

Assignment 1 - Report

- Basic implementation of the forward propagation was simple and straight forward. The implementation of the various activation functions and their derivatives was cross checked using manual calculations as well as code based tests, including numerical differentiation. The tests returned positive results.
- Implementation of backward propagation was surprisingly difficult and consumed the greatest portion of the time spent on the assignment. There were two primary challenges in the implementation of backward propagation.
- The first challenge was to map the backpropagation formulae of the various derivatives from scalar values to matrices. This was my first attempt at actually implementing such a scenario and it took a lot of time and effort to understand the nuances of both linear algebra in programming as well as of numpy.
- The second challenge was to achieve the required accuracy in calculation of backpropagation gradients. This part of the implementation turned out to require more time and effort than the rest of the assignment put together, possibly due to a bug that has yet to be caught. The current status of the implementation is that whereas all the layers calculate the gradients correctly when tested individually, backpropagation seems to accumulate the error in gradient approximation. Thus, when the network as a whole is tested for accuracy, the approximation error exceeds the acceptable range.
- In order to calculate loss and gradient of loss w.r.t. the input of the softmax output layer, literature was consulted and it was determined that the most simplified version of the equation is:
$$\text{inp_grad} = Y_{\text{pred}} - Y$$
- Implementation of the Neural Network itself seems to be fairly simple once the layers have been implemented, however, due to the aforementioned bug that causes gradient errors to accumulate, it has been impossible to confirm if the application of SGD and GD are 100% accurate.
- In conclusion, whereas implementation of the Neural Network has been completed, there exists a bug that has not been identified as of yet. Hence, the NN is unable to be trained and it is impossible to check the accuracy of the implementation.