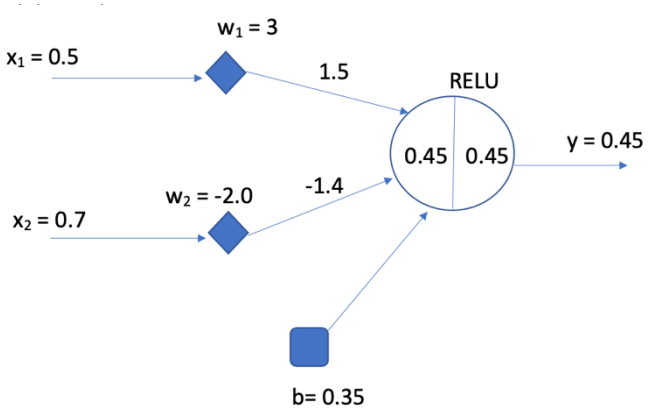
Instructions

The coursework assignment must be completed strictly individually. The answers do not require writing or executing code and should be submitted electronically in a pdf file via Moodle, on the corresponding assignment page. Please include your student id in the file name.

Marking scheme

There are seven questions. For questions 1-5, each worths 10 marks. Question 6 has 50 marks. Partial marks will be awarded for partially correct answers. The total mark is between 0 and 100.

1) Consider the ReLU neuron on the right, for which we already calculated the forward pass. [10 marks]
   a) What are the values of $\partial y/\partial x_2$ and $\partial y/\partial w_2$? [3 marks]
   b) What is the effect of doubling w_1 on $\partial y/\partial x_1$? [1 marks]
   c) What would be the values of y, $\partial y/\partial w_1$, and $\partial y/\partial w_2$ if the neuron's activation function were sigmoid instead of ReLU? [6 marks]
      - Tip: the sigmoid is defined as $g(z) = 1/(1+e^{-z})$ and its derivative is $\partial g(z)/\partial z = g(z)(1-g(z))$



2) Describe the RMSProp and ADAM optimisation algorithms. What is the main difference between them? [10 marks]

3) In a neural network, what is a ReLU neuron and what is a Hyperbolic Tangent neuron? [5 marks] For each type of neuron, what is the gradient of the output with respect to the total input? [5 marks]

4) What is the dropout regularisation technique? What is L_2 regularisation? If L_2 regularisation is applied with parameter 0.001, what is the adjustment to the gradients of the weights during training? [10 marks].

5) Explain carefully, with the aid of diagrams and equations, a recurrent net, a Long-Short-Term Memory (LSTM) net and discuss their differences. Why does a LSTM network need some sigmoid units? Indicate on your diagram how new memories can be 'stored', and how. [10 marks]

6) Implement a NN classifier for image classification (50 marks). The detailed requirements are as follows.

   You will need to implement a NN classifier, and in selecting appropriate levels of regularisation. Completing this project requires mainly code that has already been given in worksheets.

   Instructions:

   - Instructions are given in the Jupyter notebook. The project work is an extension of the work on the exercise sheets.
   - You may use any code that has been provided; if you use code from online tutorials, please reference the tutorial that you use.
   - Marks will be given for the procedure of fitting a model, tuning the regularisation, investigating how the accuracy and loss of the model relate to training set size, and for clearly displaying the results.

Notebook submission:

- Please make sure that you keep the results in all the cells. That is, have the results of your calculations and your graphs shown in the notebook. (So we don't have to re-run your code.)
- For this implementation component for Q6, please submit the Python notebook.
- **Please generate a compressed file** (e.g. a zip file) by including **(1) the notebook for Q6, and (2) a PDF file containing the answers for Q1-Q5**.