Machine Learning CSE 574 Programming Assignment 2

Handwritten Digits Classification

Group 55

Pratibha Arjun Barsale – 50247005

Venkata Krishnan Anantha Raman - 50246287

Overview

The assignment aims at implementing a Multilayer Perceptron Neural Network for evaluating its performance for classifying handwritten digits and face recognition. Also, the performance of this Single Neural Network is compared with the Deep Neural Network and Convolution Neural Network using the TensorFlow library.

Implementation

The Neural Network implemented is evaluated on the real data. The data is loaded from the MNIST dataset containing the 10 matrices for training set and testing data. The training data is further split into training and validation data. Thus, the experiments have been carried out on the three datasets namely, training, validation and testing.

**Feature selection** is done on the input data. All features that don’t carry any value are deleted. We are using the numpy.ptp function, which will give us the peak to peak value. A zero peek to peek value implies that the feature doesn’t carry any value. If feature’s peak to peak is zero on the input and test data we delete the feature from the input.

The Neural Network consists of three layers, the input layer, the hidden layer and the output layer. Feature selection is done. The two weight vectors namely w1 and w2 are assigned a random weight given the number of unit in the input and output layer. The data label is then predicted using the w1 and w2 weight vectors. Then the value of objective function and gradient are computed using the backpropagation algorithm.

By setting different values to the regularization hyper-parameters of the network, namely lambda and number of hidden unit nodes, various comparisons have been made determining the accuracy and the runtime of the neural network for the three data sets.

Choosing Hyper parameters:

In this we choose λ from 0 to 60 for various hidden units ranging from 1 to λ versus the accuracy for hidden units like 1, 10, 25, 40,50

**Hidden Units vs Training Time:**

The graph plots the Training time with the number of hidden nodes, for various lambda values. Based on the above graphs we decide ideal λ as 30, Number of hidden nodes as 25.

**Number of Layer (vs) Accuracy on Celeb Dataset:**

The above plot compares neural network with single layer and deep neural network with 3,5,7 hidden layers. It could be seen that accuracy drops from 84.63 for single layer to 79%, 77% and 76% for 3,5,7 Hidden Layers. This is because with the addition of more layers we are overfitting the problem and this results in reduced testing accuracy.

**Analysis of Convolutional Neural Network:**

The plot is done between number of hidden layers and the Time Taken & Accuracy. It could be seen that the Accuracy starts with 9.7% for 1 iteration, increases and reaches 98.7% with 10k Iterations. It can also be observed that the time taken increases with the increasing the number of iterations.

|  |  |
| --- | --- |
| **Confusion Matrix for Iteration: 1**  [[ 0 0 0 979 0 0 0 1 0 0]  [ 0 0 0 1134 0 0 0 1 0 0]  [ 0 0 0 972 0 56 0 4 0 0]  [ 0 0 0 997 0 12 0 1 0 0]  [ 0 0 0 976 0 5 0 1 0 0]  [ 0 0 0 880 0 10 0 2 0 0]  [ 0 0 0 942 0 0 0 16 0 0]  [ 0 0 0 1027 0 1 0 0 0 0]  [ 0 0 0 971 0 1 0 2 0 0]  [ 0 0 0 1006 0 3 0 0 0 0]] | **Confusion Matrix for Iteration: 100**  [[925 0 1 27 1 0 9 6 11 0]  [ 0 960 1 18 1 0 6 0 144 5]  [ 23 3 815 77 23 0 24 28 37 2]  [ 12 5 37 852 2 3 5 45 40 9]  [ 1 2 8 0 840 0 22 6 31 72]  [ 28 21 21 242 64 355 21 62 55 23]  [ 36 14 15 3 64 4 777 1 44 0]  [ 0 12 19 7 11 0 0 915 23 41]  [ 14 6 19 130 24 6 7 43 703 22]  [ 12 5 11 17 200 0 0 140 23 601]] |
| **Confusion Matrix for Iteration: 1000**  [[ 958 0 1 2 0 4 8 3 4 0]  [ 0 1122 1 3 1 0 3 0 5 0]  [ 13 17 902 28 19 1 12 20 18 2]  [ 2 8 7 939 0 19 0 16 11 8]  [ 1 3 3 0 937 0 9 1 2 26]  [ 6 3 0 27 8 817 18 2 7 4]  [ 7 7 1 2 22 14 903 0 2 0]  [ 0 13 22 7 6 0 0 937 2 41]  [ 6 11 3 27 14 16 7 11 859 20]  [ 8 8 2 11 39 7 1 15 3 915]] | **Confusion Matrix for Iteration: 10000**  [[ 975 0 1 0 0 0 0 1 3 0]  [ 0 1125 4 0 0 0 2 2 2 0]  [ 2 1 1022 0 1 0 0 1 5 0]  [ 1 0 0 1001 0 3 0 2 3 0]  [ 0 0 1 0 976 0 0 1 0 4]  [ 2 0 0 5 0 879 2 1 3 0]  [ 10 2 0 0 4 3 938 0 1 0]  [ 1 2 9 2 0 0 0 1010 1 3]  [ 5 0 2 1 1 1 0 2 960 2]  [ 5 3 2 2 7 3 0 3 1 983]] |

**Conclusion and Inferences:**

* We could see form the celeb dataset with single and multilayer neural network that with the increasing the hidden units the accuracy decreases. This is a typical case of overfitting.
* Generally, with the increase in the hidden nodes the Accuracy should improve, stabilize and decrease if it is increased beyond a specific level.
* On a convolutional neural network, with the increasing the number of iterations we will get better accuracy.