HW 3 – Neural Networks

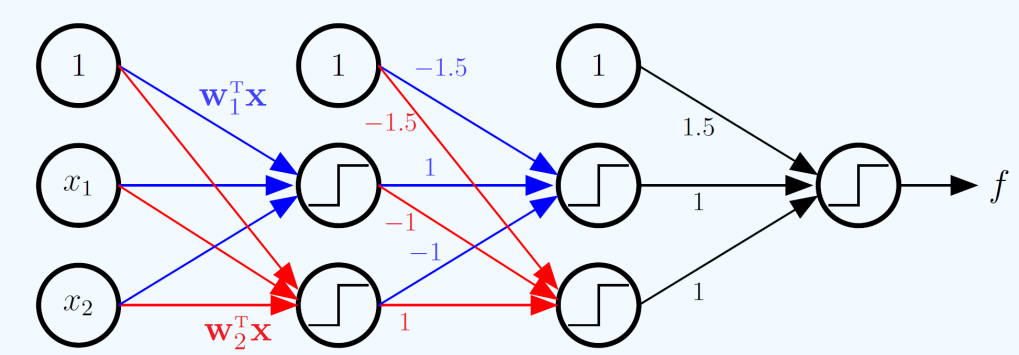
Understanding of neural networks and implementing one from scratch

First Submission Due: Monday, November 20th, 11:59pm Pacific Time

Revision Due Dates will be updated after the grades are released

Task 1

[LP] Modified version of exercise 7.3 in textbook – use the below graph representation to compute an explicit formula for f in terms of $\vec{x} = [1 \ x_1 \ x_2]^T$.



Task 2

[HP 1]

Part 1: Write a function to perform forward propagation and back propagation to compute the gradient on a neural network with 2 layers. Your neural network should have 2 input units ($d^{(0)} = 2$), m hidden units ($d^{(1)} = m$), 1 output unit ($d^{(2)} = 1$). Use tanh as the activation function on the neurons. For gradient, use squared error as the error measure.

If you would like to make the code generic, think about how you need to specify the neural network with the layers and the number of neurons per layer. E.g., the functions for forward and back propagation may take an array of $W^{(l)}$ matrices and the input vector $x^{(0)}$ as the parameters.

Note: There is a lot of code available online and implementations that you can easily copy from. I highly discourage you to do so for two main reasons: (i) We will use software similarity checkers and you risk getting caught for plagiarism. Please use your creativity to solve the homework and not to get around the system.; (ii) The course is meant for you to learn about machine learning and get a peek behind the curtain on what exactly happens in the libraries you might end up using in practice. Although you may never need to write anything from scratch, these exercises are meant to solidify your fundamental understanding of the topic.

Part 2: Implement gradient descent to train the neural network you implemented. Make your code capable of accepting any number of samples, so that you can easily perform stochastic gradient descent using this function.

Part 3: Use your neural network and gradient descent implementation from above to classify digits data (1 vs 5) from HWs 1 and 2. Perform stochastic gradient descent using only one sample in the training dataset to compute the gradient descent during each epoch. You may choose to randomly shuffle your dataset or pick a random sample in each epoch.

Part 4: Plot $E_{in}(w)$ verses iteration number for each experiment, stopping after $2*10^6$ iterations.

Part 5: Report the test error. Use the sign function in the output layer to report Ein and the test error for classification.

[HP] Perform mini-batch gradient descent (also called stochastic gradient descent) using 32 of the samples in the training dataset to compute the gradient descent during each epoch and repeat parts 4 and 5 from above.

Submission instructions

As mentioned in class activity 1, we will use GitHub to share code and track progress in this class. If you have not already done so, create a GitHub private repository for the course and name it "CMPE257-Fall23-FirstName-LastName". Add the following users to the repository: mahima-as and rbpravin. You are also welcome to add the instructor and ISA to your Colab notebooks.

In your GitHub repo, create a branch called *homework-3*. Frequently commit your code and make sure it is shared with the instructor and ISA. Include a link to the GitHub repository in your submission pdf.

Specifications

Task 1 only has one task labeled [LP] and task 2 has two components labeled [HP]. If you complete ALL the LP components satisfactorily, you will receive a grade of "low pass" on the homework. If you complete ALL the LP components and at least 1/2 HP components satisfactorily, you will receive a grade of "high pass". If you do not meet the criteria for a "low pass", the submission will be marked as "revision needed".

Note the following statements from the syllabus:

If a student receives a "low pass" or "revision needed" grade, the student may revise and resubmit their homework assignment by using one "token".

For homework assignments, if the student fails to submit their assignment by the posted deadline, their submission will receive a grade of "revision needed". If they fail to submit the assignment by the revision deadline, the submission will receive a grade of "fail".

At most two tokens may be used for the one-day deadline extensions (one token for each one-day extension), including the revision deadlines. Tokens will be automatically removed from your wallet if you submit late and/or resubmit.

VERY IMPORTANT: Include ALL the references you used for this assignment, including names of classmates you discuss with. Failure to cite your sources counts as an act of academic dishonesty and will

be taken seriously without zero tolerance. You will automatically receive a “fail” grade in the homework and further serious penalties may be imposed.

NOTE: You can look for help on the Internet but refrain from referencing too much. Please cite all your sources in your submission.

When you submit your assignment, you automatically agree to the following statement. If you do not agree, it is your responsibility to provide the reason.

“*I affirm that I have neither given nor received unauthorized help in completing this homework. I am not aware of others receiving such help. I have cited all the sources in the solution file.*”