

# Analyzing the Effect of Regularization in DNN

*(Accelerating Machine Learning Algorithms)*

(Ritika Sharma)

TEQIP\_INTERN

## **Internship Duration:**

One Month (11th June'19 - 10th July'19)

## **ABOUT:**

## **EDUCATION:**

Ritika Sharma is an undergraduate in Computer Science and Engineering at Atal Bihari Vajpayee Govt. Institute of Engineering and Technology , Shimla , Himachal Pradesh.

## **Working under Faculty Supervisor:**

Dr. Sparsh Mittal.

Assistant Professor, Department of Computer Science and Engineering, IIT Hyderabad

### **Research Interests :**

Processor architectures for machine learning, neural network accelerators, computer architecture (CPUs and GPUs), VLSI, high-performance computing, approximate computing.

## **Guided by:**

Subhrajit Nag

Pursuing Ph.D. in Computer Science and Engineering, IIT Hyderabad

## **Internship Provided by:**

Technical Education Quality Improvement Programme (TEQIP) Cell, IIT Hyderabad.

## **Acknowledgments:**

I am very thankful to the TEQIP team and also to the MHRD, Govt. of India for such an excellent internship program.

I would like to thank my mentor Dr .Sparsh Mittal and my guide Subhrajit Nag at IIT Hyderabad .

I would also like to thank Sharath R and Ritesh Gupta for giving knowledge about neural networks and various deep learning frameworks .

## **Abstract:**

Deep learning has become the most popular class of machine learning in recent times. However, deep learning techniques face overfitting problems, which is a hindrance to the model performance. Regularizing neural networks is an important task to reduce overfitting. Regularization is intended to reduce generalization error and not training error of the model .

We have analyzed the effect of various regularization techniques(L1 regularization , L2 regularization ,Dropout,Alpha Dropout,Feature Alpha Dropout,Data augmentation ) on two models . First model has ten convolution layers ,ReLU activation function ,batch normalisation and uses CIFAR10 dataset. Second model is made on the basis of architecture of Alexnet and also has CIFAR10 dataset.

## **Overview of Regularization:**

Regularization is a technique of reducing error on the test dataset . In other words ,we increase generalization of the model so that it will predict well on new data which is the data that is not given to model during train cycle .

### **Types of Regularization :**

1. L1 Regularization
2. L2 Regularization
3. Max Norm Constraint
4. Dropout
5. Data Augmentation
6. Early Stopping
7. Bagging
8. Adversarial Training

## **My contribution:**

We have found out that various regularization techniques have varying effect on the accuracy of DNNs . L1 regularization gives the least

accuracy on the testing data whereas Dropout(after fully connected (FC) layers) gives the maximum accuracy(In first model alpha dropout gives maximum accuracy whereas feature alpha dropout gives maximum accuracy in case of second model ).Also dropout after convolutional layers and fully connected layers gives less accuracy as compared to dropout after FC layers . Hence dropout should be preferred after FC layers .

Table below shows changes in of test dataset in first model (Table 1) and second model(Table 2) after introducing different parameters(basis) .

BASIS	ACCURACY(%)
Dropout at FC layers	70
Dropout at Convolution and FC layers	40
Alpha Dropout at FC layers	70
Alpha Dropout at Convolution and FC layers	10
Feature Alpha Dropout at FC layers	75
Feature Alpha Dropout at Convolution and FC layers	10
L1 Regularization	21
L2 Regularization	61
Max Norm	70
Data Augmentation	69

**Table 1**

BASIS	ACCURACY(%)
Dropout at FC layers	75
Dropout at Convolution and FC layers	73
Alpha Dropout at FC layers	80
Alpha Dropout at Convolution and FC layers	59
Feature Alpha Dropout at FC layers	78
Feature Alpha Dropout at Convolution and FC layers	50
L1 Regularization	12
L2 Regularization	60
Max Norm	61
Data Augmentation	60

**Table 2**

**Conclusion:**

For solving the major problem of all DL models, the choice of regularizers does matters. Dropout and Data Augmentation are two best regularizers for image classification problem . There are other factors that can also play a major role in increasing the speed of computation of the model like ReLU (activation function), batch normalization , weight initialization , stride ,gradient descent algorithm etc .

### **Learning Accomplishments:**

I started my internship with not having any background from the ML and DL and from that day till now I have learned a lot. I started with learning Linux and then Python programming language .I also learned some tools which Pytorch provides. For doing all these things, my guide, Subhrajit Nag had guided me.He gave me the right way to learn. I also have done some hands-on experience with Pytorch, Numpy and Pandas which I learned from the teaching session taken by Ritesh Gupta and Sharath R . I have learned writing research and survey papers from seminars taken by Dr. Sparsh Mittal. He also gave the Latex and Grammarly for writing such papers. I have also learnt about Convolutional Neural Network (CNN).