Claude Shyaka

ID#: 801326243

ECGR-5105

**Homework 5: Gradient Descent, Backpropagation, and ANN**

**GitHub Repo: https://github.com/claudeshyaka/ml/tree/main/**

In this exercise, we used Fully connected neural networks (NN) and convolution neural networks from the Pytorch neural network library to develop models for image classification using the CIFAR10 dataset. In the first part, we developed models using fully connected neural networks, then we used convolution neural networks to develop more efficient image classification models. The results of our analysis are presented below.

1. In this section, a fully connected neural network was created with 3072 input features and 10 output (classes) labels. The network only had one hidden layer with a size of 512 nodes. We trained the NN for 300 epochs and obtained a training accuracy of 1.0 and a validation accuracy of 0.47. These results suggest that the model is overfitting.
2. Next, we developed another fully connected NN with 3072 input features and 10 output classes, however, this time we extended the network with two more hidden layers. The first hidden layer has 1024 nodes and the second has 512 nodes. The NN was trained for 300 epochs and the training and validation accuracies were reported as 1.0 and 0.47 respectively. Thus, these results suggest that the fully connected neural network is overfitting the data. In conclusion, a better model of NN would be required for better-generalized results.
3. In this section, we developed convolution NN models to classifier images using the CIFAR10 dataset. Here the Cond2D module from Pytorch was used to develop the NN models. We also used max pooling for down-sampling. In addition, the activation function used was Tanh. The CNN model we built had one convolution hidden layer and one linear fully connected hidden layer. After training for 300 epochs, the training and validation accuracies were reported as 0.82 and 0.61 respectively.
4. The model in part 3 was extended to a neural network with two hidden convolution layers and two hidden linear fully connected layers. After training for 300 epochs, the training and validation accuracies were reported as 1.0 and 0.7 respectively. In addition, the training time was also reported as approximately 4 hours. Thus, these results suggest that the model is overfitting the data.