Project 4b: Neural Nets

Introduction

In this project you will be implementing neural nets, and in particular the most common algorithm for learning the correct weights for a neural net from examples. Code structure is provided for a Perceptron and a multi-layer NeuralNet class, and you are responsible for filling in some missing functions in each of these classes. This includes writing code for the feed-forward processing of input, as well as the backward propagation algorithm to update network weights.

Files you will edit

NeuralNet.py	Your entire Neural Net implementation will be within this file

Files you will not edit

NeuralNetUtil.py	Functions for converting the datasets into python data structures
Testing.py	Helper functions for learning a neural net from data
autograder.py	A custom autograder to check your code with

Evaluation: Your code will be graded for technical correctness and performance on the given datasets. Your write up will be graded based on your interpretation of these results.

Academic Dishonesty: We will be checking your code against other submissions in the class for cheating. If you copy someone else's code and submit it with minor changes, we will know. These cheat detectors are quite hard to fool, so please don't try. We trust you all to submit your own work only; *please* don't let us down. If you do, we will pursue the strongest consequences available to us.

Getting Help: You are not alone! If you find yourself stuck on something, contact the course staff for help either during Office Hours or over email/piazza. We want these projects to be rewarding and instructional, not frustrating and demoralizing. But, we don't know when or how to help unless you ask.

Part 1: Neural Network Learning Implementation (60%)

This project follows the same terminology as in the lectures and Ch 18.7 in your book. Neural networks are composed of nodes called perceptrons, as well as input units. Every perceptron has inputs with associated weights, and from this it produces an output based on its activation function. Thus you will be implementing a feed-forward multi-layer neural net.

As in the last project, we will be training the neural nets to be classifiers. Inputs will be in the form of sets of examples that have an assignment of values to various features and corresponding class values. The datasets used for this project include the cars dataset from the last project and a new dataset of pen handwriting values. For the latter, numeric data from images is stored to train a classifier of handwritten digits.

Instead of converting the stored examples into dictionaries as in the last project, each example will be parsed into lists of numeric values. Each possible classification for each class corresponds to a single output perceptron, so in addition to the list of inputs each example includes the list of outputs for the output layer. The Pen dataset has 16 inputs and 10 output perceptrons, since there are 16 different features for the handwriting recognition data input and 10 possible classifications of the input (corresponding to the values 0-9). In the case of a discrete-valued examples such as in the cars dataset, distinct arbitrary numeric values are assigned to every value of every feature.

The code we provide you has the methods for parsing the datasets into python data structures, and the beginning of the Perceptron and NeuralNet classes. A Perceptron merely stores an input size and weights for all the inputs, as well as methods for computing the output and error given an input. An object of the NeuralNet class stores lists of Perceptrons and has methods for computing the output of an entire network and updating the network via back propagation learning. The network consists of inputs (just a list of inputs that is a parameter to feed forward), an output layer, and 0 or more hidden layers. Although the structure and initialization is written, all the actual functionality will be implemented by you.

Question 1 (2 points): Feed Forward

Implement sigmoid and sigmoidActivation in the Perceptron class. Then, implement feedForward in the NeuralNet class. Be sure to heed the comments - in particular, don’t forget to add a 1 to the input list for the bias input. You now have a Neural Net classifier! However, the weights are still randomized so it is rather useless…

Question 2 (2 points): Weight Update

Implement sigmoidDeriv, sigmoidActivationDeriv, and updateWeights in Perceptron according to the equation in the book. Note that delta is an input to updateWeights, and will be the appropriate delta value regardless of whether the Perceptron is in the output or a hidden layer; its computation will be implemented later in backPropLearning.

Question 3 (4 points): Back Propagation Learning

Implement backPropLearning in NeuralNet using the methods you implemented in questions 1 and 2. Note that this is a single iteration of the back propagation learning, and the loop to perform the full learning algorithm will be implemented in the next question. You can largely follow the pseudocode in your book, though note that you should not be updating weights until you have computed the error delta values that use those delta values. You code does not have to exactly follow our suggestions in the comments, so long as it correctly implements back propagation.

Question 4 (4 points): Back Propagation Learning Loop

Lastly, implement buildNeuralNet to actually train a good neural network classifier. The stopping condition for training should be the average weight modification of all edges going below the passed in threshold, or the iteration going above the maximum number of iterations also passed in. See the comments in the code for more detail. You should now have a working neural net classifier! If your solutions are right, then calling testPenData in Testing.py should result in output similar (since we are starting from random weights, the numbers will not be exactly the same) to this:

Starting training at time 16:10:05.910131 with 16 inputs, 10 ouputs, hidden layers [24], size of training set 7494, and size of test set 3498
. . . . . ! on iteration 10; training error 0.015139 and weight change 0.000118
. . . . . ! on iteration 20; training error 0.011576 and weight change 0.000095
. . . . . ! on iteration 30; training error 0.010015 and weight change 0.000088
. . . . . ! on iteration 40; training error 0.009674 and weight change 0.000085
. . . . . ! on iteration 47; training error 0.008954 and weight change 0.000077
Finished after 47 iterations at time 16:13:15.706756 with training error 0.008954 and weight change 0.000077
Feed Forward Test correctly classified 2921, incorrectly classified 577, test accuracy 0.835049

Part II: Analysis (40%+20% Extra Credit)

The analysis should be short and concise - we primarily care that you report the performance statistics asked for accurately.

Question 5 (4 points): Learning With Restarts

Neural Network as we have implemented them work by gradient descent from a random starting point. This is a form of local search, so we have the typical local search problem of landing in a local maxima that may or may not be close to the global maxima. As in other forms of local search, the the solution to this is random restarts. Run 5 iterations of testPenData and testCarData with default parameters and report the max, average, and standard deviation of the accuracy.

Question 6 (4 points): Varying The Hidden Layer

Vary the amount of perceptrons in the hidden layer from 0 to 40 inclusive in increments of 5, and get the max, average, and standard deviation of 5 runs of testPenData (you'll want just let your computer run this one for a while) and testCarData for each number of perceptrons. Report the results in a table. Additionally, produce a learning curve with the number of hidden layer perceptrons being the independent variable and the average accuracy being the dependent variable. Briefly discuss any notable trends you noticed related to increasing the size of the hidden layer has on your neural net.

Question 7 (2 points Extra Credit): Learning XOR

As you've learned in class, adding the hidden layer allows Neural Nets to learn non-linear functions such as xor. To show this in effect, produce the set of examples needed to train a Neural Net to compute a 2 variable xor function. Train a neural net without a hidden layer with it and report the behavior. Then, run it on neural nets starting with 1 perceptrons in the hidden layer and increasing until you get a neural net that works well. Are the results what you expected?

Question 8 (2 points Extra Credit): Novel Dataset

Explore the UCI Machine Learning Repository or other ML datasets found online, and select a new dataset to try your Neural Network with. Write a method in NeuralNetUtil.py called buildExamplesFromExtraData that is akin to the other get methods we wrote. Answer question 5 for this dataset, and submit your NeuralNetUtil.py in addition to NeuralNet.py for the option of extra credit.