Report:: TIPR Assignment - II

Aniket Pramanick
SR. No. 14686
Monday 4th March, 2019

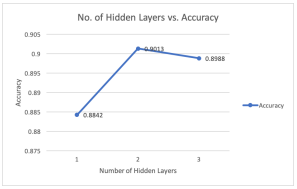

Figure 1: (MNIST Data)Task1(i): Number of Hidden Layers vs Accuracy

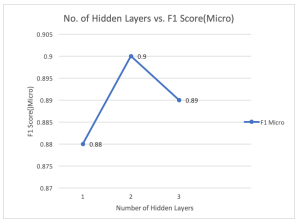

Figure 2: (MNIST Data)Task1(ii): Number of Hidden Layers vs F1 Score(Micro)

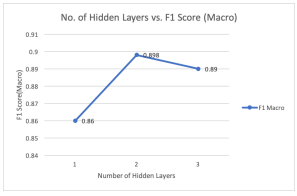

Figure 3: (MNIST Data)Task1(iii): Number of Hidden Layers vs F1 Score(Macro)

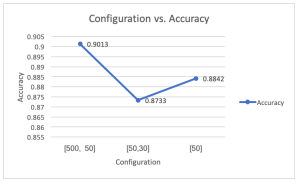

Figure 4: (MNIST Data)Task2(i): Configuration vs Accuracy

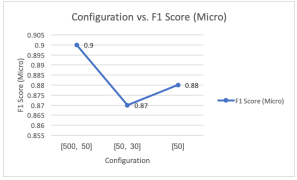

Figure 5: (MNIST Data)Task2(ii): Configuration vs F1 Score(Micro)

Abstract

The objective of this assignment is to understand the mechanism of building, training and testing multi-layer *Neural Network*. To understand the effect of various *Hyper-parameters*, such as *Number of Layers*, *Number of Nodes per Layer*, *Activation functions*, various experiments are performed, and the results are furnished below. To reveal the effects of various Hyper-parameters, the following tasks were performed with two given data sets, *MNIST* and *Cat-Dog*. The data set descriptions were shared in the assignment manual.

1 Experiments with MNIST Dataset:

The neural network trained with Learning Rate **0.001**, Regularization Parameter **0.00001**, Epochs **5000** and Number of Batches **1**. For training and validation, the entire given data set has been randomly in **9:1** ratio.

1.1 Task I: Number of Layers vs. Prediction Metrics

The plots are furnished in Figure 1, 2 and 3.

1.2 Task II: Number of Neurons vs. Prediction Metrics

The plots are furnished in Figure 4, 5 and 6.

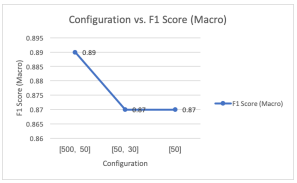

Figure 6: (MNIST Data)Task2(iii): Configuration vs F1 Score(Macro)

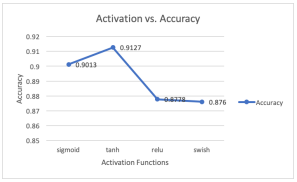

Figure 7: (MNIST Data)Task3(i): Activation vs Accuracy

1.2.1 Task III: Different Activation Functions

Two Hidden Layers of size 500 and 50 has been used to train and test the Network. The network is tested on *sigmoid*, *tanh*, *relu* and *swish* activation functions.

The plots are furnished in Figure 7, 8 and 9.

1.3 Task IV: Various Initialization Techniques.

Two Hidden Layers of size 500 and 50 has been used to train and test the Network. The network is tested on *zero*, *uniform random* and *gaussian random* initialization of weights and biases.

The plots are furnished in Figure 10, 11 and 12.

1.4 Task V: Keras

Two Hidden Layers of size 500 and 50 has been used to train and test the Network.

The plot is furnished in Figure 13.

2 Experiments with Cat-Dog Dataset:

The neural network trained with Learning Rate **0.001**, Regularization Parameter **0.00001**, Epochs

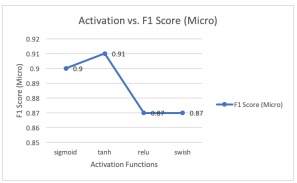

Figure 8: (MNIST Data)Task3(ii): Activation vs F1 Score(Micro)

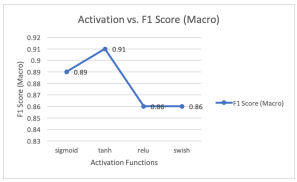

Figure 9: (MNIST Data)Task3(iii): Activation vs F1 Score(Macro)

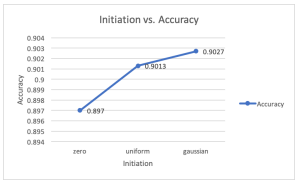

Figure 10: (MNIST Data)Task4(i): Initiation vs Accuracy

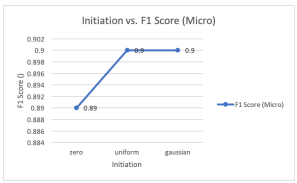

Figure 11: (MNIST Data)Task4(ii): Initiation vs F1 Score(Micro)

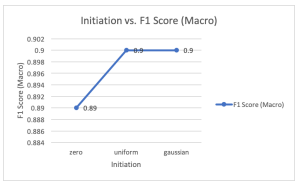

Figure 12: (MNIST Data)Task4(iii): Initiation vs F1 Score(Macro)

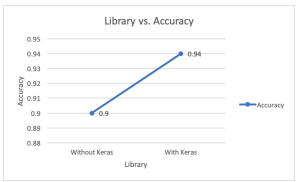

Figure 13: (MNIST Data)Task5: Library Usage vs Accuracy

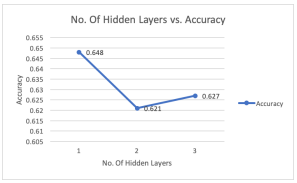

Figure 14: (CatDog Data)Task1(i): No. of Hidden Layers vs Accuracy

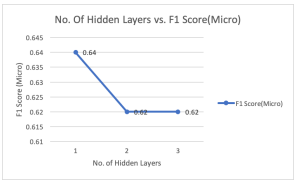

Figure 15: (CatDog Data)Task1(ii): No. of Hidden Layers vs F1 Score(Micro)

3000 Number of Batches **1**. For training and validation, the entire given data set has been randomly in **9:1** ratio. Moreover the 200x200x3 images are converted into *Monochrome*, then flattened into 40000 dimensional points and then reduced into 64 dimensional points using PCA.

2.1 Task I: Number of Layers vs. Prediction

Plots are furnished in Figure 14, 15 and 16

2.2 Task II: Number of Neurons vs. Prediction Metrics

The plots are furnished in Figure 17, 18 and 19.

2.2.1 Task III: Different Activation Functions

Two Hidden Layers of size 50 has been used to train and test the Network. The network is tested on *sigmoid*, *tanh*, *relu* and *swish* activation functions.

The plots are furnished in Figure 20, 21 and 22.

2.3 Task IV: Various Initialization Techniques.

Two Hidden Layers of size 50 has been used to train and test the Network. The network is tested

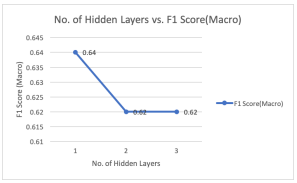

Figure 16: (CatDog Data)Task1(iii): No. of Hidden Layers vs F1 Score(Macro)

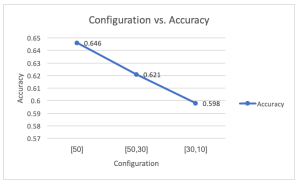

Figure 17: (CatDog Data)Task2(i): No. of Neurons vs Accuracy

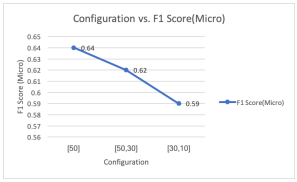

Figure 18: (CatDog Data)Task2(ii): No. of Neurons vs F1 Score(Micro)

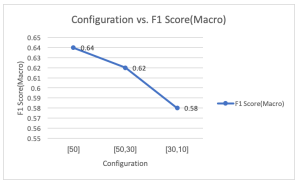

Figure 19: (CatDog Data)Task2(iii): No. of Neurons vs F1 Score(Macro)

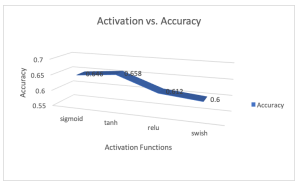

Figure 20: (CatDog Data)Task3(i): Activation vs Accuracy

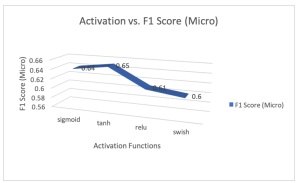

Figure 21: (CatDog Data)Task3(ii): Activation vs F1 Score(Micro)

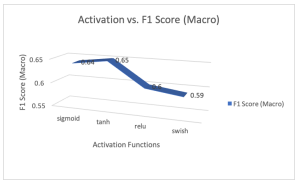

Figure 22: (CatDog data)Task3(iii): Activation vs F1 Score(Macro)

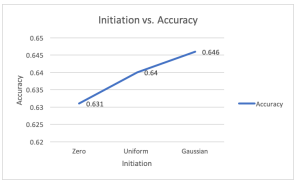

Figure 23: (CatDog Data)Task4(i): Initiation vs Accuracy

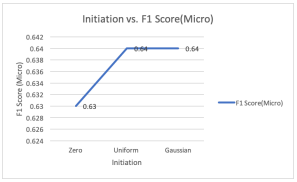

Figure 24: (CatDog Data)Task4(ii): Initiation vs F1 Score(Micro)

on *zero*, *uniform random* and *gaussian random* initialization of weights and biases.

The plots are furnished in Figure 23, 24 and 25.

2.4 Task V: Keras

Two Hidden Layers of size 50 has been used to train and test the Network.

The plot is furnished in Figure 26.

3 Performance on Dolphins and PubMed Datasets

MLP outperforms Bayes and Nearest Neighbour Algorithms, which is evident from the Figure 27.

References

Christopher Olah. 2017. *https://colah.github.io*.

Daniel Jurafsky. 2017. *Speech and Language Processing*. Stanford University, Stanford, USA.

Eduard Ma, Xuezhe; Hovy. 2016. *End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF*. ACL 2016.

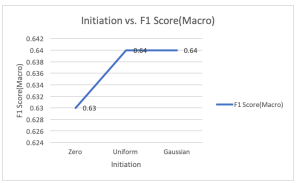

Figure 25: (CatDog Data)Task4(iii): Initiation vs F1 Score(Macro)

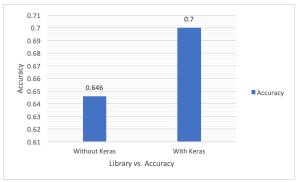

Figure 26: (CatDog)Task5: Library Usage vs Accuracy

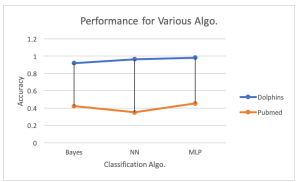

Figure 27: Performance on various Classification Algo.