# Data Classification with Deep Learning using Tensorflow

Fatih Ertam
Informatics Department
Firat University
Elazig, Turkey
fatih.ertam@firat.edu.tr

Galip Aydın
Computer Engineering
Firat University
Elazig, Turkey
gaydin@firat.edu.tr

***Abstract*— Deep learning is a subfield of machine learning which uses artificial neural networks that is inspired by the structure and function of the human brain. Despite being a very new approach, it has become very popular recently. Deep learning has achieved much higher success in many applications where machine learning has been successful at certain rates. In particular It is preferred in the classification of big data sets because it can provide fast and efficient results. In this study, we used Tensorflow, one of the most popular deep learning libraries to classify MNIST dataset, which is frequently used in data analysis studies. Using Tensorflow, which is an open source artificial intelligence library developed by Google, we have studied and compared the effects of multiple activation functions on classification results. The functions used are Rectified Linear Unit (ReLu), Hyperbolic Tangent (tanH), Exponential Linear Unit (eLu), sigmoid, softplus and softsign. In this Study, Convolutional Neural Network (CNN) and SoftMax classifier are used as deep learning artificial neural network. The results show that the most accurate classification rate is obtained using the ReLu activation function.**

***Keywords— Deep Learning; Tensorflow; CNN; ReLu; Softmax Classifier***

## I. INTRODUCTION

Big data is a term that we can use for data sets with large, diverse and complex data structures that are difficult to analyze or visualize using traditional computing methods and approaches [1]. The widespread use of the Internet has caused frequent use of social media and increased data production such as photo, video, text sharing and internet log records, which in turn resulted in the formation of very large data sets called big data.

The term big data does not only express the data collected from the internet sources but also very different types of data produced by various sources such as mobile phones, climate sensors etc [2]. It is not possible to express the term big data and big data anlysis with a single definition [3]. However big data analysis can be defined as providing usability for the data received from various sources through the use of data mining operations. The fact that the size of the data and the time it takes to process it, is another significant aspect of the big data. There are 5 components called 5-V for big data. These components are variety, velocity, volume, verification and value [6].

In addition to the machine learning approaches used in the big data analysis, deep learning approaches are frequently used today. In recent years, studies using deep artificial neural networks have been successful in many competitions [7].

Deep learning, which is becoming more popular every day, can be regarded as a learning technique on artificial neural networks. While classification methods in classical machine learning approaches are considered to be 0 or 1, deep learning can provide numerical results between 0 and 1 as well. Thus, more accurate answers can be obtained for the current problem, and faster and higher accuracy values can be achieved in classification approaches.

By using a CNN in deep learning, a model class can be created to enable powerful and often correct assumptions by changing various parameters [8]. There are several libraries used in deep learning studies.

In this study, classification was made using the Tensorflow library. Tensorflow is an open-source software library developed by the Google for numerical computation, which is now widely used by many large companies. Tensorflow provides an interface for expressing machine learning algorithms and an application for executing these algorithms [9].

A calculation expressed using TensorFlow can be carried out with little or no modification in a wide range of heterogeneous systems, from mobile devices such as phones and tablets, to large scale distributed systems of hundreds of machines, and to various computing devices such as GPU cards.

In this study, the MNIST dataset was used to measure the performance of the Tensorflow library. MNIST is an abbreviation of "Modified National Institute of Standards and Technology" and is a large data set consisting of handwritten numbers widely used for the training of image processing systems [10].

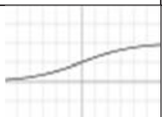
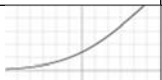
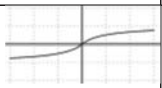
In this dataset, there are 60.000 figures for training and 10.000 figures for the test. Fig-1 shows the numbers written with different handwritings. Each picture consists of 28x28 pixels.

978-1-5386-0930-9/17/$31.00 ©2017 IEEE

Fig. 1. Examples of the MINIST dataset

The correct classification performance of the data set was measured by selecting different activation functions in the software prepared with the Tensorflow library. The graphs and equations for the selected activation functions are given in Table-1.

TABLE I. GRAPHS AND EQUATIONS OF ACTIVATION FUNCTIONS USED

| Activation Function | Plot | Equation |
|---|---|---|
| Rectified linear unit [11] |  | $f(x) = \begin{Bmatrix} 0\ , x < 0 \\ x\ , x \ \geq 0 \end{Bmatrix}$ |
| Exponential linear unit [12] |  | $f(x) = \begin{Bmatrix} \propto (e^{x} - 1)\ , x < 0 \\ x\ , x \ \geq 0 \end{Bmatrix}$ |
| TanH [13] |  | $f(x) = \frac{2}{1 + e^{-2x}} - 1$ |
| Sigmoid[13] |  | $f(x) = \frac{1}{1 + e^{-x}}$ |
| SoftPlus [13] |  | $f(x) = \ln(1 + e^{x})$ |
| Softsign[14] |  | $f(x) = \frac{x}{1 + \lvert x \rvert}$ |

Softmax classifier is used to classify the data set. Softmax is a classifier and function that is often used in deep learning applications [15]. With Softmax, the data can be categorized by direct classifier.

In the first part of this paper, big data, deep learning, TensorFlow, MNIST data set, activation functions used and general information about softmax classifier are given. In the second section, experimental work is described. The final section contains the results and the discussion.

## II. EXPERIMENTAL STUDIES

The flow diagram for the prepared system is given in Figure-2. The images contained in the MNIST data set are stored in the IDX file format. This is a simple format for vectors and multidimensional matrices of various digital types. There are 4 files in the dataset, including image and label information for testing and training.

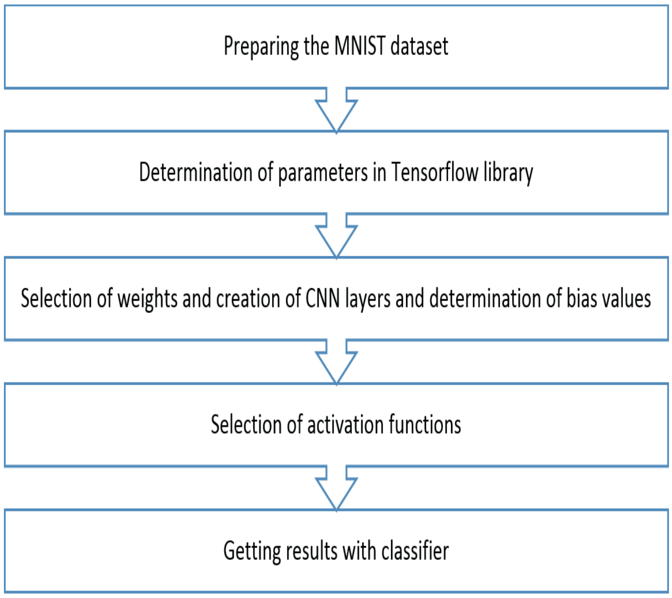



Fig. 2. Proposed System

The parameters for the Tensorflow Library for deep learning in the second phase are specified. For this study, 100,000 iterations were processed. Since each image is composed of 28x28 pixels, the network input parameter value is taken as 784. Output class is selected as 10, because the digits are expected to be within digits from 0 to 9.

In determining the weights in the third stage, 5x5 convolutional layer and 32 outputs were selected. Then a 5x5 convolutional layer and 64 outputs were applied again. In the fully connected layer, values of 3136 and 1024 are selected as input and output, respectively. In the output layer, the output value in the previous layer is 1024, which is input and taken as 10 outputs with the total number of classes. Bias values are taken as output values in the weight layers.

In the fourth step, the performance of the classifier was measured during the selected iteration by selecting different activation functions. Table-2 shows the accuracy rates obtained from using different activation functions.

TABLE II. CLASSIFIER PERFORMANCE ACCORDING TO DIFFERENT ACTIVATION FUNCTIONS

| Activation Function | Testing Accuracy |
|---|---|
| ReLu | 0.9843 |
| TanH | 0.7591 |
| eLu | 0.9687 |
| Sigmoid | 0.7851 |
| SoftPlus | 0.9765 |
| SoftSign | 0,7773 |

The accuracy values obtained according to the iteration numbers of the ReLu activation function obtained as the best result are given in Figure-3. The graphs for the other activation functions are given in Figures 4-8.

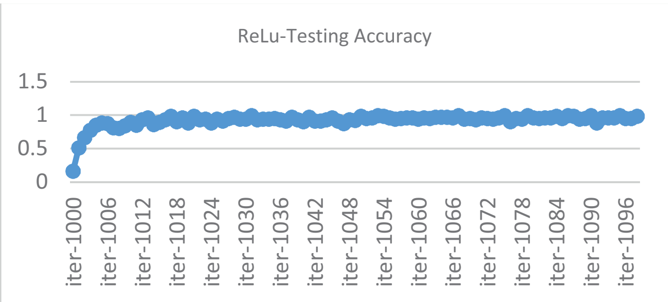



Fig. 3. ReLu activation function Test accuracy values

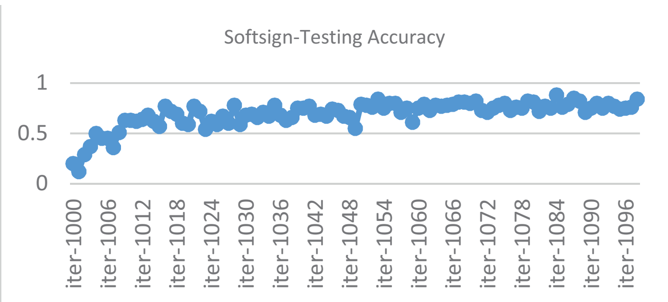



Fig. 4. Softsign activation function Test accuracy values

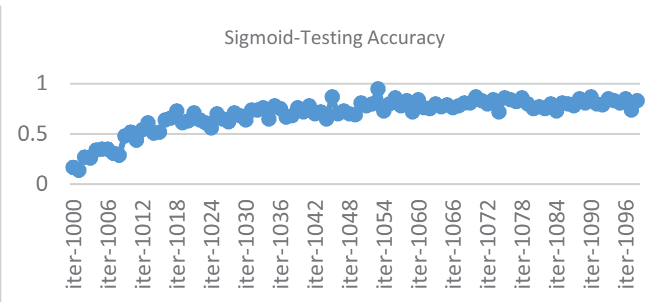



Fig. 5. Sigmoid activation function Test accuracy values

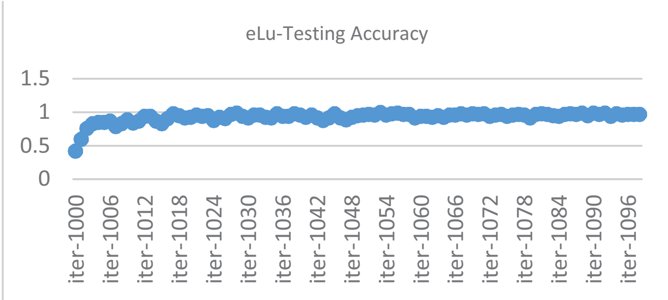



Fig. 6. eLu activation function Test accuracy values

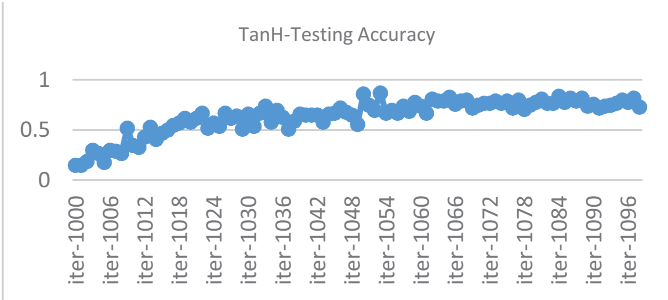



Fig. 7. TanH activation function Test accuracy values

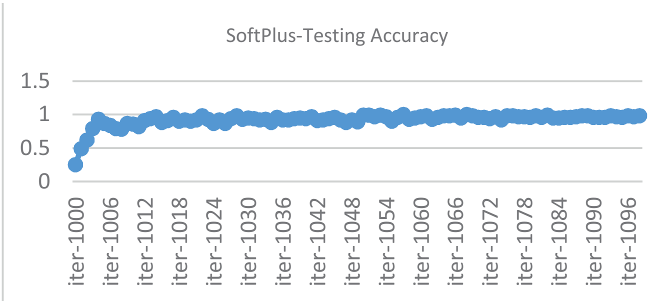



Fig. 8. SoftPlus activation function Test accuracy values

## III. RESULTS AND DISCUSSION

Deep learning approaches are increasing in their popularity every day. Deep learning provides fast and effective solutions especially in the analysis of big data. In this study, a classification task was carried out on the MNIST data set which is widely used in deep learning applications. Tensorflow was used for this purpose.

Different activation functions were selected in the system to test the accuracy of the classification of the system. ReLu, eLu, tanH, sigmoid, softPlus and softsign activation functions were used for this purpose. SoftMax is used as the classifier function. It has been observed that the best accuracy is achieved when the ReLu activation function is selected. In the study conducted with ReLu activation function, 98.43% classification accuracy is obtained on the test data.

The increase in the number of iterations showed an increase in the accuracy values, but the total classification time is also increased. In subsequent studies we aim to increase accuracy by applying different neural network architectures.

## REFERENCES


[1] S. Sagiroglu and D. Sinanc, "Big data: A review," in 2013 International Conference on Collaboration Technologies and Systems (CTS), 2013, pp. 42–47.

[2] C. Snijders, U. Matzat, and U. Reips, "'Big Data': Big Gaps of Knowledge in the Field of Internet Science," Int. J. Internet Sci., vol. 7, no. 1, pp. 1–5, 2012.

[3] V. Mayer-Schönberger and K. Cukier, Big Data: A Revolution That Big will Transform How we Live, Work and Think, vol. 26. 2013.

[4] I. S. Rubinstein, "Big Data: The End of Privacy or a New Beginning?," Int. Data Priv. Law, vol. 3, no. 2, pp. 74–87, 2013.

[5] M. a. Beyer and D. Laney, "The Importance of 'Big Data': A Definition," 2012.

[6] M. M. Gobble, "Big Data: The Next Big Thing in Innovation.," Res. Technol. Manag., vol. 56, no. 1, pp. 64–66, 2013.

[7] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85–117, Jan. 2015.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process. Syst., pp. 1–9, 2012.

[9] M. Abadi et al., "TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), 2016, pp. 265–284.

[10] Y. Lecun, C. Cortes, and C. J. C. Burges, "The MNIST Database," Courant Institute, NYU, 2014. [Online]. Available: http://yan.lecun.com/exdb/mnist.

[11] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," Proc. 27th Int. Conf. Mach. Learn., no. 3, pp. 807–814, 2010.

[12] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)," pp. 1–14, 2015.

[13] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," AISTATS '11 Proc. 14th Int. Conf. Artif. Intell. Stat., vol. 15, pp. 315–323, 2011.

[14] P. Le and W. Zuidema, "Compositional Distributional Semantics with Long Short Term Memory," arXiv1503.02510 [cs], 2015.

[15] R. . Salakhutdinov and G. . Hinton, "Replicated softmax: An undirected topic model," Adv. Neural Inf. Process. Syst. 22 - Proc. 2009 Conf., pp. 1607–1614, 2009.