CSE474/574: Introduction to Machine Learning (Fall 2017)

Project 3: Classification

Presented By:
Anirban Chatterjee – 50249214
Vikram Singh – 50247207

In this project, we are supposed to implement and evaluate different classification algorithms. The classification task will be to recognize a 28x28 grayscale handwritten digit image and identify it as a digit among 0,1,2, ...,9.

We start with the Logistic Regression model. Logistic regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. In logistic regression, the dependent variable is binary or dichotomous, i.e. it only contains data coded as 1 (TRUE) or 0 (FALSE).

Here we train the model on MNIST dataset and tune the hyper parameters to lower the error. We found that using small batch size and higher number of iterations result in a more accurate model. In our model, we have kept the learning rate as low as 0.5. Our accuracy for this model is 0.92 for MNIST and 0.10 for USPS

O/P-> Logistic Regression accuracy on MNIST: 0.9251

O/P-> Logistic Regression accuracy on USPS: 0.1079

Then we move on to a single layer neural network. Here the main work is done by the hidden layers. Reducing the range of values of initial weight improves the results. We have kept the value at 30. Our learning rate is 3.0 and here we have kept the minibatch size as 10. We found that as there are hidden layers, keeping the batch size small results in better accuracy as increasing the number of hidden layers increases the complexity and running time as well. Our accuracy for this model is 0.94 for MNIST and 0.46 for USPS

O/P-> Single layer neural network accuracy on MNIST: 0.9493

O/P-> Single layer neural network accuracy on USPS: 0.4655

In machine learning, a convolutional neural network (CNN) is a class of deep, feed-forward artificial neural networks that has successfully been applied to analyzing visual imagery. Finally, we implement a convolutional neural network model and train it using our MNIST data set. Here we kept the batch size at a higher end at 100 and used 500 and the number of steps. Our accuracy for the model is 0.97 for MNIST and 0.36 for USPS.

O/P-> Convolutional neural network accuracy on MNIST: 0.9703

O/P-> Convolutional neural network accuracy on USPS: 0.3667

Conclusion:

The "No Free Lunch" theorem states that there is no one model that works best for every problem. The assumptions of a great model for one problem may not hold for another problem, so it is common in machine learning to try multiple models and find one that works best for a particular problem. In our project, we used 3 different models on two different datasets yet no particular model works best on all. So we can infer that the No Free Lunch theorem is correct.