Assignment III) Neural Networks

ELL-409 10 Marks

Submission date: 9/11/2022 - 11:59 pm

I) Data Description:

The dataset to be used in this assignment is adapted from Zillow's Home Value Prediction Kaggle competition data. The number of input features have been reduced and the task is to predict whether the house price is above or below median value. Please visit the link below to download the modified dataset.

Download Dataset

Note that to download this notebook from Github, you have to go to the front page and download ZIP to download all the files.

II) Create a Training and Testing Neural Network:

The input features are given in the first 10 columns and the last column contains the features that need to be predicted i.e.. **Is the house price above the median or not?** (1 for yes and 0 for no)

Steps to create and train a neural network are:

1) Import the data
2) **Think how you should normalize the data.**
3) Initialize the network parameters.(This requires you to initialize number of hidden layers and to choose an activation function)
4) Initialize weight and bias of the network
5) Train the network using Training data and Training Targets.

III) Parameter Optimisation:

An important issue in neural networks is parameter optimization. First, select a performance measure which will be used to compare the performance of the different parameters on the validation set. You should use stochastic gradient descent with momentum for all experiments, but you can experiment with other training algorithms if you wish so. Then you should optimize the parameters as follows:

1) First select a network architecture and an activation function that you believe is a reasonable starting point. Explain your motivation for selecting this architecture and activation function. Define a stopping criterion, a weight initialisation method and a momentum and learning rate update schedule.

2) Optimize the initial learning rate (disable regularization). Explain how you found a good initial value for learning rate. Save the plot of the training and validation loss and training and validation classification error (from epoch 1 until the stopping criterion is met).
3) Optimize the learning rate update schedule. Include a table where you present the results on the validation set of the different learning rate schedules you have tested.
4) Use dropout and report if there is any improvement in the validation performance. Explain what changes you have made to the network and/or training procedure when you use dropout.
5) Use two other types of regularization (any two you wish) and compare their performance with dropout. Explain which regularization parameters you optimized and present a plot which shows the performance on the validation set as a function of the parameters that needs to be optimized. Also discuss why regularization is needed.
6) Optimize the topology of the network, i.e. the number of hidden layers and the number of neurons in each hidden layer, the size of your input, and the number of neurons in the output layer. Include a plot which shows the performance on the validation set as a function of the number of layers. Include another plot which shows the performance on the validation set as function of the hidden layers size.
7) Train a network using the optimal set of parameters you have found so far and a different activation function than the one you defined in step 1. Comment on the performance on the validation set and discuss if it is any different than the initial activation function you used.
8) Train a network using the optimal set of parameters and include a figure which shows the training and validation loss and another figure which shows the training and validation prediction error. Present in the same figure the curves you saved in step 2. Comment on their differences. Save also this network since you should include it in your submission.

IMPORTANT NOTE: It is impossible to run an exhaustive search, so think of a reasonable strategy when you optimize parameters (e.g. don't explore too many similar parameter values) and put more effort on optimizing parameters you think are more important. Also whenever possible use some default values, e.g., start with a momentum of 0.5 and increase it (at the same epoch that learning rate begins to decrease) linearly to 0.9.

IV) Performance Estimation:

Test the performance of the network trained with the optimal set of parameters on the test set and report the confusion matrix.

V) Deliverables:

Submissions will be done on moodle in the form of a pdf report and a zip containing source code(with compiled source) and all output files generated. Only one of the group members will upload the assignment.