Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 132 (2018) 437–443

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2018)

Automated Classification Method for Early Diagnosis of Alopecia Using Machine Learning

Ishita Kapoor,Anju Mishra

Amity University, Sector-125, Noida (U.P.), India - 201313

Abstract

Artificial neural networks are being used in many fields for different purposes. Medical diagnosis is one of the major purposes. In the field of medical, classification plays an important role as the main aim of the doctor is to classify whether a person is suffering from the disease or not. The objective of this paper is to evaluate neural network for the detection of alopecia in human beings and to find the accuracy. With the help of proposed model, the clinical experts will be able to get a second opinion that will help them to take proper decisions for diagnosing the presence of this disease in patients. This second opinion is crucial due to many factors while doing disease identification. These factors include increased population, environmental pollution, growing demands for proper medication, and less availability of medical experts to cope up with this increasing demand. Also, the dynamic nature of disease symptoms plays an important role in correct diagnosis of a certain disease. The proposed system uses a feedforward artificial neural network and backpropagation algorithm to classify patients with alopecia and without alopecia. The evaluation of the proposed system is done with the help of performance plot as well as regression plot. Experimental results show that the accuracy of proposed system is 91% which is reliable enough for a clinical expert to make his decisions.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2018).

Keywords: Alopecia; Neural Network; Performance plot;Disease classification; Hairfall.

1. Introduction

Disease diagnosis is one of the active research topic in today's world because medical is one of the areas where researchers are continuously making efforts to develop automated systems for disease diagnosis. These diagnostic systems are not to replace expert clinicians rather to provide them second opinion while deciding for the presence of a certain disease in patient. This second opinion can be crucial in situations when there is lack of proper medical support staff. Also, ever increasing population is also a major factor to impact the quality of medical assistance.

*Corresponding Author: ishitakapoor11@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2018).
10.1016/j.procs.2018.05.157

These factors contribute to high demands of such systems to overcome the pressure on medical staff. With the advancement of technology disease can be cured or diagnosed with the help of different methodology.

Artificial neural network is very helpful in developing such systems. With their powerful non-linear nature,they are very useful for solving complicated problems which are otherwise difficult to solve with deterministic techniques. [1] Majority of the problems which can be solved by using artificial neural network fall under classification category. As in the field of medical, a patient is classified as either affected with the disease or unaffected with the disease (healthy). Artificial neural networks are best suited for this type of problem category. Previous studies and experiments have proven that there were serious performance issues with these automated systems. But recent technological development helped in overcoming these issues and providing for more accurate systems to diagnose a certain disease. Nevertheless, the developed systems for disease diagnosis have limitations as e.g., lack of generalization i.e., a disease diagnosis system can only work upon a specific disease. Therefore, to observe the same disease in different people, a dataset is usually maintained to facilitate further analysis of disease symptoms and patients' history to prescribe medicines accordingly.

Due to growing population, hard timelines and increased pressure on experts, it is a real challenge for experts to detect the disease timely and accurately under such pressure.[2]In case of hair fall, alopecia is sometime being confused with a temporary imbalance of hormones or due to stressful conditions, thyroid, or other hormonal issues, leading to wrong medications which can worsen the patient's condition. To avoid such misinterpretations of disease symptoms and false diagnosis, an automated system can be of a great help.

Alopecia is an auto-immune disorder that concludes in hair loss on the scalp and other parts of the body.[3] It can range from small, individual smooth patches of hair loss to total loss of all hair of the body. It can impact individuals of all age groups, sexes and ethnicity,and generally affects amid adolescence. Figure 1 shows the effect of alopecia on the scalp.

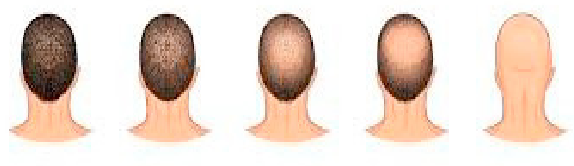

Figure 1: Effect of alopecia on scalp

There are three types of alopecia as shown in figure 2 namely, *Alopecia Areata* in which two to three patches of hair loss on the scalp. *Alopecia Areata Totalis*, in which there is complete baldness on the scalp. *Alopecia Areata Universalis* in which there is complete baldness on the body. It has been shown by previous studies that patients with A*lopecia Totalis* or *Alopecia Universalis* more often are affected by a poorer forecast, and wrong treatment.[2,3,5,8]

Figure 2: Types of alopecia

The impact of alopecia includes a high probability of unconstrained hair reduction. The more drawn out the time frame of hair loss and the bigger the region of scalp gets affected by this disease.Consequently, there are some medications available to treat this disease, but none of these is effective after delayed diagnosis of alopecia.

2. Neural Network

Neural network is a computational tool made up of interconnecting components called artificial neurons which process data dynamically and generates output corresponding to external inputs. It is a network of neurons operating in parallel and encouraged by biological nervous system. The type of connections between these neurons determines the neural function. The whole system is composed of multiple neurons arranged in columns. Each column is called a layer in the network. The first layer corresponds to input layer, and the last layer corresponds to output layer and in the middle of these two layers lie one or more hidden layers. Figure.3 shows a typical neural network topology.

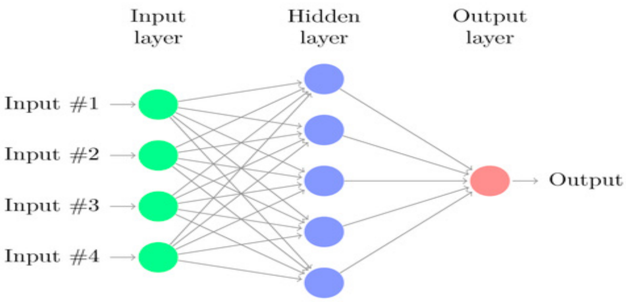

Figure 3: Typical neural network typology

Neural network is ideal for recognition and classification tasks. It learns with the help of presented examples.There are two types of learning paradigms for artificial neural networks: Supervised and unsupervised.

- Neural networks with supervised learning: it adjusts the connection of the weights to match its output with the real output in a repetitive process until a desired result is obtained.
- Neural networks with unsupervised learning: self-organization is done and neural network only has the input and the output is not known to it. In this clustering of the input layer is done and the output generated is function of results obtained after clustering of inputs into different classes.

2.1 Proposed Model

In the proposed model feed forward network is used for the purpose of classification. In feed forward approach the data flow is unidirectional. There are no evaluation circles i.e. no feedback loops are present in the network.

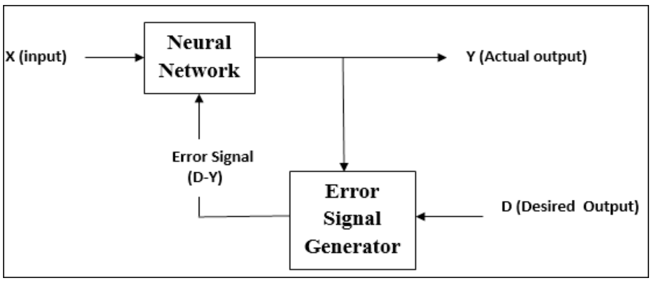

Figure 4: Backpropagation

A typical feed forward backpropagation model is used for the detection of alopecia. In figure 4 the general architecture of a feed forward backpropagation network is shown. X refers to the system's input while Y is the output generated by the system and known as the actual output. The system is presented with a set of examples to facilitate learning. D is desired output that is expected to be generated by the system and presented as the example output. The actual output is also connected to the error generating node where the delta rule is applied and calculates error as the difference between actual and desired output of the system. Apart from the actual output, the error generating signal is connected to the desired output as well. It receives the output from both the nodes and then apply delta rule which calculates the difference between both the outputs and that difference is the error from both the outputs [4,6,7] The error calculation is the main step so as to minimize the mis-prediction i.e., minimizing the difference between actual and target output. This minimization step is repeated several times until the output is optimized.

An artificial neural network works in three steps namely training, validation and testing. Training is the initial phase and is used to train the network on the sample data taken from the actual problem. The network is trained and then the performance of system is tested on a validation set. Further, the values of these parameters were analysed and those parameter values were selected which turn out to be having less errors and contributes more in determining the correct output, this is called the feature extraction and selection step. [5] The last phase involves testing of network over real data. Test data is utilized in the network to evaluate the performance of the network.

The proposed model consists of three layers: input, output and the hidden layer. The proposed system uses two hidden layers, each having 10 hidden neurons. The network is trained on the dataset of alopecia disease.

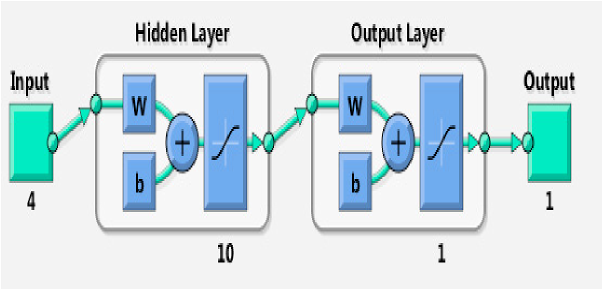

Figure 5: The proposed model

In figure 5 the proposed model is shown with one input layer, two hidden layers and one output layer. Each entity (neuron) in the hidden layer makes use of relocation function to process the data which it receives from previous layer, then relocate the processed information to the output entity for more processing using a relocation function in each neuron of output layer.

3. Experimental Results

3.1 Data Analysis

The dataset of alopecia was obtained from a recent experiment that is being conducted by researchers to obtain the cure of alopecia. 4 main attributes were considered in the dataset.

1. Length
2. Nail brittleness
3. Damage
4. Hair Follicle

Length of the hair was considered along with nail brittleness because both nails and hair are made up of keratin and only these are contains keratin as their main content. The amount of Damage made to the hair was also taken into consideration and lastly and more importantly hair follicle was taken into consideration which is the root of the hair, for identifying the disease. All the above factors are the main contributing factors which leads to severe damage to the hair.

Patients were considered as healthy or unhealthy on the basis of the above attributes. The dataset of 100 patients were considered out of which 80 samples were employed for training the network and 20 were employed for testing the networks.

3.2 Performance Evaluation

Neural network toolbox of MATLAB R2017 was used for simulating the proposed model. A two-layer feed forward network was created in which consists of 4 input neurons, 10 hidden neurons, and a linear output neuron. Feed forward backpropagation technique was used in the proposed model. [6] The result of applying the artificial neural network methodology to distinguish between affected and unaffected person based upon selected symptoms showed very good abilities of the network to learn the patterns corresponding to symptoms of the person. The accuracy of the given data is observed and evaluated for further use. Three plots were generated by the system: Training state, Performance plot and the regression plot.

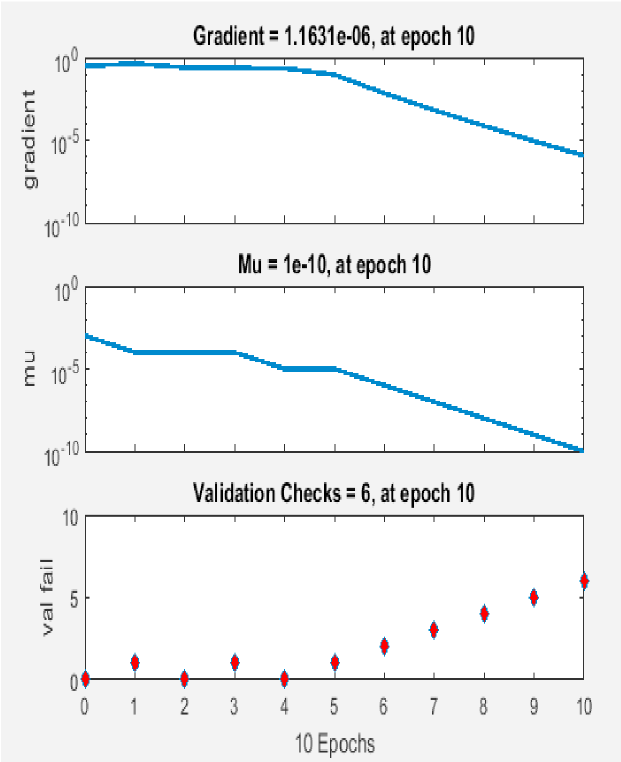

Figure 6: Training state

Figure 6 shows the training state in which the gradient is calculated. The value of the gradient is very important as it determines the condition of training and testing state. [7] Minimum the value of the gradient the better is the condition of training and testing set. In the above figure the value of gradient is 1.1631e-06 which means 0.011631. As the value of gradient is come closer to zero, the training and testing set are considered to be in good condition.

Performance plot is the plot that calculates the best validation performance during a certain epoch. The good performance is taken from the epoch with the lowest validation error. Mean square error (MSE) is the average square difference between output and target. Lower the value, better is the result. A value of 0 signifies no error i.e., a perfect classification or prediction. [7]

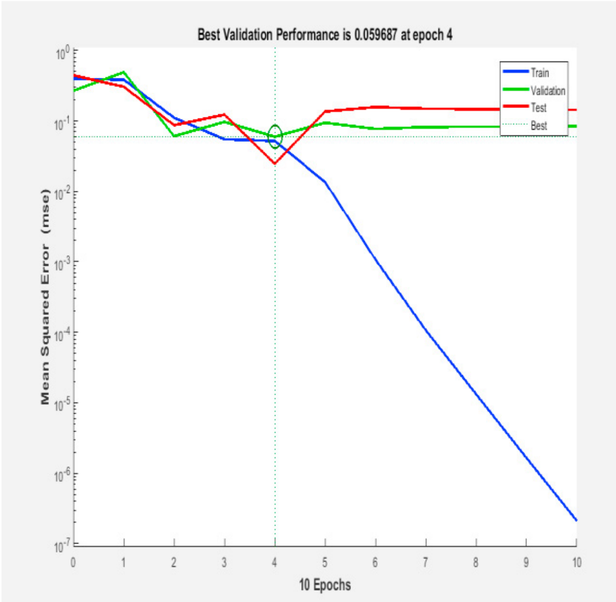

Figure 7: Performance plot

The best validation performance is found to be 0.059687 at epoch 4 during our experiments on the data. As the value obtained by proposed system is found to be closer to 0 it indicates a less amount of errors leading to near

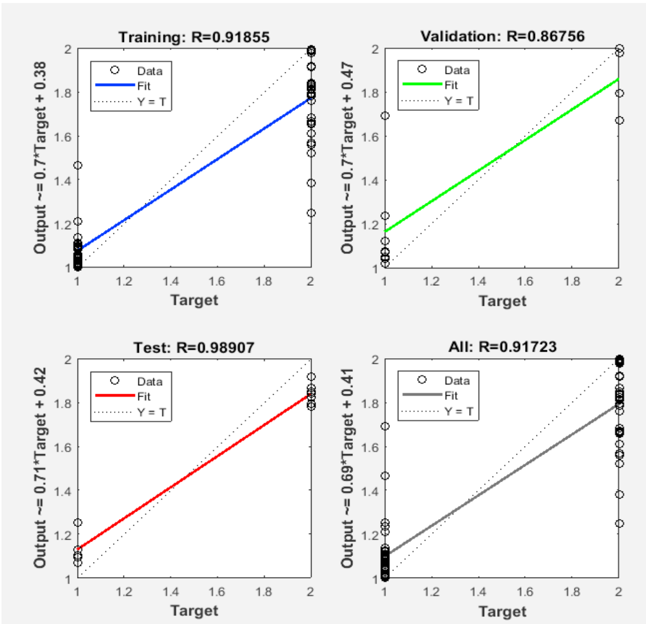

Figure 8: Regression plot

perfect classification accuracy. Figure 8 shows the regression plot contains 4 plots: Training, validation, and testing.The dashed line represents the perfect result in which the output is equal to the target. The solid line represents the best curve line between the output and the target. The value of R is the main value to be observed. The values of R indicate the relationship between output and the target. If the value of R is 1, it means that there is exact linear relation between output and the target. If the value of R is 0 or somewhere close to 0, it means that there is no linear relation between outputs and the targets. From the above regression table, it is seen that the final value for regression is 91%. Therefore, the proposed system achieved a 91% of accuracy over the data presented to it.

TABLE I. Results

	Regression (R)
Training	0.91855
Validation	0.86756
Testing	0.98907
All	0.91723

4. Conclusion

Early disease diagnosis is a compulsion as well as a necessary requirement in today's scenario. In developing countries, the problem becomes more severe because offaster growing population and lack of trained medical staff. This puts pressure on medical expert to diagnose the disease in a very short time and prescribe appropriate medication to help resolve the disease.In such scenario, an automated disease diagnosis system will help cater these problems and assisting the medical expert by providing a second opinion about the presence of disease. This will make the disease diagnosis process faster and less time consuming. The presented system will help find out early signs of alopecia disease by using an artificial neural network (ANN). The data was differentiated by the ANN based on the features extracted from the data and then the subject is categorized in one of the two categories: affected and unaffected patients. The data was analysed using feed forward backpropagation neural network. The performance of ANN was evaluated with respect to the learned features to differentiate healthy and affected patients. The above model can be very useful for clinical experts who generally wants a second opinion because of the dynamic nature of visible symptoms.

References

[1] R. Dybowski and v. Gant, clinical applications of Artificial neural networks, cambridge university press,2007.

[2] R. Dey-Rao, A.A. Sinha, A genomic approach to susceptibility and pathogenesis leads to identifying potential novel ther- apeutic targets in androgenetic alopecia Genomics 2017 Mar 2. pii: S0888-7543(17)30011-3

[3] Bhargava K, Fenton D; Chambers J, Tomlinson J. Alopecia areata and its impact on young people. Dermatological Nursing, March 2015, Vol 14, No 1

[4] N. Al-Milli, "Backpropagation neural network for prediction of heart disease," *Journal of theoretical and applied information Technology,* vol. 56, pp.131-135, Oct 10, 2013.

[5].Bhargava K, Fenton D; Chambers J, Tomlinson J. Alopecia areata and its impact on young people. Dermatological Nursing, March 2015, Vol 14, No 1

[6] S. Moein, S. A. Monadjemi and P. Moallem, "A Novel Fuzzy-Neural Based Medical Diagnosis System", International Journal of Biological &MedicalSciences, Vol.4, No.3, 2009, pp. 146-150.

[7] P. Francisco, G. C. Juan Manuel, S. P. Antonio and RF. Daniel, "A robust model of the neuronal regulator of the lower urinary tract based on artificial neural networks", Neurocomputing, Vol.71, No.4-6, 2008, pp.743-754.

[8]Gupta, Mrinal, and Venkataram Mysore. "Classifications of Patterned Hair Loss: A Review." Journal of Cutaneous and Aesthetic Surgery 9.1 (2016): 3–12. PMC. Web. 4 Dec. 2017.