Jared Miller Proj3 Answers

<u>Written Question</u>
1. **Describe the process you went through in determining different architectures and hyper-parameters to try for the regression model. Do you feel you had a systematic approach and what would you change if you started over? (4 pt)**

If I'm being completely honest at the start, deciding parameters was quite random. When I was trying to create the parameters at the start I wrote down the structure of the neural network on a sheet of paper. I had the neural network start simple. (1 hidden layer, 3 neurons per hidden layer) This was helpful because I could make sure that all the pieces fit together before trying to tune it. After this, I started to tune it. I used 250 hidden layer neurons because it was in the middle of the recommended range. For deciding my batch size and my learning rate I did some trial and error. I incremented the batch size in chunks of 50 once I got a learning rate I was happy with. I decided on the learning rate by trial and error. I did this by setting the learning rate towards the higher end of the recommended and if it felt like I was overshooting I lowered it. I did all this trial and error until I was within spec. I didn't adjust the amount of hidden layers because I could remain in spec without messing with it. In terms of other architecture, I started with a ReLU for my activation functions and it was within spec so I didn't tune it further. So to start I didn't have a systematic approach and it felt more like trial and error. As I learned more about how the networks work I was able to develop a process.

If I were able to start over I think that I would have some default parameters and tune them based on the performance of the model. This is what I did for questions 3 and 4 and it seemed to work perfectly.

Batch Size = 100
Learning Rate = 0.1
Hidden Layer Neurons = 200
Hidden Layers = 1
Activation Unit = ReLU

2. **Did your experience with getting the regression model to work inform your architecture and hyper-parameter search for the digit classification problem? What about the RNN model? Explain why and how. (6 pt)**

Yes, my experience with getting the regression model to work informed my architecture and hyperparameter tuning. Once I got my RegressionModel working I used that as a starting point for my parameters.

From the Regression Model, I was pretty sure that with enough tuning my structure could work for the digit classification problem. From the values I got from my Regression Model, I started by tuning my learning rate. Because the digit classification was taking significantly longer than the other problem I increased my learning rate significantly. This same intuition also made me decide to increase my batch size and the number of hidden neurons. I thought that because I was

having to make more dynamic predictions all of the variables would need to be higher than in the RegressionModel.

A similar thing can be said for the RNN. I started with the default values (the values from the regression model) and then proceeded to tune them based on my observations. I increased the number of hidden neurons once again because this problem was more dynamic. I kept the batch size the same. After this, all that was needed was to get the recurrent nature of the RNN functional and I passed the tests.

This assignment opened my eyes when it comes to training with neural networks. I would have never thought there was a general structure that could learn such dynamic and different problems. The reason why the choices from the Regression Model affected all my tuning is because they use the same structure but with different tuning.

Self Analysis
1. **What was the hardest part of the assignment for you?**

The hardest part of the assignment for me was understanding neural networks. I've never implemented a neural network before and so it was quite difficult to get a hold of the concepts.

2. **What was the easiest part of the assignment for you?**

Using the assignment as a reference was the easiest part of the assignment for me. I was super impressed with the detail in the assignment description and the code comments. It helped me understand the concepts significantly better.

3. **What problem(s) helped further your understanding of the course material?**

This was one of the coolest assignments I've ever done. Each problem helped further my understanding of the course material.

4. **Did you feel any problems were tedious and not helpful to your understanding of the material?**

No

5. **What other feedback do you have about this homework?**

None