# ΥΣ19 Artificial Intelligence II (Deep Learning for Natural Language Processing)
Fall Semester 2021
Homework 2
20% of the course mark
Announced: November 25, 2021 Due: December 19, 2021 at 23:59

1. Compute the gradient of the cross entropy loss function with respect to its logits. See case (7), page 4 of the matrix calculus notes of the Stanford class CS222n. Remember that we have defined cross entropy loss in our lecture on softmax regression. Note that the notes from Stanford use an alternative notation which you are expected to follow.

   (1/10 marks)

2. Apply backpropagation to the following computational graph to compute all the gradients with respect to the scalar inputs x, m, b and y^*. Assume the loss function is $MSE(y, y^*) = (y - y^*)^2$ where y^* is the true label and y is the predicted output from the neural network. Note also that $ReLU(x) = \max(0, x)$.
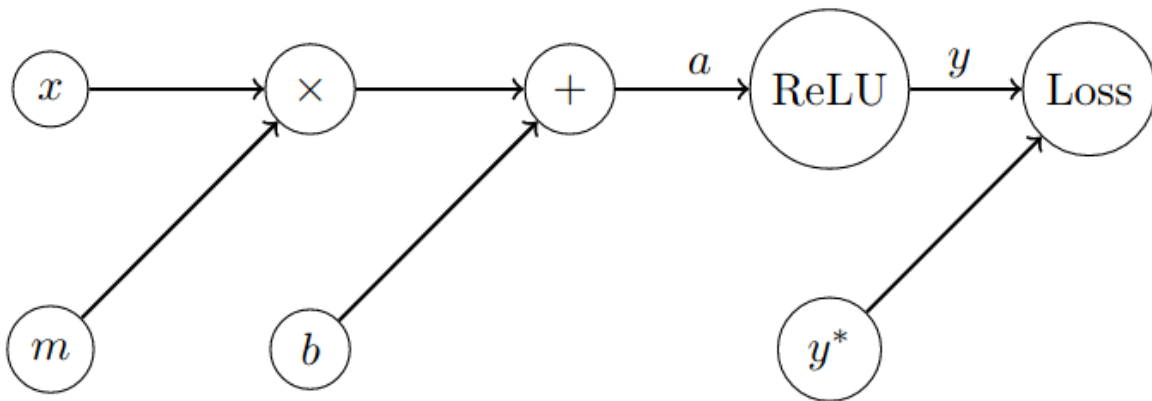
   (1/10 marks)

3. Develop a vaccine sentiment classifier with 3 classes (neutral, anti-vax and pro-vax) using feed-forward neural networks and the datasets of the previous homework provided on e-class. We expect you to experiment with at least two models and compare them using F1 score, recall and precision. For the development of the models, you are expected to experiment at least with the following hyperparameters:

- the number of hidden layers and the number of their units
- the activation functions
- the loss function
- the optimizer

Furthermore, in at least one of your models, you must use pre-trained word embedding vectors using GloVe.

For your best model, you must also provide the loss vs epochs plot along with the ROC curve (on the validation set, after training has been completed) and explain what you see. Also, compare your best model with the model you obtained using softmax regression in Homework 1. Your solution should be implemented in PyTorch and we expect your report to be well-documented.

(8/10 marks)

Note: You should hand in: (i) a pdf with your answers to questions 1 and 2 and a detailed report of your solution for question 3, including an explanation of the methods you used, an evaluation and comparison of different approaches you tried and citations to relevant literature that you might have used in developing your solutions. (ii) one Colab notebook (ipynb files using https://colab.research.google.com/) containing your code for question 3. You should use Python 3.6 or a later version.