UNIVERSITY AT BUFFALO

CSE – 574 INTRODCUTION TO MACHINE LEARNING

PROGRAMMING ASSIGNMENT-2 REPORT

HANDWRITTEN DIGITS CLASSIFICATION

GROUP 11

GROUP MEMBERS:

SIDDHESH CHOURASIA: 50415033

TANUSH TRIPATHI: 50411177

SARVESHWAR SINGHAL: 50418642

1. Introduction

This project implements a Multilayer Perceptron Neural Network and evaluates its performance in classifying handwritten digits. Further, same neural network implementation is used to analyze a more challenging face dataset and compare the performance of the neural network against a deep neural network and convolutional neural network which are implemented using the TensorFlow library. Our Neural Network Model Implements Forward feed and back propagation methodology to adjust weights of model for predictive learning. The parameters in Neural Network model are the weights associated with the hidden layer units and the output layer units. This project discusses relations and dependency of accuracy and training time of Neural network on parameters such as number of Hidden nodes, regularization coefficient, maximum iterations of conjugate gradient algorithm to perform optimization task.

2. Feature Selection

In the dataset provided, one can observe that there are many features which values are the same for all data points in the training set. With those features, the classification models cannot gain any more information about the difference (or variation) between data points. Therefore, we can ignore those features in the pre-processing step.

Below Uninformative Pixels are removed before machine learning algorithm is applied:
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 52, 53, 54, 55, 56, 57, 82, 83, 84, 85, 111, 112, 140, 141, 168, 196, 476, 532, 560, 615, 643, 644, 645, 671, 672, 673, 699, 700, 701, 727, 728, 729, 730, 754, 755, 756, 757, 758, 759, 780, 781, 782, 783]

3. Summary of Best Accuracies

Below we present summary of achieved results before we start discussing them in more detail.

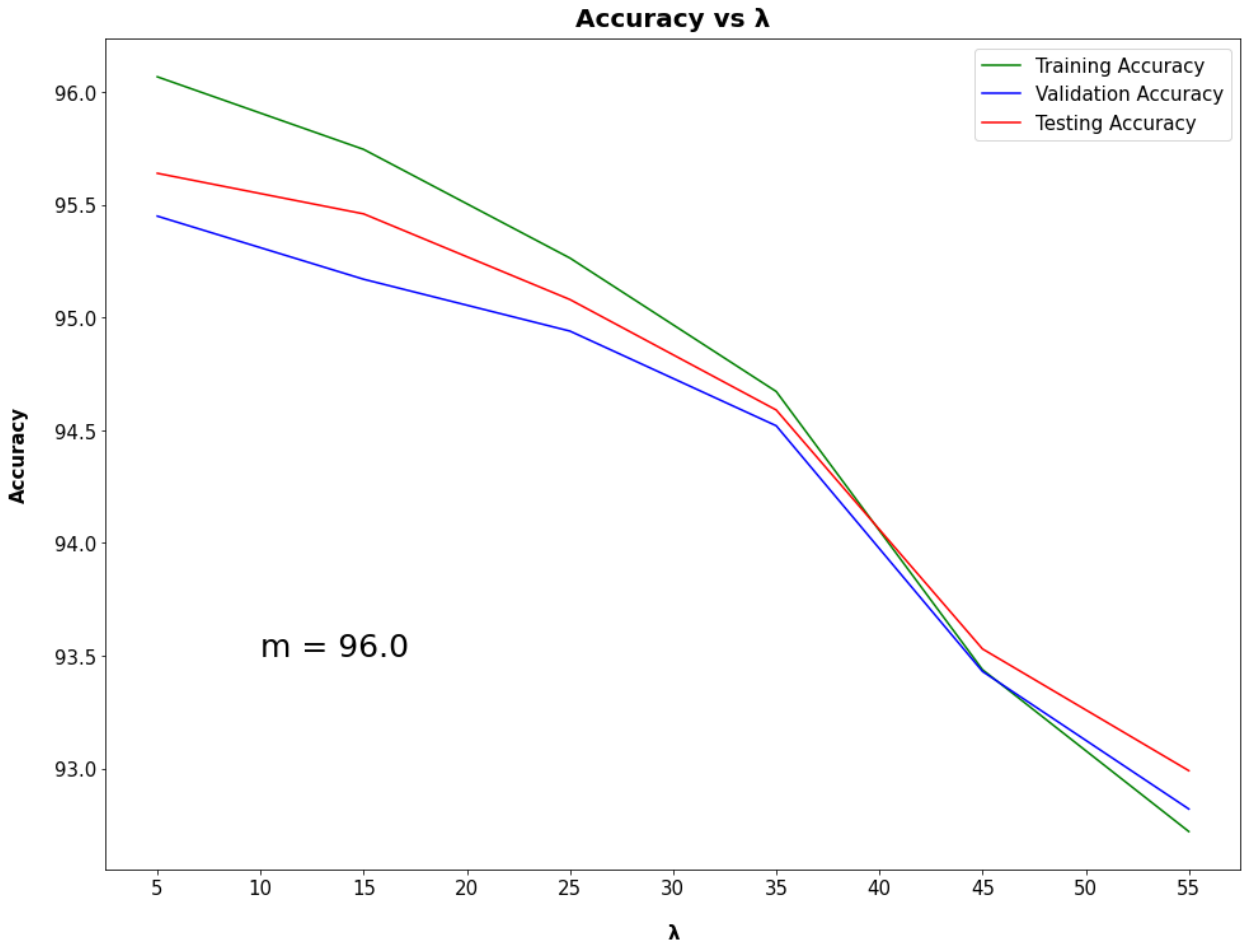
Maximum Iterations	Hidden Nodes	Regularization coefficient	Training Accuracy	Validation Accuracy	Test Accuracy
50	100	5	94.78%	93.99%	94.61%
100	120	5	98.338%	97.05%	97.19%
150	100	5	98.992%	97.36%	97.42%
200	100	5	99.422%	97.37%	97.75%
250	100	5	99.512%	97.85%	97.87%

Table 1: Summary for Best achieved Accuracies for different Parameters

4. Accuracy analysis of Neural Network for handwritten digits

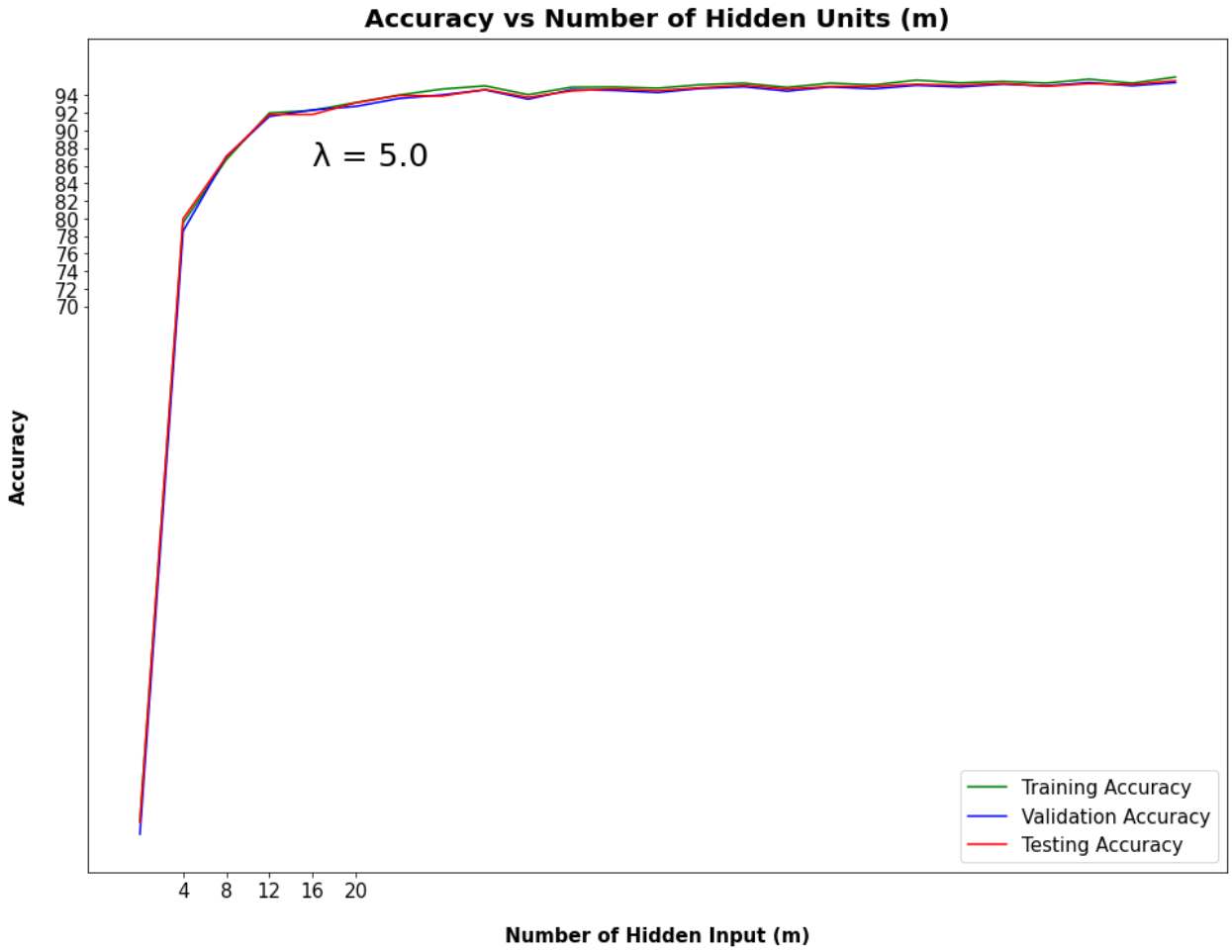
A) Accuracy V/S λ:

We are using regularization in Neural network so that overfitting problem will be avoided. According to our data set, validation set and test set accuracy is decreasing as we are increasing the (λ).

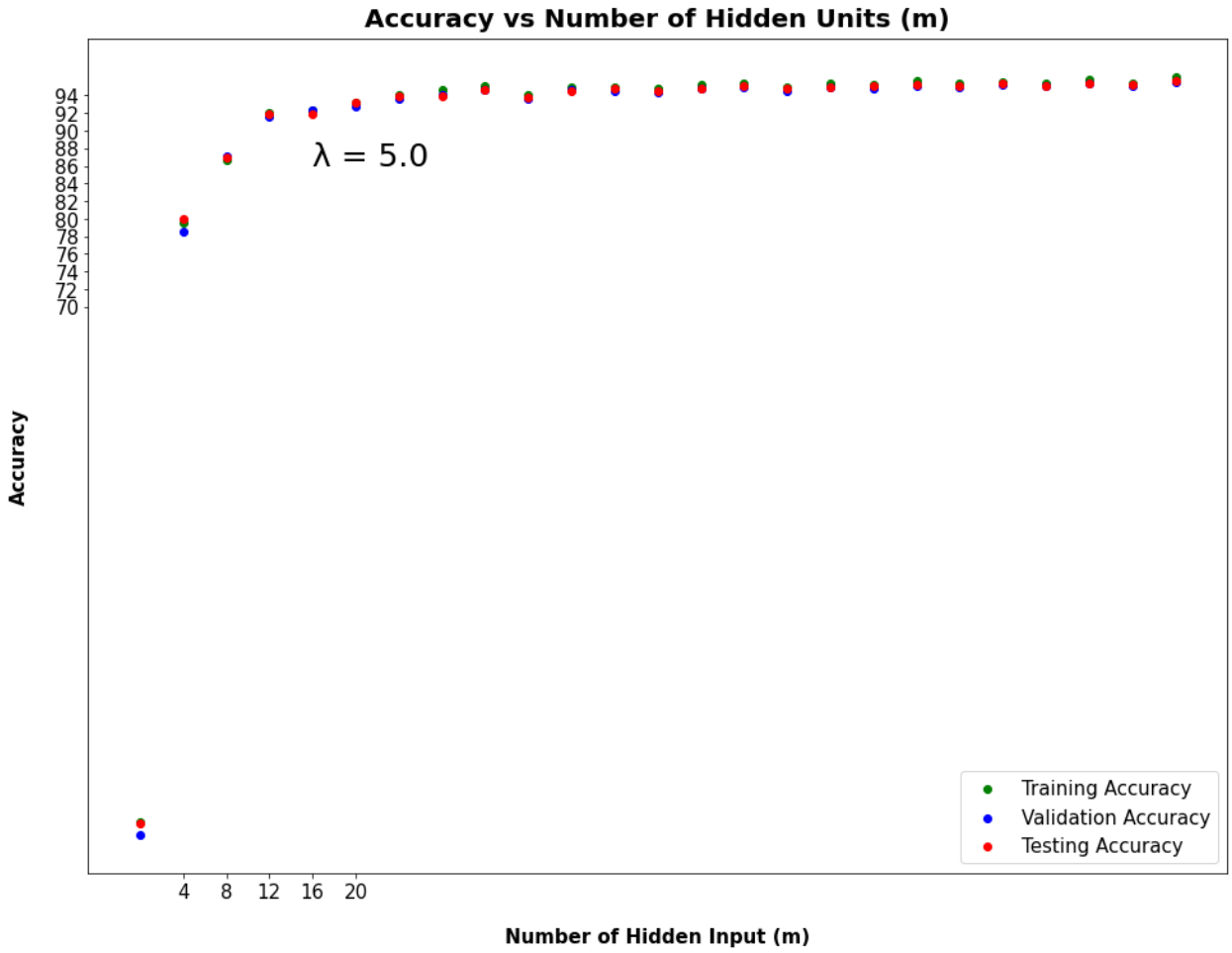


From the above figure, we can infer that although the regularization coefficient(λ) is being used to solve the overfitting problem by controlling the magnitude parameters., due to which it lowers the accuracy of the training set.

B) Accuracy V/S Hidden Nodes:



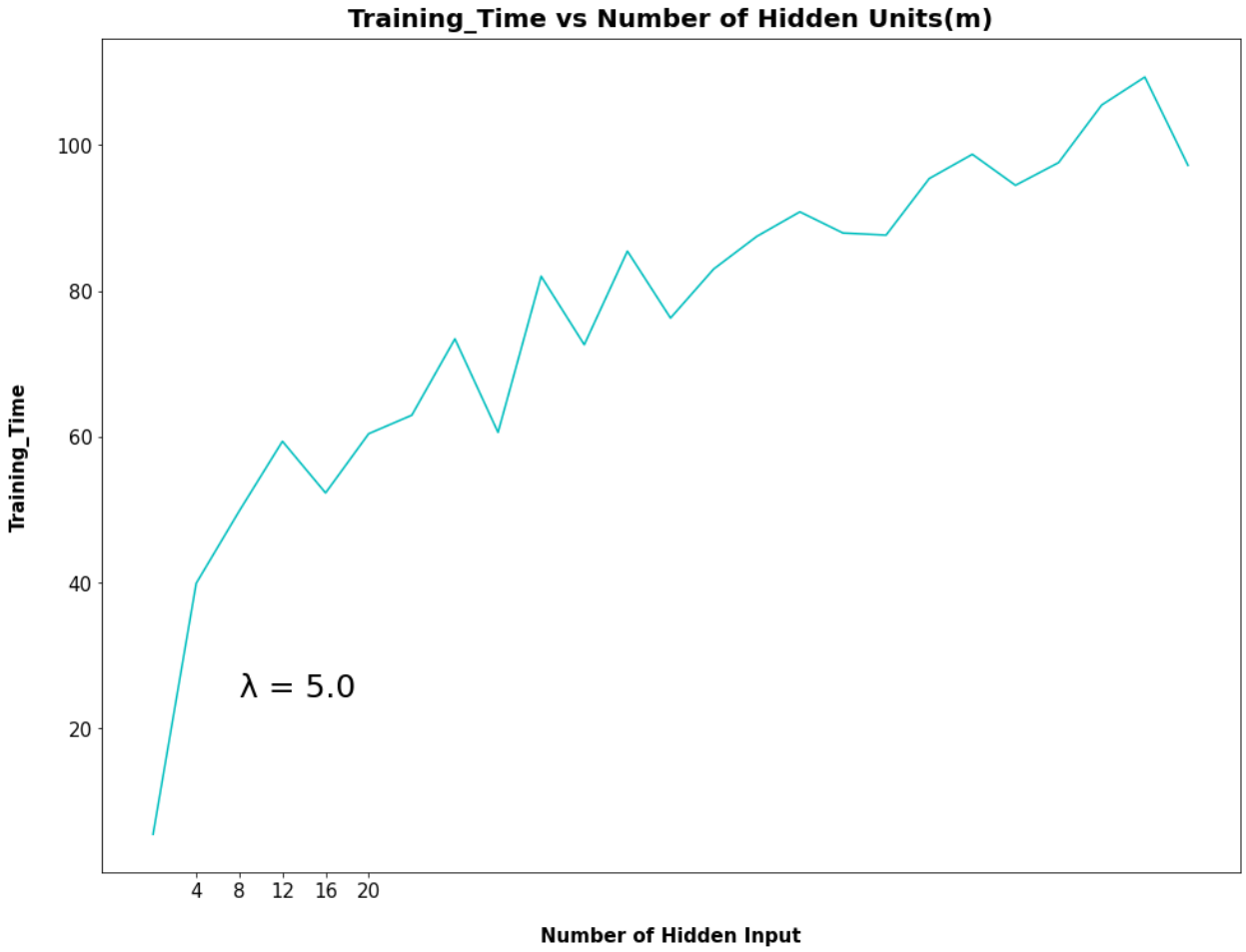
From the above figures, test data, validation data and train data accuracy is increasing with increase in no of hidden nodes and max number of iterations.



From this we can conclude that accuracy of the neural network increases as no of hidden units and layers increase.

C) Hidden Nodes V/S Training Time:

In this we have plot the graphs for training time for all three training set, validation set, and test set against number of hidden nodes.



From the above plots, we can extrapolate that the training time of the neural network increases as we increases the number of nodes. It also increases as the number of iterations of optimization function increases.

5. Comparative Analysis of Single Hidden layer implementation and Deep Neural Network on the CelebA data set

5.1 Performance of Implemented algorithm on face Dataset

Training Set Accuracy	91.241%
Validation Set Accuracy	89.681%
Test Set Accuracy	90.310%
Evaluation Time	456

5.2 Deep Neural Network Performance on face Dataset

Number Of Hidden Layers	Learning Rate	Training Epochs	Batch Size	Execution Time	Accuracy
2	0.001	200	200	270	84.82%
3	0.001	200	200	300	84.2%
5	0.001	200	200	315	83.9%
7	0.001	200	200	544	83.1%

Observation: In general, the accuracy of Neural Network tends to increase for increase in Neural Network units. But, after certain threshold addition of many hidden nodes can lead to decreasing accuracy. This is seen above. Hence, accuracy reduces with increase in number of hidden layers

Comparing Section 5.1 and 5.2, we present below conclusion:

Result1: Our implemented algorithm of Single Layer Forward feed and backward propagation surpasses Deep Neural Networks in terms of accuracy percentage. Our implementation algorithm gives accuracy of 90.310% on test data whereas Deep Neural Networks gives best accuracy of 84.82% with 2 Hidden Layers.

Result2: Deep Neural Networks Implementation performs better than our implementation of forward feedback and backward propagation in terms of training time, though time difference is not very significant.

6. Analysis of Convolutional Neural Network on the MNIST data set (Present in TensorFlow Library)

6.1 Convolutional Neural Network Performance on MNIST Dataset

Size of Filters	Number of Filters 1	Number of Filters 2	Execution Time	Accuracy
5*5	16	36	160	98.8%
5*5	25	36	222	98.8%
5*5	50	75	396	99%
10*10	50	75	1560	99%
5*5	100	100	1020	99.1%
5*5	200	200	3606	99.3%

Observation: In general, the accuracy of Convolutional Neural Network tends to increase for increase in Number of Filters. Also increasing the Number of Filters increases the execution time. Also increasing the size of the filters have no impact on the accuracy , as seen above. Although the one with the Number of Filters equal to 200 gives the best accuracy , it takes a lot of time to execute , comparing to the one with Number of Filters 100 , which gives almost the same accuracy with drastic reduction in execution time.

Result1: Our implemented algorithm of Single Layer Forward feed and backward propagation falls behind Convolutional Neural Networks in terms of accuracy percentage. Our implementation algorithm gives an accuracy of 90.310% on test data whereas Convolutional Neural Networks gives best accuracy of 98.3% with 2 Layers.

Result2: Convolutional Neural Networks Implementation performed poorer than our implementation of forward feedback and backward propagation in terms of training time.

7. Conclusion and Summary

- Our Implemented Neural Network Training time increases with more number of Maximum iterations (maxiter) in optimization function.
- Training time is also directly proportional to number of hidden units and hidden layers, i.e it increases with increase in the number of hidden units and hidden layers.
- The accuracy of the implemented neural network on Training data is inversely proportional to the regularization coefficient, i.e it decreases as the value of regularization coefficient (λ) increases.
- Like Training data, even Validation data and Test data decreases as the value of Regularization coefficient value(λ) increases, although it is theoretically expected to increase.
- In general, the accuracy of Neural network tends to increase up for increase in neural network units, which reduces after certain addition of hidden nodes.