

## **Assignment #5 Neural Networks**

In this assignment, you are to implement the Backpropagation algorithm. In order to get full credit, you are expected to cover the first 4 requirements of the following. The last component of the question is a bonus requirement:

- 1- Correctly implement the Backpropagation algorithm. We will assume discrete classification. As a result, you should make sure that your output layer has one node per class and the predicted label is that of the node with the highest activation. You may also need to reformat the input files to reflect this design decision. Your implementation must support the following:
  - Ability to create arbitrary network structure (number of hidden layers, number of nodes per layer, etc.).
  - Random weight initialization (small random weights with mean 0).
  - Incremental weight update.
  - A reasonable stopping criterion.
  - Training set randomization at each epoch.
  - An option to include a momentum term in training.
- 2- Use your backpropagation algorithm on the Iris problem, with a random 70/30 split.
  - With a single hidden layer and a fixed number of hidden nodes (of your choice), experiment with different learning rates. Graph training and test set accuracy over time for several different learning rates. Based on these results, select a reasonable learning rate.
  - With the learning rate selected and a single hidden layer, experiment with different numbers of hidden nodes, starting from 1 and adding 1 each time until you see no improvement on the training set's accuracy. For each choice of number of hidden nodes, graph training and test set accuracy over time.
  - Record your best number of hidden nodes (i.e., the one resulting in highest accuracy on the test set).
- 3- Use your backpropagation algorithm on the Vowel problem, with a random 75/25 split. (Note: Make sure you ignore the "Train or Test" attribute).
  - Repeat the above experiments.
  - With the learning rate selected, induce a 2-hidden layer neural network with 6 hidden nodes in the first layer and 4 hidden nodes in the second. Graph training and test set accuracy over time.
- 4- Using only the best number of hidden nodes as recorded above for 1 hidden layer and the same training/test splits, re-run your backpropagation algorithm with the momentum term option to induce a neural network for both Iris and Vowel. Graph training and test set accuracy over time.
- 5- Analyze the data you have collected and thoughtfully answer the following questions:
  - Discuss the effect of different learning rates on the algorithm's performance.
  - Discuss the effect of different numbers of hidden units on the algorithm's performance (1-hidden layer case).
  - Compare your recorded best numbers of hidden nodes for each problem with the following heuristic value:  $H=N/(10(I+O))$ , where N is the size of the (training) data set, I is the number of network inputs and O is the number of outputs.
  - How did the momentum term affect the learner's behavior (number of epochs to convergence, final accuracy, etc.)?

- 1- This is an individual task.
- 2- The source code as well as the report describing your functions and output should be submitted through Google Classroom.
- 3- The due date for the submission of this phase is Wednesday, December, 23, 2020 at 12:00 am.
- 4- Please Review the definition of cheating in the first presentation.

Best Regards,

**Inas Yassine**