A Model of Deep Neural Network for Iris Classification With Different Activation Functions

Ayşe ELDEM
Department of Computer Engineering
Karamanoğlu Mehmetbey University
Karaman / TURKEY
ayseeldem@kmu.edu.tr

Hüseyin ELDEM
Department of Technology and Computer Programming
Karamanoğlu Mehmetbey University
Karaman / TURKEY
heldem@kmu.edu.tr

Deniz ÜSTÜN
Department of Computer Engineering
Karamanoğlu Mehmetbey University
Karaman / TURKEY
denizustun@kmu.edu.tr

***Abstract*—In recent years, deep neural network (DNN) has been frequently used for classification. In this study, iris flowers having 3 different types are classified by using DNN which are utilized the width and length of petal and sepal features as input. Some experiments are made for the iris dataset by using different activation functions and different epoch numbers. Then, the activation function which gives the best result is determined. The classification success of the developed model is achieved 96% for the iris dataset.**

***Index Terms*— Classification, Deep Learning, Deep Neural Network, İris**

I. INTRODUCTION

Deep learning has begun to be use from the beginning of 2010's. Learning and interpretation with DNN is better than the other methods. Deep learning algorithms are used in many areas such as classification, recognition, interpretation, image analysis, sound analysis, gene analysis, time based operations, medical, etc. Altough machine learning algorithms pre-process the data, DNN algorithms allows you to manipulate raw data [1, 2]. DNN algorithms can solve complex problems at high speed and high performance using by hidden layers.

There are different types of classification in the literature [3, 4]. Basically, a system is trained with a train data in the classification. Then, the system determines the classes of test data that have not been used before. For example; it is possible to classify by using width and length of sepal and petal for iris flowers. Many studies have been done iris classification until today. The iris flowers were classified by using the neuro-fuzzy classification method [5], a hybrid neuro-fuzzy classifier based on nefclass model [6], k-means algorithm and neural network [7], Bayesian classifier which run as regression algorithm [8], ν-K-SVCR method [9], decision trees [10], k nearest neighbor algorithm on feature projections [11], structural learning with forgetting (SLF) network [12].

In this study, a deep neural network model is developed for the classification of iris dataset taken from UCI Machine Learning Respository. There are 3 different types for iris called iris-setosa, iris-versicolour and iris-virginica. Many experiments is done for the model with different activation functions and different epoch numbers.

II. MATERIAL AND METHOD

A. Iris Flowers

Iris flowers are based in colors such as blue, yellow, purple, pink, orange and white in nature. So they are called as rainbow. Irises are grown in many areas like deserts, swamps.

B. Dataset

Iris dataset is taken from UCI Machine Learning Repository [13]. It was introduced by Ronald Fisher in 1936 [14].

There are 4 input parametres in the dataset. These parameters are sepal length, sepal width, petal length and petal width. Sepal and petal sizes are shown in Fig. 1.

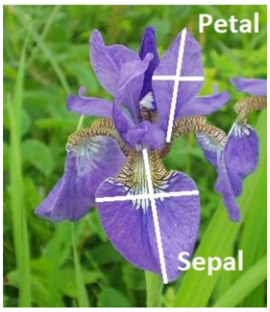

Figure 1. Sepal and Petal Sizes

There are three class for iris flower like iris setosa, iris versicolour and iris virginica. Each group consists of 50 data. There are 4 input parameters and 1 output parameters.

978-1-5386-6878-8/18/$31.00 ©2018 IEEE

Detailed information about dataset is given in Table I and Table II. In Table I, the type of iris and size of sepal and petal are shown, the minimum and maximum values for each parametres are shown in Table II.

TABLE I. DATASET INFORMATION

Sepal Length (SL)	Sepal Width (SW)	Petal Length (PL)	Petal Width (PW)	Type
5.1	3.5	1.4	0.2	Iris-setosa (1)
4.9	3.0	1.4	0.2	Iris-setosa (1)
…	…	…	…	…
7.0	3.2	4.7	1.4	Iris-versicolour (2)
6.4	3.2	4.5	1.5	Iris-versicolour (2)
…	…	…	…	…
6.2	3.4	5.4	2.3	Iris-virginica (3)
5.9	3.0	5.1	1.8	Iris-virginica (3)

TABLE II. ATTRIBUTE INFORMATION

Attribute	General Informaiton	
	Max	*Min*
SL	4.3	7.9
SW	2.0	4.4
PL	1.0	6.9
PW	0.1	2.5

C. Application

A model is developed for the classification of iris flowers by using deep neural network that is shown in Fig. 2. 2 hidden layers are used which have 50 and 20 neurons (dense). In this study, 80% of the iris dataset is allocated as training data and the remaining 20% of dataset as the test data. Keras library was used and the success rates of different activation functions of different epoch numbers were examined. Tanh, relu, sigmoid activation functions are used. The graphics of activation functions are shown in Fig. 3.

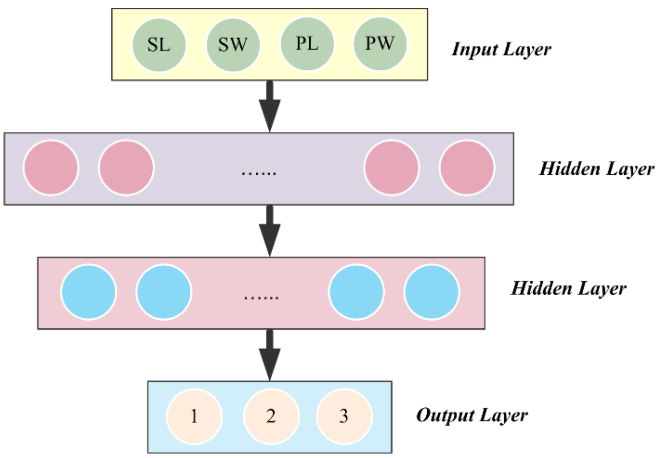

Figure 2. The Architecture of the Model

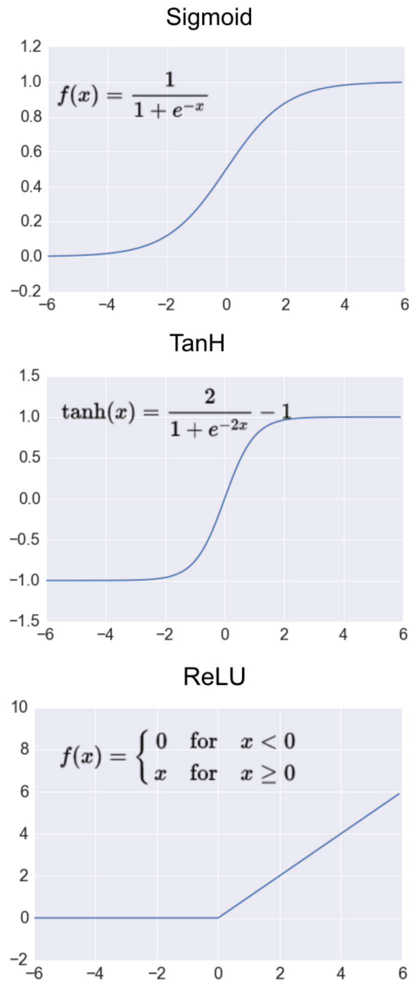

Figure 3. Activation Functions [15]

Sigmoid and tanh are non-linear functions. As seen in Fig.3, output value range is between 0 and 1 for sigmoid activation function. It is between -1 and 1 for tanh and between 0 and x for Relu activation function. Relu is faster than sigmoid and tanh. Because its mathematical operations are similar.

Iris dataset has 3 type of iris flower. Iris-setosa is labeled as '1', Iris-versicolour is labeled as '2' and Iris-virginica is labeled as '3'. Categorical_crossentropy is chosen as the loss function which is used for multiclass classification.

As shown in Table III, Table IV and Table V, different activation functions are used with different epoch numbers in the 1^{st} and 2^{nd} hidden layers and the success rates are examined.

As shown in Table III, relu activation function in the first hidden layer and sigmoid, tanh and relu in the second hidden layer are used. It is observed to have the same accuracy rates for different epoch numbers. The best classification accuracy rate is observed in Table IV. The best result is obtained by using sigmoid activation function in the 1st hidden layer, tanh activation function in the second hidden layer and 300 epoch.

TABLE III. RELU ACTIVATION FUNCTION IN 1ST HIDDEN LAYER

1st Hidden Layer	2nd Hidden Layer	Epoch Number	Accuracy Rate
Relu	Sigmoid	100	86%
		200	86%
		300	86%
		400	86%
Relu	Tanh	100	90%
		200	90%
		300	90%
		400	90%
Relu	Relu	100	90%
		200	90%
		300	90%
		400	90%

TABLE IV. SIGMOID ACTIVATION FUNCTION IN 1ST HIDDEN LAYER

1st Hidden Layer	2nd Hidden Layer	Epoch Number	Accuracy Rate
Sigmoid	Relu	100	93%
		200	93%
		300	93%
		400	90%
Sigmoid	Tanh	100	93%
		200	93%
		300	96%
		400	90%
Sigmoid	Sigmoid	100	90%
		200	90%
		300	86%
		400	86%

TABLE V. TANH ACTIVATION FUNCTION IN 1ST HIDDEN LAYER

1st Hidden Layer	2nd Hidden Layer	Epoch Number	Accuracy Rate
Tanh	Relu	100	93%
		200	90%
		300	90%
		400	86%
Tanh	Sigmoid	100	90%
		200	86%
		300	86%
		400	86%
Tanh	Tanh	100	93%
		200	86%
		300	86%
		400	86%

The used workstation features are Intel Xeon E5-1620 v4 3.5 GHz processor, 64GB RAM and NVIDIA® Quadro® M4000 8 GByte graphics card.

The classification results for the test data is shown Table VI as detailed. In test data, there are 12 iris setosa, 10 iris versicolor and 8 iris virginica samples. All iris setosa and iris versicolor samples are classified correctly. But, only one sample of iris virginica is classified as wrong. In Table VII, the accuracy rate as percent are given for each type of iris.

TABLE VI. RESULTS OF CLASSIFICATION

Test Data	Class		
	Iris Setosa	***Iris Versicolor***	***Iris Virginica***
Number of Classification	12	0	0
	0	10	0
	0	1	7

TABLE VII. ACCURACY RATE FOR THE IRIS TYPES

Type	Accuracy Rate (%)
Iris Setosa	100%
Iris Versicolor	100%
Iris Virginica	87,5%

In the developed model, the behavior of different activation functions and different epoch numbers are examined for iris dataset. As a result of this classification, the model is achieved a 96% success rate with sigmoid, tanh and 300 epoch.

III. RESULTS

In this study, a novel deep neural network model is proposed for the classification of the iris dataset. The dataset is divided into 80% train data and 20% test data. Different experiments have been carried out to make determination of the activation function and epoch number. Different activation functions are used with different epoch numbers. The best result is obtained with sigmoid in the first hidden layer, tanh activation function in the second hidden layer and 300 epoch numbers. The classification success rate of the developed model is 96%.

ACKNOWLEDGMENT

The authors thank to the department of Computer Engineering in Karamanoğlu Mehmetbey University for providing high performance computer used in training and testing process of classification.

REFERENCES

[1] Ö. İnik and E. Ülker, "Deep Learning and Deep Learning Models Used in Image Analysis," Gaziosmanpasa Journal of Scientific Research, vol. 6(3), 2017, pp. 85-104.

[2] A. Kamilaris, and F.X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, 2018, pp. 70-90.

[3] C. Coşkun and A. Baykal, "An Application for Comparison of Data Mining Classification Algorithms," XIII. Akademik Bilişim Konferansı, 2011, pp. 51-58.

[4] K. Yıldız, Y. Çamurcu, and B. Doğan, "A Comperative Analize of Principal Component Analysis and Non-Negative Matrix Factorization Techniques in Data Mining," XII. Akademik Bilişim Konferansı, 2010, pp. 207-213.

[5] V. Arya and R.K. Rathy, "An efficient Neuro-Fuzzy Approach for classification of Iris Dataset," 2014 International Conference on Reliability Optimization and Information Technology (ICROIT), 2014, pp. 161-165.

[6] B. Gliwa and A. Byrski, "Hybrid Neuro-Fuzzy Classifier Based on NEFCLASS Model," Computer Science, vol. 12, 2011, pp. 115-135.

[7] V. Poojitha, M. Bhadauria, S. Jain and A. Garg, "A collocation of IRIS flower using neural network clustering tool in MATLAB," 6th International Conference - Cloud System and Big Data Engineering (Confluence), 2016, pp. 53-58.

[8] S. B. Kotsiantis and P. E. Pintelas, "Logitboost of Simple Bayesian Classifier," Informatica (Slovenia), vol. 29, 2005, pp. 53-60.

[9] P. Zhong and M. Fukushima, "Regularized nonsmooth Newton method for multi-class support vector machines," Optimization Methods and Software, vol. 22(1), 2007, pp. 225-236.

[10] L. O. Hall, N. Chawla, and K.W. Bowyer, "Decision tree learning on very large data sets," Systems, Man, and Cybernetics, IEEE International Conference, 1998, pp. 2579-2584.

[11] H. A. Güvenir and A. Akkus, "Weighted k Nearest Neighbor Classification On Feature Projections," 1998.

[12] W. Duch, R. Adamczak and K. Grabczewski, "Extraction of crisp logical rules using constructive constrained backpropagation networks," Neural Networks International Conference, 1997, pp. 2384-2389.

[13] UCI, Iris Dataset, 30.03.2018, Available from: https://archive.ics.uci.edu/ml/datasets/iris.

[14] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics, vol. 7(7), 1936, pp. 179-188.

[15] https://www.kdnuggets.com/2017/09/neural-network-foundations-explained-activation-function.html, 01.06.2018