

DEEP NEURAL NETWORK AND ITS APPLICATIONS FOR TRAINING DATA SET BY USING NORMALIZATION

Ms. K. VAISHNAVI¹, Ms. R. HARINI², Ms. S.S. SUVALAKSHMI³

¹Assistant Professor, Department of Information Management System,

Mahalashmi Women's College of Arts and Science, Chennai

^{2,3}Research Scholar, Department of Computer Science, Shri Shankarlal Sundarbai Shasun Jain College for Women, Chennai

Abstract: Now-a-days, deep learning plays a vital role in the field of machine learning. Deep learning is derived from the concept of artificial intelligence. The major advantages of using deep learning is effective, supervised, reduced cost and time efficient machine. Deep learning allows various procedures and topographies for the input data to be sent into an algorithm for processing to produce required output. Deep learning is used in various applications and mainly for security purposes. Domains which are running by using deep learning are education, science and technology, medical science in the job of cancer detection, stock market analysis, natural language processing, face recognition and many more. The most complicate process in deep learning is training deep neural networks, because, each and every input layer has to be trained in the manner of using parameters to produce required output. This process may slow the entire process, in the form learning rates and initialization of parameters. By using an appropriate algorithm like Multilayer perceptron neural network (MLPNN) which is forward technique of Artificial Neural Network (ANN) that can be used for back propagation training, convolutional neural network (CNN) which is used for analyzing visual image based on shared-weights architecture and translation invariance characteristics, recurrent neural network (RNN) which develops a connection between the nodes in the directed graph, generative adversarial network (GAN) which acts as a two neural network and used for gaming process, deep belief network (DBN) which helps to develop a graphical model of an input data under several layers of perceptron and many more. Deep learning mechanism allows user to intercept an input to many layers of perception until the required output is obtained. This paper is designed in the process of deriving the super concept of normalization in deep learning architecture for training the input data sets. Normalization is a new technique for training data set and also activating hidden data sets in the neural network. It uses the unbiased technique in gradient variations for input data sets. Gradient variation helps to develop the model in various layered perceptron. This normalization can also be implemented for statistical data analysis. Normalization is also helps to identify the minimum and maximum values among the data sets. If huge number of data sets is processed, then normalization allows batch process to train the data sets.

Keywords: Artificial Intelligence, Machine Learning, Neural Networks, Generative Adversarial Network, Normalization, Gradient Variations.

1. INTRODUCTION:

Deep learning is the subset of machine learning which an advanced technology of the system. Deep learning is designed for automatic progression of knowledge which does not programmed explicitly. The concept of deep learning starts with the patterns and features of the input data set. Deep learning uses advanced algorithms with technical transformations. It takes learning algorithm from the mechanism of machine learning which is classified into four categories. [1] They are supervised learning which produce output for each and every input, unsupervised learning which produce the classification of output rather than exact output of input, semi – supervised which is the combination of supervised and unsupervised learning i.e. produce output for some input data set only and finally reinforcement learning which uses the agent and environment for regression and classification of the trained data set. Time consumption varies for different amount of input data set. As high number of data set is progressed, then levels of perceptron are increased as hidden layers in the deep learning. [2]

The training part of the data set helps to compare the input data set before trained and after trained and the knowledge level of the input data. Deep learning can also hold the large volume of data set. It evolves for feature extraction. The word deep points the hidden layers produced by the input data set. The transformation of input data set to hidden layers consists of Credit Assignment Path (CAP). Deep learning has n number of representation level in its

abstract. Deep learning concept is involved by non – linear process which use hierarchical model for multiple layers. These layers pass the derived output to another layer and so on. [3] This paper proposed by using both supervised and unsupervised learning from the machine learning mechanism. This paper proposed some advanced technical network model which is designed in the concept of deep learning and some advanced application which are used in medical science, science and technology, education field, government functions, and many more. Stochastic Gradient Variation (SGV) proves the effective way for training the input data set. If the network of layers becomes deeper, then the training data set will become complicate because of the simplification of parameters. [4]

2. TECHNICALLY ADVANCED NETWORKS OF DEEP LEARNING:

Neural network provides many flexible network models, in which deep learning is designed with some advanced technical network model for future machine applications. [5]The following explains about the advanced technical network model proposed by deep learning.

ALEXNET – The net is name of a Convolution Neural Network (CNN) which is designed by Alex Krizhevsky. The architecture of AlexNet consists of eight layers in which five for convolutional layers and three are fully – connected layers. AlexNet is high speed for training huge data set. [6]

VISUAL GRAPHIC GROUP NET (VGG) – This network is proposed specially for visual representation and visual recognition. The website of this network uses Google Analytics for website content. The architecture of VGG is developed for evaluate the growth of the network depth using the data items. [7]

GENERATIVE ADVERSARIAL NETWORK (GAN) – This network is the abstract class of machine learning. This network is designed by GoodFellow and his colleagues in 2014. This network is specially designed for gaming process. It consists of two neural networks which contest with each other. [8]

3. DATA NORMALIZATION IN DEEP LEARNING:

Data normalization is used for preparing data set in machine learning. The main goal of data normalization is to change the data values in the column without change in the range value produced by the original dataset. [9] For an example, an input dataset with data in one column ranges from 0 to 10 and data in another column ranges from 100 to 1000. The column data set differs high in the scale of numbers. Modeling is used for combining the feature values of the dataset. Error may occur in the process of modeling because of the huge difference between the column dataset. To sort out the error, normalization is used. [10]

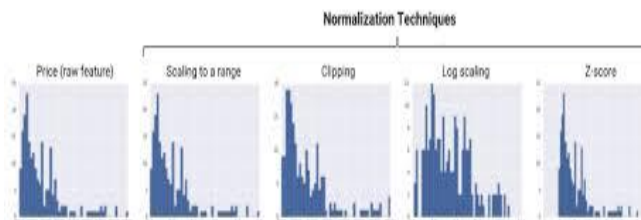


Fig 1 – Various Normalization techniques for Training Dataset

4. VARIOUS APPROACHES OF NORMALIZATION IN DEEP LEARNING:

Normalization is used in deep learning for training dataset by various advanced approaches. They are as follows:

4.1 BATCH NORMALIZATION:

Batch normalization is flexible and feasible among normalization methods used for training dataset under deep learning. It enables faster and stable training of deep neural networks by stabilizing the distributions of layer inputs during the training phase. Internal Covariate Shift (ICS) is used to make the changes in the distribution of inputs processed under hidden layers which is deeply relate with batch normalization approach. [11]

The main advantage of using batch normalization for training data set under deep learning are Batch normalization reduces the internal covariate shift (ICS) and accelerates the training of a deep learning, batch normalization is strongly used for reducing the gradients values in the range of parameters in the input dataset which produces the output under learning mechanism of machine learning without any error under and it also helps in prevention of network from the modes under saturation. [12]

4.2 WEIGHT NORMALIZATION:

Weight normalization under deep learning is processed by using supervised learning from machine learning mechanism. The main use of weight normalization is to repeat the parameterization process. This normalization processed

by uncouple the length of vectors which are weights by its direction. It helps to improve the optimization of weights of the trained dataset under deep learning. [13]

The advantages of using weight normalization for training dataset under deep learning are Weight normalization improves the conditioning of the optimization problem as well as speed up the convergence of SGV. [14]

4.3 LAYER NORMALIZATION:

Layer normalization is the best approach of normalization methods. It is used to improve the speed of the training process of the input dataset from n number of neural networks. Layer normalization is used to process the estimation of statistical report of the summarized input to the output through hidden layer in direct manner. This approach is proposed to make the advanced process of batch normalization. [15]

The advantages of using layer normalization for training dataset under deep learning are Layer normalization can be easily applied to recurrent neural networks by computing the normalization statistics separately. [16]

4.4 GROUP NORMALIZATION:

Group normalization is developed on the basis of batch normalization. It is similar like batch normalization, but it processes the input dataset under the division of layers into groups and then mean and variance of the dataset is computed. Among the normalization approaches, group normalization provides the accuracy range of batch volume which is stable for further processing. [17]

5. NORMALIZE THE DATASET USING KERAS:

It is very important to scale the pixel value of an image much before pass the image under deep learning during the training process of the input dataset as an image. [18]

5.1 IMAGE CLASSIFICATION DATASET USING MNIST HANDWRITTEN:

The main aim of this method is to range the handwritten digit. Value should range from 0 to 9. This dataset is provided as part of the Keras library and can be automatically downloaded (if needed) and loaded into memory by a call to the `keras.datasets.mnist.load_data()` function. For example:

```
1 # example of loading the MNIST dataset
2 from keras.datasets import mnist
3 (x_train, y_train), (x_test, y_test) = mnist.load_data()
```

Fig 2 – Loading the MNIST dataset.

Load the MNIST dataset and summarize the dataset. [19]

```
1 # load and summarize the MNIST dataset
2 from keras.datasets import mnist
3 # load dataset
4 (train_images, train_labels), (test_images, test_labels) = mnist.load_data()
5 # summarize dataset shape
6 print('Train', train_images.shape, train_labels.shape)
7 print('Test', test_images.shape, test_labels.shape)
8 # summarize pixel values
9 print('Train', train_images.min(), train_images.max(), train_images.mean(), train_
10 print('Test', test_images.min(), test_images.max(), test_images.mean(), test_image
```

Fig 3 – Loading Dataset and Summarize the Dataset shape and Pixel values.

```
1 Train (60000, 28, 28) (60000,)
2 Test ((10000, 28, 28), (10000,))
3 Train 0 255 33.318421449829934 78.56748998339798
4 Test 0 255 33.791224489795916 79.17246322228644
```

Fig 4 – Output of the Coding.

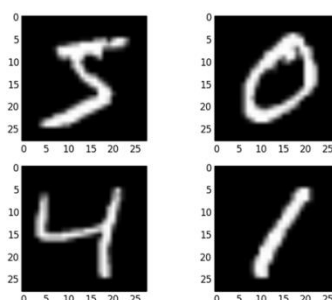


Fig 5 – MNIST Handwritten

5.2 PIXEL SCALING BY IMAGEDATAGENERATOR:

This method is used to scale the pixel values in the input image dataset by modeling process. Load the input dataset and configure the ImageDataGenerator (e.g. construct an instance) and finally calculate image statistics (e.g. call the fit() function). [20]

Techniques of ImageDataGenerator for pixel scaling are given below:

NORMALIZATION OF PIXEL – Fixing the pixel values.

CENTER POSITION OF PIXEL – Computing the mean value of pixel to zero.

STANDARDIZATION OF PIXEL – Computing Variance for pixel values.

There are two step processed for standardizing the pixel values. They are selecting sample for image and fixing feature for image. This process is helps to compute the mean and standard deviation of the pixel values. [21] For example:

```
1 # create and configure the data generator
2 datagen = ImageDataGenerator(...)
```

Fig 6 – Creating Data Generator

After creating the data generator, it is very important to calculate the statistical report of the trained dataset. To calculate the statistical report, fit() function is used.

```
1 # calculate scaling statistics on the training dataset
2 datagen.fit(trainX)
```

Fig 7 – Calculation Statistics on Training Dataset

Once computation of statistical report, data generator of the dataset is to processed by using flow() function.

```
1 # get batch iterator
2 train_iterator = datagen.flow(trainX, trainy)
3 # fit model
4 model.fit_generator(train_iterator, ...)
```

Fig 8 – Code for Fitting Model

If the trained dataset should be validate, fig 9 can be used.

```
1 # get batch iterator for training
2 train_iterator = datagen.flow(trainX, trainy)
3 # get batch iterator for validation
4 val_iterator = datagen.flow(valX, valy)
5 # fit model
6 model.fit_generator(train_iterator, validation_data=val_iterator, ...)
```

Fig 9 – Code for Training and Validating Dataset

It very important to test the batch process and data generator by using evaluate_generator() function. [22]

```
1 # get batch iterator for testing
2 test_iterator = datagen.flow(testX, testy)
3 # evaluate model loss on test dataset
4 loss = model.evaluate_generator(test_iterator, ...)
```

Fig 10 – Code for Testing the Dataset

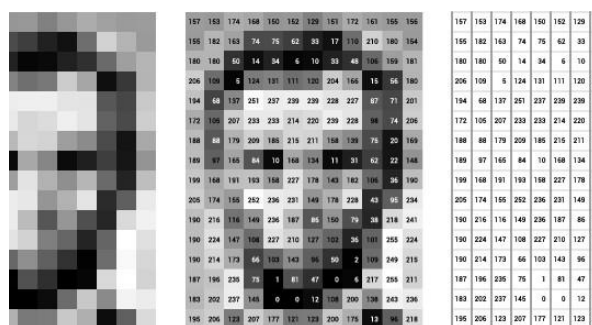


Fig 11 – Keras Scaling Image Pixels for Training

CONCLUSION:

This paper presents about the optimal and feasible mechanism for training the dataset under deep neural network with high speed of process. Time consumption varies for different amount of input data set. As high number of data set is progressed, then levels of perceptron are increased as hidden layers in the deep learning. The training part of the data set helps to compare the input data set before trained and after trained and the knowledge level of the input data. Deep learning can also hold the large volume of data set. It evolves for feature extraction. The word deep points the hidden layers produced by the input data set. The transformation of input data set to hidden layers consists of Credit Assignment



Path (CAP). This paper also proposed the model of using various normalization approaches for training data set under deep learning mechanism. This paper also established about the normalization of datasets under Keras by providing the methods of MNIST handwritten and pixel scaling by ImageDataGenerator.

REFERENCES:

- [1] Raiko, Tapani, Valpola, Harri, and LeCun, Yann. Deep learning made easier by linear transformations in perceptrons. In International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 924–932, 2012.
- [2] Russakovsky, Olga, Deng, Jia, Su, Hao, Krause, Jonathan, Satheesh, Sanjeev, Ma, Sean, Huang, Zhiheng, Karpathy, Andrej, Khosla, Aditya, Bernstein, Michael, Berg, Alexander C., and Fei-Fei, Li. ImageNet Large Scale Visual Recognition Challenge, 2014.
- [3] Saxe, Andrew M., McClelland, James L., and Ganguli, Surya. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. CoRR, abs/1312.6120, 2013.
- [4] Shimodaira, Hidetoshi. Improving predictive inference under covariate shift by weighting the log-likelihood function. Journal of Statistical Planning and Inference, 90 (2):227–244, October 2000.
- [5] Srivastava, Nitish, Hinton, Geoffrey, Krizhevsky, Alex, Sutskever, Ilya, and Salakhutdinov, Ruslan. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res., 15(1):1929–1958, January 2014.
- [6] Sutskever, Ilya, Martens, James, Dahl, George E., and Hinton, Geoffrey E. On the importance of initialization and momentum in deep learning. In ICML (3), volume 28 of JMLR Proceedings, pp. 1139–1147. JMLR.org, 2013.
- [7] Desjardins, Guillaume and Kavukcuoglu, Koray. Natural neural networks. (unpublished). [8] Duchi, John, Hazan, Elad, and Singer, Yoram. Adaptive subgradient methods for online learning and stochastic optimization. J. Mach. Learn. Res., 12:2121–2159, July 2011. ISSN 1532-4435.
- [8] Gulcehre, C, aglar and Bengio, Yoshua. Knowledge matters: Importance of prior information for optimization. CoRR, abs/1301.4083, 2013.
- [9] He, K., Zhang, X., Ren, S., and Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. ArXiv e-prints, February 2015.
- [10] Hyvärinen, A. and Oja, E. Independent component analysis: Algorithms and applications. Neural Netw., 13(4-5): 411–430, May 2000.
- [11] Jiang, Jing. A literature survey on domain adaptation of statistical classifiers, 2008.
- [12] Serizel RGD (2016) Deep-neural network approaches for speech recognition with heterogeneous groups of speakers including children. Nat Lang Eng 1(3):1–26
- [13] Soniya, Paul S, Singh L (2015) A review on advances in deep learning. In: Proceedings of IEEE workshop on computational intelligence: theories, applications and future directions (WCI), pp 1–6. <https://doi.org/10.1109/wci.2015.7495514>
- [14] Sudholt S, Fink GA (2017) Attribute CNNs for word spotting in handwritten documents. Int J Doc Anal Recognit (IJ DAR). <https://doi.org/10.1007/s10032-018-0295-0>
- [15] B (2017) Deep learning in the automotive industry: applications and tools. In: Proceedings of the IEEE international conference on big data, pp 3759–3768
- [16] Makhmudov AZ, Abdurkarimov SS (2016) Speech recognition using deep learning algorithms. In: Proceedings of the international conference on informatics: problems, methodology, technologies, pp 10–15
- [17] Markovnikov N, Kipyatkova I, Karpov A, Filchenkov A (2018) Deep neural networks in russian speech recognition. ArtifIntellNat Lang CommunComputInfSci 789:54–67. https://doi.org/10.1007/978-3-319-71746-3_5
- [18] Mohamed A, Dahl G, Geoffrey H (2009) Deep belief networks for phone recognition. In: Proceedings of the nips workshop on deep learning for speech recognition and related applications, pp 1–9
- [19] Azar MY, Hamey L (2017) Text summarization using unsupervised deep learning. Expert SystAppl 68:93–105
- [20] Chen XW, Lin X (2014) Big data deep learning: challenges and perspectives. IEEE 2:514–525. <https://doi.org/10.1109/ACCESS.2014.2325029>
- [21] Chen CH, Lee CR, Lu WCH (2016) A mobile cloud framework for deep learning and its application to smart car camera. In: Proceedings of the international conference on internet of vehicles, pp 14–25. <https://doi.org/10.1007/978-3-319-51969-22>
- [22] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2018) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans Pattern Anal Mach Intell 40(4):834–848