Neural Network Model for Heart Disease Diagnosis

Seda Postalcıoğlu
Computer Engineering
Bolu Abant Izzet Baysal University
Bolu, Turkey
postalcioglu_s@ibu.edu.tr

Abstract— the heart disease diagnosis system is proposed in this study. This kind of diagnosis systems enhance medical care and helps doctors. In this paper, heart disease dataset from kaggle web site is used. Neural Network is examined and analyzed for different structures as an optimizer, loss function, and batch size. The simulation results show that the proposed neural network model has 90,16% accuracy.

Keywords—neural network, optimizer, loss function

I. INTRODUCTION

Many systems are proposed for the diagnosis of diseases. There is a need to develop diagnosis systems because it helps medical experts. This kind of diagnosis systems enhance medical care [1]. Figure 1 shows factors that have led to deaths [2]. Heart diseases are the main cause of death according to Figure 1. For this reason, the diagnosis of heart disease is one of the important issues.
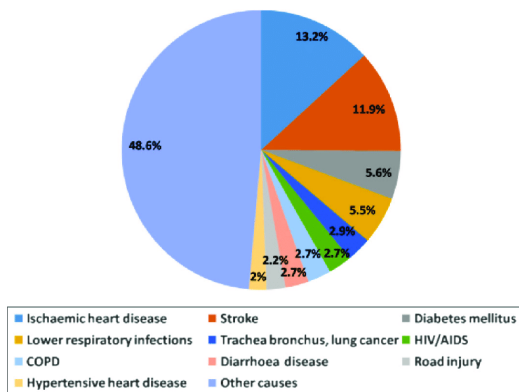

Fig. 1. Factors that have led to the deaths [2]

Neural Network is very popular for modeling. For example, MLPNN is preferred to classify the web sites in the reference 3 [3]. Anomaly detection in crowded videos is examined using neraul network in reference 4 [4]. Comprehensive area of applicability of Neural Network (NN) and its ability to learn complicated and nonlinear relationships, neural networks are used in biomedical engineering in reference 5 [5]. The single layer NN solves only linearly separable problems. Most of the comprehensive problems are nonlinear. To deal with this kind of problems one or more layers are added and known as a multilayer perceptron neural network (MLPNN) [6].

This paper includes the diagnosis of heart disease using MLPNN. This kind of diagnosis system helps the doctor for making decisions about the disease. The benefit of this kind of structure is that it forecasts the disease in a less amount of time [7].

II. MULTILAYER PERCEPTRON NEURAL NETWORK (MLPNN)

The configuration of the MLPNN is seen in Figure 2. During the learning stage, the network learns by adapting the weights to find the correct class [5].
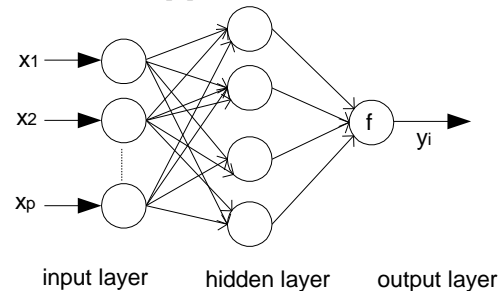

Fig. 2. MLPNN Model [8]

The computation of MLPNN is given in Equation (1-2) [1].

$$V_j = \sum_{i=1}^{P} W_{ji}X_i + \theta_j \qquad (1)$$

$$y_i = f_j(V_j) \qquad (2)$$

V_j is the linear combination of inputs $x_1, x_2 \ldots x_p$, Θ_j the bias, W_{ji} the connection weight between the input X_i and the neuron j, $f_j()$ the activation function, y_j is the output [1]. The weights of the neural network are updated using Back Propagation (BP) algorithm and according to the equation (3) [1].

$$W_{ji}(t+1) = W_{ji}(t) - \varepsilon \frac{\partial E_f}{\partial w_{ji}}(t) \qquad (3)$$

ε is the learning rate, E_f is the error function [1]. The performance of the system is analyzed. The data is obtained from kaggle web site. Heart disease diagnosis system with MLPNN denotes whether the heart disease is existing or absent.

III. NEURAL NETWORK MODEL FOR HEART DISEASE DIAGNOSIS SYSTEM

In this paper heart disease dataset from kaggle web site is used [9]. It contains 303 patient records. Each record contains 14 attributes. A block diagram of the MLPNN based heart disease diagnosis is seen in Figure 3. The advantage of this system is which detects the disease in a short time.
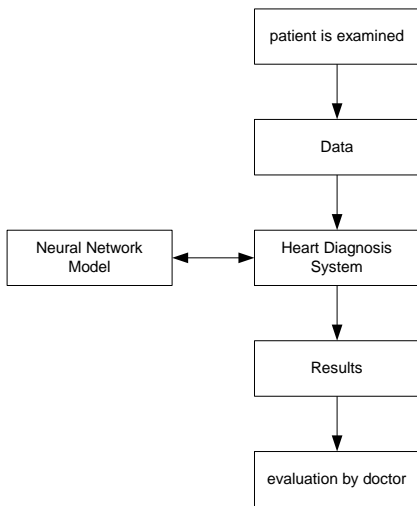
Fig. 3. Block diagram of the heart disease diagnosis system

Figure 4 shows the correlation between variables. It is clear that every metric is perfectly correlated with itself. There's not a lot of orange squares in the chart. Hence, a single value not gives heart disease results.
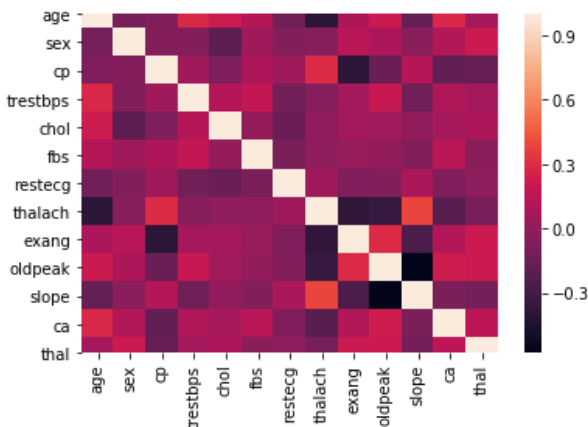

Fig. 4. Correlation between variables

The Rectified Linear Unit activation function and Sigmoid function are used in the first two layers and in the output layer, respectively. Figure 5 and Figure 6 shows the accuracy and loss results for the model, respectively. The blue line shows the training result, the black line shows the test results for Figures 5-6.
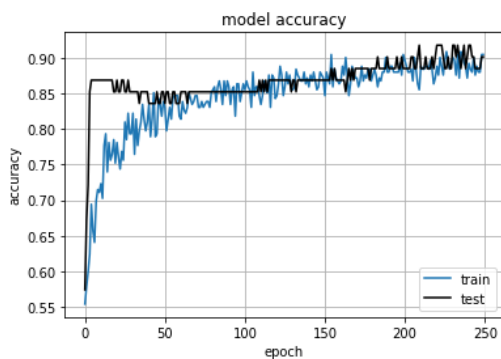

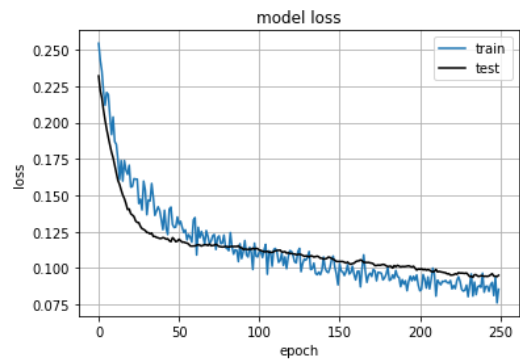Fig. 5. Accuracy result for the model

Fig. 6. The accuracy and the loss results for the model

Figure 7 shows the confusion matrix. The confusion matrix contains the four entries, TP (true positive), FP (false positive), FN (false negative), TN (true negative). The model accuracy is obtained as 90,16%.
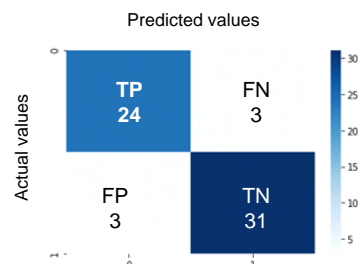

Fig. 7. The confusion matrix

Table 1 gives the classification efficiency according to the samples and layers. It is clear that MLPNN's performance is better according to the Single Layer Neural Network (SLNN). The reason of high accuracy is large number of samples are used to training. The accuracy is 82.42% when using 70% of data for testing the system for MLPNN.

TABLE I. CLASSIFICATION EFFICIENCY FOR THE SYSTEM

Training samples	Test samples	Multi Layer (Accuracy)	Single layer (Accuracy)
242	61	90.16%	86,89%
227	76	86.84%	84,21%
212	91	82,42%	80,22%

Loss function and optimizers are determined for the construction of MLPNN. The loss function is used to calculate a set of weights. To search through different weights for the network, the optimizer is used [10]. Cross Entropy and Mean Square Error (MSE) are used as a loss function. The average of the squared differences is calculated by MSE between the predicted and actual values [11]. A score gives the average difference between the actual and predicted probability distributions is done by Cross Entropy [11]. To make predictions as correct as possible optimizers are used. RMSprop only accumulates gradients in a fixed window. Adam stands for adaptive moment estimation, it is using past gradients to calculate current gradients. SGD calculates on a small subset or random selection of data examples [12]. There is a study about the optimizers performance examined by Postalcioglu [13].

Batch size and dropout techniques are important and effective for construction to model. The batch size is a

hyperparameter of gradient descent and it controls the number of training samples [14]. At the end of the batch, an error is calculated. Batch gradient descent technique is used to improve the model by this error [14]. There are some studies about batch size in the references [15,16]. NN includes multiple non-linear hidden layers and learns very difficult relationships between their inputs and outputs. Limited training data causes to overfitting and many techniques have been developed for reducing it [17]. The choice of which units to drop is random as a known dropout technique. Dropout is an algorithm for training neural networks [18].

In this study, optimizer, loss function, batch size parameters are examined and their effects for the accuracy is given in Table II.

TABLE II. THE PERFORMANCE RESULTS

Optimizer	Loss function	Batch size	Accuracy %
Rmsprop	Binary cross entropy	1	85,25
		32	86,89
		64	90,16
		128	90,16
	Mean square error	64	90,16
Adam	Binary cross entropy	64	86,89
	Mean square error	64	88,52
SGD	Binary cross entropy	64	88,52
	Mean square error	64	80,33

Table 2 shows that the best accuracy is obtained using RMSprop. Accuracy is obtained as 90,16% for the heart diagnosis system.

IV. CONCLUSION

The heart disease diagnosis system is proposed in this study. Using this kind of system helps the doctors to a diagnosis of the disease. Firstly the system detects the disease, then the doctor evaluates the results. Evaluation of patients is done according to the system result. In this study, NN is used. Different optimizers, loss function, and batch size are used and analyzed. The simulation results show that the proposed NN model has 90,16% accuracy for the heart disease diagnosis system.

REFERENCES

[1] T.T. Hasan, M.H. Jasim, and , I.A., Hashim , "Heart Disease Diagnosis System based on Multi-Layer Perceptron neural network and Support Vector Machine", International Journal of Current Engineering and Technology, Vol.7, No.5,pp.1842-1853, 2017.

[2] A. Farah, "Achieving sustainable cultivation of coffee", Burleigh Dodds Science Publishing Limited, Chapter taken from: Lashermes, P. (ed.), Achieving sustainable cultivation of coffee, Burleigh Dodds Science Publishing, Cambridge, UK ,ISBN: 978 1 78676 152 1, 2018.

[3] Ş. Kayıkçı, E .Akyazı, "Classification of Open Directory Web Pages Using Artificial Neural Networks", International Journal of Scientific and Technological Research, ISSN 2422-8702 (Online) Vol 4, No.9,pp. 112-119, 2018, https://doi.org/10.7176/JSTR.

[4] A. Ratre, "Taylor series based compressive approach and Firefly support vector neural network for tracking and anomaly detection in crowded videos", Journal of Eng.Research Vol.7 No 4 December 2019, pp.115-137.

[5] K.U. Rani, "Analysıs of heart dıseases dataset Usıng Neural Network Approach", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.1, No.5, pp.1-8, 2011. DOI : 10.5121/ijdkp.2011.1501

[6] J. S. Sonawane, V. Patil, "Prediction of heart disease using multilayer perceptron neural network", International Conference on Information Communication and Embedded Systems (ICICES2014), Chennai, pp. 1-6, 2014. doi: 10.1109/ICICES.2014.7033860

[7] K. Subhadra, B. Vikas, "Neural Network Based Intelligent System for Predicting Heart Disease", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-5 March,pp.484-487, 2019.

[8] M.M. Kirmani, "Heart Disease Prediction using Multilayer Perceptron Algorithm", International Journal of Advanced Research in Computer Science, Volume 8, No. 5, pp.1169-1172, 2017.

[9] UCI, https://www.kaggle.com/ronitf/heart-disease-uci, Access date: 2019.

[10] J. Brownlee, web site: https://machinelearningmastery.com/tutorial-first-neural-network- python-keras/ accesed on 2019.

[11] J. Brownlee,https://machinelearningmastery.com/how-to-choose-loss-functions-when-training-deep-learning-neural-networks/ accesed on 2019.

[12] T. Peng, web site: https://syncedreview.com/2019/03/07/iclr-2019-fast-as-adam-good-as-sgd-new-optimizer-has-both/ accessed on 2019.

[13] S. Postalcıoğlu, "Performance Analysis of Different Optimizers for Deep Learning based Image Recognition", International Journal of Pattern Recognition and Artificial Intelligence Vol. 34, No. 2 2020, pp:1-12.

[14] J. Brownlee, https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/ accesed on 2018

[15] L. Balles, J. Romero, P. Hennig, "Coupling adaptive batch sizes with learning rates", 2016. arXiv preprint arXiv:1612.05086.

[16] S.L. Smith, P.J. Kindermans, Q.V. Le, "Don't decay the learning rate, increasethe batch size", 2017, arXiv preprint arXiv:1711.00489.

[17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever , R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research 15, pp.1929-1958, 2014.

[18] P. Baldi,, P.J. Sadowski, "Understanding dropout", Proceedings of the NIPS, pp.1-9, 2013.