Migraine Diagnosis by Using Artificial Neural Networks and Decision Tree Techniques

Ufuk CELIK, *Sakarya University, Turkey, ufuk.celik1@ogr.sakarya.edu.tr*

Nilufer YURTAY, *Sakarya University, Turkey, nyurtay@sakarya.edu.tr*

Ziynet PAMUK, *Sakarya University, Turkey, ziynet@sakarya.edu.tr*

Keywords:

ABSTRACT   *Computer supported studies in wide range of medical fields have been greatly expanded in recent years. Also, many medical organizations continue to build databases for different diseases. This medical database for artificial intelligence techniques for the determination of the disease is invaluable. As a subset, artificial neural networks and decision tree techniques are used for disease diagnosis. In this study Gini algorithm from decision trees and distributed delay network, probabilistic neural network, feed-forward network and learning vector quantization from artificial neural network have been used in order to diagnose migraine and probable migraine. Performance of these techniques has been compared and distributed delay network technique is observed as the best diagnosis with 95.45% accuracy.*

Keywords   *: headache, migraine, decision-tree, artificial neural networks, medical diagnosis*

Migraine Diagnosis by Using Artificial Neural Networks and Decision Tree Techniques

Headache diagnosis and treatment with computer support for the realization of the first studies began in Italy in 2002, with a study in collaboration with nine different clinical and patient records and statistical information were recorded with a program. The program compared the patient records according to International Headache Society criteria with simple methods and positive results have been obtained (Gallai et al., 2002). There is another program that was written in Microsoft Access database system in Naples University Federico II Headache Center in Italy followed by an article in 2004 and there was another article as verification for this program which says this program has more positive results than standard clinical diagnoses (Simone, Marano, & Bonavita, 2004). There was a web based study which was working as rule based in America in 2004 and it was used for patient

medical records (Kopec, Shagas, Selman, Reinharth, & Tamang, 2004). There was another web based system in America 2007 which is still online but it has not a user friendly design and it is not easy to use (Maizels & Wolfe, 2008). "Superbase" database system was used for the headache prevalence study and patient records were examined in Turkey. 54% of the doctors indicated that the error in the diagnosis of migraine, a part of luxury has been expressed that it was because of the density of the patient (Karlı et al., 2006). In another study artificial neural network analysis has been used for prediction of headache prognosis in elderly patient. Neural network model for grouped survival data can be used as a prognostic model. If the prevalence of a disease is low, the sensitivity of the model for detection of the patients with disease will be low (Taşdelen, Helvaci, Kaleağasi, & Özge, 2009). With an another study artificial neural network used for headache diagnosis and its performance was observed and results show that a neural network using binary vectors as input variables and five neural networks as a single output can obtain the best results in diagnosing for some frequents types of headache (Mendes, Fiuza, Teresinha, & Steiner, 2010).

Methods

By the International Headache Society, headaches criteria were defined for the first time in 1988 and revised in 2004 (Olesen, 2004). According to explanations there are two main types of headache. The primary type of headaches is not due to an illness other than pain. The secondary type of headaches is associated with a disease.

Migraine without aura and probable migraine without aura is in primary headaches and it is classified in 1.1 and 1.6.1 according to International Headache Society criteria. It is diagnosed according to criteria explained below.

1.1. Migraine without aura diagnosis criteria

A. At least 5 attacks fulfilling criteria B-D

B. Headache attacks lasting 4-72 hours (untreated or unsuccessfully treated)

C. Headache has at least two of the following characteristics:

unilateral location

pulsating quality

moderate or severe pain intensity

aggravation by or causing avoidance of routine physical activity (eg, walking or climbing stairs)

D. During headache at least one of the following:

nausea and/or vomiting

photophobia and phonophobia

E. Not attributed to another disorder

1.6.1 Probable migraine without aura diagnosis criteria

A. Attacks fulfilling all but one of criteria A-D for 1.1 Migraine without aura

B. Not attributed to another disorder

Dataset used in this study contains 70 training data and 30 test data and has 8 attributes and 3 types of classification knowledge. Attributes of dataset have been given in table 1. There are 16 patients with migraine, 33 patients with probable migraine and 21 patients without migraine in training set. The test dataset has 6 patients with migraine, 16 patients with probable migraine and 8 patients without migraine (Celik).

In this study decision tree has been obtained with Gini algorithm by using RapidMiner software. Gini algorithm is a method based on binary division of dataset. Attribute values of assets are owned by the Gini algorithm so that each one is divided into two groupings. Branches, divisions occur as a result of these groupings. Each attribute value of the binary elements of the group shall be deemed to have separated branches. This branch is used for attribute values in the group's formulation element numbers (Özkan, 2008). Gini formulation provides the first calculation of the left and right values. Each attribute for the left and right divisions and the Ginileft Giniright expressions are calculated as follows:

Li: On the left branch i the sample group (s) number

Ri: i group in the right branch sample (s) number

k: the number of classes

T: node samples

|Tleft|: Left branch sample (s) number

|Tright|: Right branch sample (s) number

Be calculated with the following definitions of relations.

$$Gini_{left} = 1 - \sum_{i=1}^{k} \left(\frac{L_i}{|T_{left}|} \right)^2 \qquad Gini_{right} = 1 - \sum_{i=1}^{k} \left(\frac{R_i}{|T_{right}|} \right)^2$$

The nature of the learning set, for each j the number of elements to be calculated the following correlation (Özkan, 2008).

$$Ginij \ = \ \frac{1}{n} \ (\ |Tleft|Giniletf + |Tright|Giniright)$$

Migraine without aura and probable migraine without aura decision tree obtained by Gini algorithm is shown in figure 1. The most important diagnostic criterion is pain duration. Another criterions are attacks number, unilateral pain, pulsative etc.

Feed-forward network shown in figure 2 consist of a series of layers. The first layer has a connection from the network input. Each subsequent layer has a connection from the previous layer. The final layer produces the network's output. Feed-forward network can be used for any kind of input to output mapping. A feed-forward network with one hidden layer and enough neurons in the hidden layers can fit any finite input-output mapping problem.

Distributed delay network are similar to feed forward network, except that each input and layer weights has a tap delay line associated with it. This allows the network to have a finite dynamic response to time series input data. This network is also similar to the time delay neural network which only has delays on the input weight (Beale, Hagan, & Demuth, 2014).

The probabilistic neural network shown in figure 3 developed by Specht (Specht, 1990) is a network formulation of 'probability density estimation'. It is a model based on competitive learning with a 'winner takes all attitude' and the core concept based on multivariate probability. The probabilistic neural network provides a general solution to pattern classification problems by following an approach developed in statistics, called Bayesian classifiers. The network paradigm also uses Parzen Estimators which were developed to construct the probability density functions required by Bayes theory. The probabilistic neural network uses a supervised training set to develop distribution functions within a pattern layer. Training of the probabilistic neural network is much simpler than that of the multi layer neural networks. However, the pattern layer can be quite huge if the distinction between categories is varied and at the same time quite similar in special areas.

 The learning vector quantization shown in figure 4 network was developed by Kohonen and its fundamental principle is to represent a vector by using a determined number of vectors (Kohonen, 1984),(Kohonen, Barna, & Chrisley, 1988). Like other multi-layered networks, the learning vector quantization network has 3 layers: Input Layer, Hidden Layer (i.e., Kohonen Layer) and Output Layer. On the input layer, the samples obtained from outside are given to the network, and the information coming with these samples comprise the input vector; no data processing is done. On the hidden layer, the closest weight vector to the input set is determined. Every element (i.e., neuron) on the hidden layer represents a reference vector. Input vector is mapped into a reference vector which consists of the weights between the input layer and the hidden layer (Alkim, Gurbuz, & Kilic, 2012).

Results

Receiver operating characteristic can be expressed as the rate of correct positives in wrong positives. The migraine diagnosis parameters used for receiver operating characteristic is given in table 2.

Parameters calculation formulas for receiver operating characteristic are given below in formula 1, 2 and 3 (Weinstein, Obuchowski, & Lieber, 2005), (Fawcett, 2006).

(1) Sensitivity = $\dfrac{TP}{(TP + TN)}$

(2) Certainty = $\dfrac{TN}{(TN + TP)}$

(3) Accuracy = $\dfrac{TP + TN}{(TP + TN + FP + FN)}$

The results obtained from decision tree and artificial neural networks studies are given in table 3. Distributed delay network algorithm gives the best results for accuracy.

When all the classifiers are compared, the best accuracy value was achieved with distributed delay network, the best sensitivity value was achieved with Feed-forward network and the best certainty values were achieved with Gini and distributed delay network.

Discussion

Another dataset which include 18-22 age people and total 535 units and prepare by physician was occurred. This study, in training phases are made up from total 381 units which include 137 migraines, 79 probable migraines and 165 no migraines. Test phases are made up from total 154 units which include 54 migraines, 32 probable migraines and 68 no migraines.

In this additional study was performed for 344 patients of migraine and probable migraine dataset contain 138 women and 206 men between the ages of 18-22. 62.99% of accuracy, 26.16% of sensitivity, 75.89% of specificity for 244 training and 100 test data using by distributed delay network technique was obtained. It was observed that the results of this study are lower than the results of the study mentioned above. The reason of this is that the data in the dataset does not show homogeneous distribution of migraine, probable migraine and no-migraine data. It is expected for the decision tree technique to show a low performance for this dataset. These rates are possible to increase with the increasing number of data and homogeneity in terms of classes. Hence, artificial neural networks software programs will be able to be used in diagnosing perfectly by using current patient diagnosis records and data.

Physicians can diagnose by taking into account their knowledge, experience and laboratory results. The doctors may make wrong decisions in diagnosing sometime although they are experienced. For example, medical diagnosis expert system can analyze the disease of patients and inform appropriate curing method by receiving the symptoms of patients and other related facts. Expert system has been greatly progressed, and most of those are being used to replace or assist human experts (Kwon, Kim, & Kim, 2009), (Cruz-Correia et al.,

2005). There have been several studies conducted to create a solution by using computer technologies in Turkey and all over world to eliminate human factor in diagnosing.

In this study, Gini algorithm from decision trees and distributed delay network, probabilistic neural network, feed-forward network and learning vector quantization from artificial neural network have been used in order to diagnose migraine and probable migraine. According to results of evaluation done by receiver operating characteristic analysis the probabilistic neural network is the lowest accuracy value whereas distributed delay network technique is 95.45% accuracy value. Distributed delay network technique can be used for migraine without aura and probable migraine without aura diagnosis.

References

1. Alkim, E., Gurbuz, E., & Kilic, E. (2012). A fast and adaptive automated disease diagnosis method with an innovative neural network model. Neural Netw, 33, 88-96. doi: 10.1016/j.neunet.2012.04.010

2. Beale, M. H., Hagan, M. T., & Demuth, H. B. (2014). Neural Network Toolbox™ User's Guide (Vol. R2014a): Matlab MathWorks.

3. Celik, U. Excel file for Gini dataset. from http://migbase.com/giniResults.xls

4. Cruz-Correia, R., Vieira-Marques, P., Costa, P., Ferreira, A., Oliveira-Palhares, E., Ara, F., . . . Costa-Pereira, A. (2005). Integration of hospital data using agent technologies - A case study. AI Commun., 18(3), 191-200.

5. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recogn. Lett., 27(8), 861-874. doi: 10.1016/j.patrec.2005.10.010

6. Gallai, V., Sarchielli, P., Alberti, A., Pedini, M., Gallai, B., Rossi, C., . . . The Collaborative Group for the Application of, I. H. S. C. o. t. I. S. f. t. S. o. H. (2002). Application of the 1988 International Headache Society Diagnostic Criteria in Nine Italian Headache Centers using a Computerized Structured Record. Headache: The Journal of Head and Face Pain, 42(10), 1016-1024. doi: 10.1046/j.1526-4610.2002.02231.x

7. Karlı, N., Zarifoğlu, M., Ertaş, M., Saip, S., Öztürk, V., Bıçakçı, Ş., . . . Uzuner, N. (2006). Economic impact of primary headaches in Turkey: a university hospital based study: part II. The Journal of Headache and Pain, 7(2), 75-82. doi: 10.1007/s10194-006-0273-7

8. Kohonen, T. (1984). Self-organization and associative memory: Springer-Verlag.

9. Kohonen, T., Barna, G., & Chrisley, R. (1988, 24-27 July 1988). Statistical pattern recognition with neural networks: benchmarking studies. Paper presented at the Neural Networks, 1988., IEEE International Conference on.

10. Kopec, D., Shagas, G., Selman, J., Reinharth, D., & Tamang, S. (2004). Development of an Expert System for Aiding Migraine Diagnosis. The Journal of Information Technology in Healthcare, 2(5), 355-364.

11. Kwon, P.-J., Kim, H., & Kim, U. (2009). A study on the web-based intelligent self-diagnosis medical system. Advances in Engineering Software, 40(6), 402-406. doi: http://dx.doi.org/10.1016/j.advengsoft.2008.07.004

12. Maizels, M., & Wolfe, W. J. (2008). An expert system for headache diagnosis: the Computerized Headache Assessment tool (CHAT). Headache, 48(1), 72-78.

13. Mendes, K. B., Fiuza, R. M., Teresinha, M., & Steiner, A. (2010). Diagnosis of Headache using Artificial Neural Networks. J. Comput. Sci, 10(7), 172-178.

14. Olesen, J. (2004). Preface to the second edition. Cephalalgia, 24, 9-10. doi: 10.1111/j.1468-2982.2003.00824.x

15. Özkan, Y. (2008). Veri madenciliği yöntemleri: Papatya Yayıncılık.

16. Simone, R., Marano, E., & Bonavita, V. (2004). Towards the computerisation of ANIRCEF Headache Centres. Presentation of AIDA CEFALEE, a computer assisted diagnosis database for the management of headache patients. Neurological Sciences, 25(3), s218-s222. doi: 10.1007/s10072-004-0290-8

17. Specht, D. F. (1990). Probabilistic neural networks. Neural Networks, 3(1), 109-118. doi: http://dx.doi.org/10.1016/0893-6080(90)90049-Q

18. Taşdelen, B., Helvaci, S., Kaleağasi, H., & Özge, A. (2009). Artificial neural network analysis for prediction of headache prognosis in elderly patients. Turkish Journal of Medical Sciences, 39(1), 5-12.

19. Weinstein, S., Obuchowski, N. A., & Lieber, M. L. (2005). Clinical Evaluation of Diagnostic Tests. American Journal of Roentgenology, 184(1), 14-19. doi: 10.2214/ajr.184.1.01840014

Table 1

Dataset attributes used in decision tree and artificial neural networks techniques

Attribute no	Attribute Name
1	min 5 attacks
2	4-72 hours duration
3	unilateral
4	pulsative
5	moderate or severe
6	increase in movement
7	nausea and/or vomiting
8	photophobia or phonophobia
Class	headache (migraine, no-migraine, propable migraine)

Table 2

Receiver operating characteristic parameters

Abbreviations	Description
TP: True Positive	The number of the patients diagnosed by the software in the list which includes the patients diagnosed as migraineurs by physician
TN: True Negative	The number of healthy people diagnosed by the software in the list which includes the healthy people decided by physician
FP: False Positive	The number of the patients diagnosed as migraineurs by the software in the list which includes the healthy people decided by physician
FN: False Negative	The number of healthy people diagnosed by the software in the list which includes the patients diagnosed as migraineurs by physician

Table 3.

Receiver operating characteristic analysis for migraine and probable migraine diagnosis

Classifier	Sensitivity(%)	Certainty(%)	Accuracy(%)
Decision tree (GINI)	83.33	93.75	90.90
DDN distributed delay network	100	93.75	95.45
PNN probabilistic neural network	82.82	88.09	84.62
FFN feed-forward network	96.96	88.88	90.30
LVQ learning vector quantization	94.44	90.47	90

Figure Captions
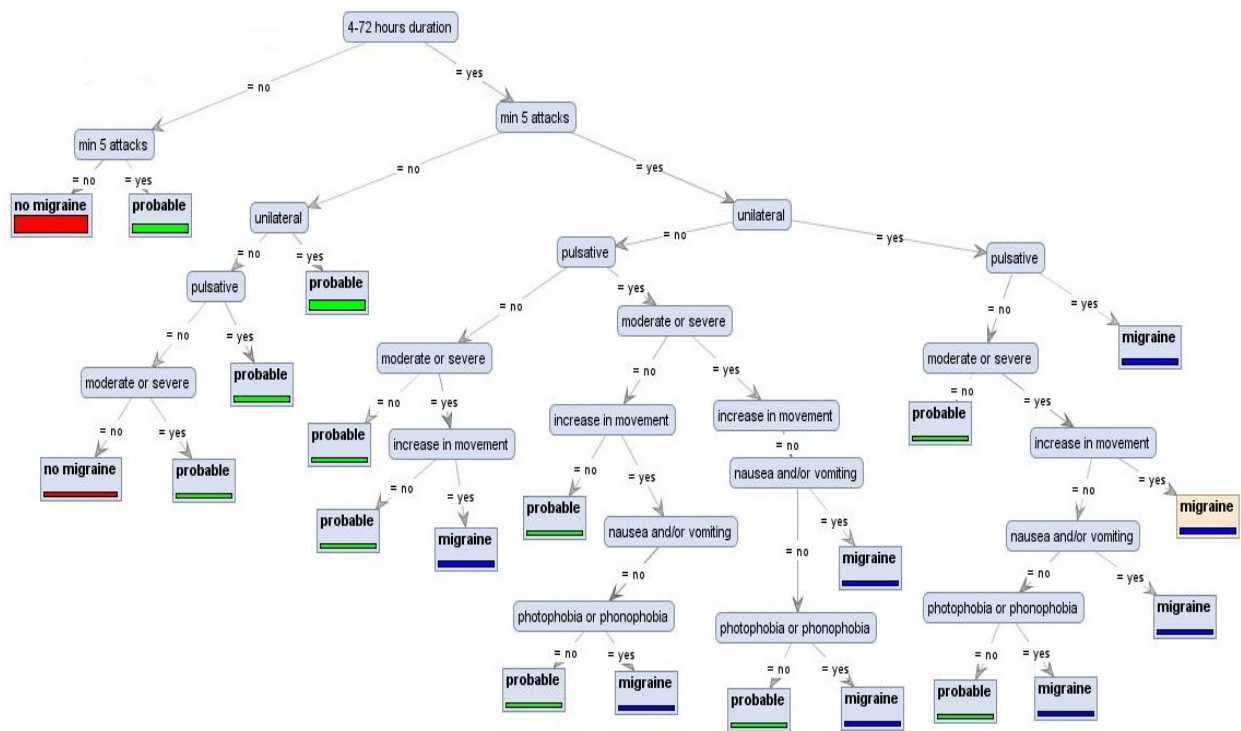
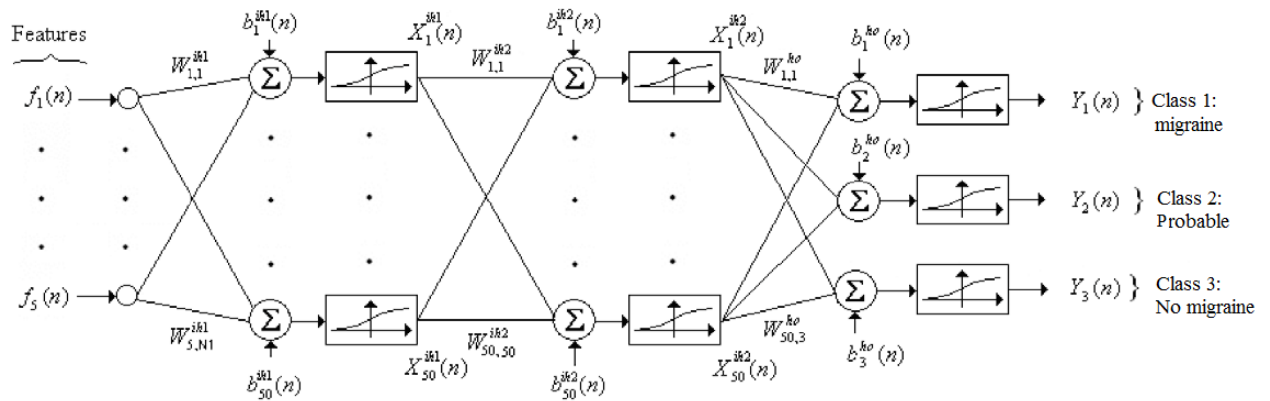Figure 1. Migraine and probable migraine decision tree obtained by using Gini algorithm

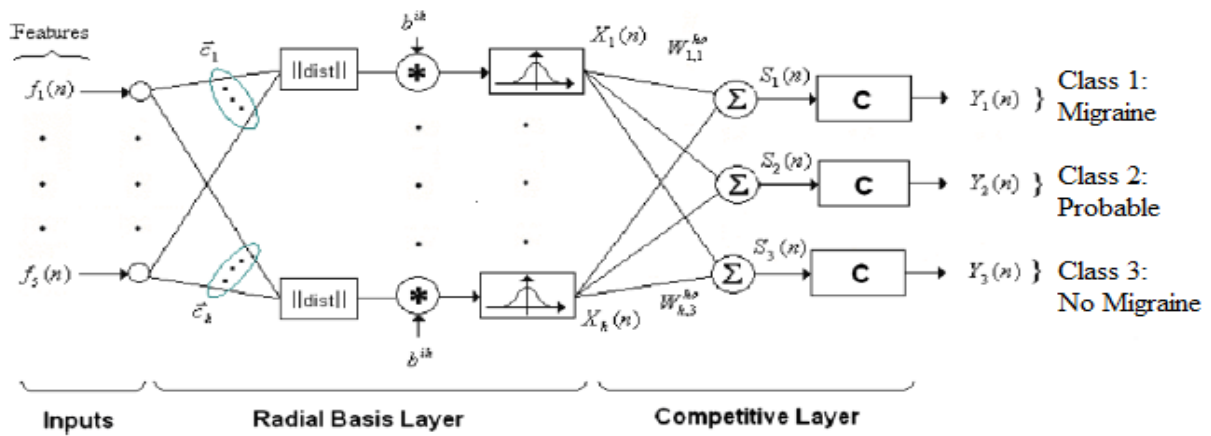Figure 2. Implementation of multilayer neural network for the migraine disease diagnosis

Figure 3. Implementation of probabilistic neural network for the migraine disease diagnosis
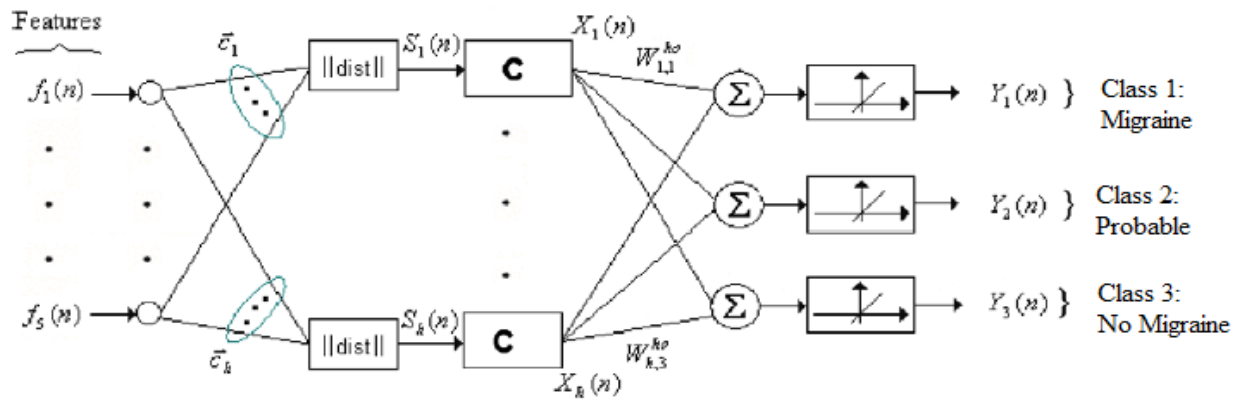

Figure 4. Implementation of learning vector quantization neural network for the migraine disease diagnosis