

MILIN DESAI

CSC 578 HW 5

Backprop Hyper-Parameters

Overview

In this home work, we have used the code from the book NND but modified as per professor's guidance. We have use different parameters in the home work like activation form, regularization, lambda, activation output and hidden output.

This home work was design in mainly two part first was making an application or modifying the application code and the second part was consist of the testing which is really crucial part in the homework .

Network Code

- Hyperparameters

There are mainly three part hyper parameter first cost function which contains of the QuadraticCost', 'LogLikelihood', 'CrossEntropy'.Hyper-parameters are consist of mainly this parameters which are really important and building block of this application cost, act_hidden, act_output, regularization, and dropoutpercent.

There are main four activation function used in this graph "Sigmoid", "Tanh", "ReLU", "Softmax".

Regularization and lambda there are two regularization in this application L2 and L1 and lambda value we have used for last 4 test cases.

Drop out function we modified in code also and we have used function like drop out mask. Basically let use no what are the percent drop out.

Completeness :

For this part, you will need to make modifications to the Network Code, adding some hyper-parameters and modifying some functions.

- Hyper-parameters:
- Cost function **Completed**
 - **QuadraticCost, CrossEntropy, LogLikelihood**
- Activation function **Completed**
 - **Sigmoid, Tanh, ReLU, Softmax**
- Regularization L1 and L2 **Completed**
 - Dropout rate **Completed**
- Functions to modify (minimally):
 - set_parameters() **Completed**
 - feedforward() **Completed**
 - backprop() **Completed**
 - update_mini_batch() **Completed**
 - total_cost() **Completed**

For one training one test Iris data

```
([0.6818657040185175, 0.6571643765632978],  
 [0, 0],  
 [0.6472763926428745, 0.5701794908614409],  
 [0, 1])
```

After modification of code I run all 12 experiments first 8 experiments doesn't require lambda values and the activation form.

For 11 12 just data set were different with one more parameter of drop out percent. Output results were similar for many outputs. Although, it wasn't the perfect match with the professor output. I try to rectify that why I am not getting the perfect out as the professor but was not able to find.

My overall Homework is Complete except few results were not matching with professor's code.

Reflection:

I really enjoyed this homework , Although I was frustrated in some part because I spent hours resolving errors but It was great deal of the learning and I felt like I am working on some project and some parts are my responsibility Toughness of the assignment was just right amount or maybe a bit more. But with enough time on research and readings was able to understand concept more throughly.