Editorial Manager(tm) for Primary Health Care Research & Development
Manuscript Draft

Manuscript Number: PHC-D-10-00010R1

Title: Diagnosis of Headache using Artificial Neural Networks

Short Title: Headache Diagnosis with Artificial Neural Networks

Article Type: Research

Keywords: Neural network, primary health care, referral and consultation, migraine, headache

Corresponding Author: Dr. Ronald Moura Fiuza, M.D.

Corresponding Author's Institution: Clínica Neurológica de Joinville

First Author: Karina B Mendes, MS

Order of Authors: Karina B Mendes, MS; Ronald Moura Fiuza, M.D.; Maria Teresinha A Steiner, PhD

Manuscript Region of Origin: BRAZIL

Abstract: We intend to create an auxiliary tool for headache diagnosis to be used in primary level of Health Care System. Data were collected from a structured questionnaire which was applied to 2177 patients with headache. Artificial neural networks were trained to reach the diagnosis performed by a single neurologist. We tested them varying their activation function, initial weights and number of elements in input, hidden and output layers. The best results were obtained using binary coordinates as input vectors (information of the questionnaire) and one individual neural network for each output (diagnosis). Sensitivity and specificity of artificial neural networks were respectively 0.93 and 0.91 for tension-type headache identification, 0.99 and 0.94 for migraine without aura, 1.0 and 0.98 for migraine with aura and 1.0 and 0.96 for medication-overuse headache. Artificial neural networks can be used as tools to support the diagnosis of these common forms of headache.

Suggested Reviewers:

Opposed Reviewers:

TITLE: Diagnosis of Headache using Artificial Neural Networks

1 INTRODUCTION

The World Health Report 2008 is titled "Primary Health Care: now more than ever". The document is the yearly report of the World Health Organization (WHO). It seeks to strengthen the concepts erected in the Alma Ata Conference (1978). This conference recommended primary health care as the most efficient basis on which a health system can be supported to provide equal access to services and better use of existing resources (World Health Organization, 1978 and 2008).

The WHO's effort to spread this concept aims to create fair and efficient systems, with universal applicability and is especially important in poor countries (Pan American Health Organization, 2007). Through the creation of a large gate to health systems, the population has the possibility to receive the concepts of health promotion and disease prevention and yet to receive attention on prevalent diseases in each region.

The headaches stand out as one of the most common groups of diseases. Some epidemiological studies estimate up to 93% of prevalence. As it affects most people, it is one of the biggest reasons for medical consultation (Boardman et al, 2003). This generates significant impact on quality of life, with interference in relationships at work, family and leisure (Dahlöf and Solomon. 2006). It still represents a considerable social and economic burden, not only through direct costs to the health system but also indirect costs, related to absenteeism and low productivity at school or work (Berg and Ramadan, 2006).

However, most headache patients do not demand or have access to the health system, and therefore they medicate themselves (Lipton, Stewart and Simon, 1998; Lipton et al, 2002). Among those who seek the doctor, most are served by primary health care systems and are not reviewed in secondary or tertiary levels. As a result of these distortions, it is estimated that about half of patients never receive a correct diagnosis and, therefore, are not properly treated (Lipton and Bigal, 2007).

To improve this situation it is necessary to increase the proportion of patients with access to diagnosis and appropriate treatment. As recommended by the WHO, it is possible to develop a system of increasing complexity, which could provide broader access and appropriate referral of more complex cases.

To introduce this model, it is necessary to develop tools which can be used by the non-medical professional at a preliminary assessment. In the case of headaches, the initial step should be a standardized method capable of distinguishing the most frequent types of headache, allowing an initial screening (Lipton and Bigal, 2007).

As structured questionnaires have been frequently used in specialized clinics as part of the medical anamnesis of these patients, it is possible to develop a treatment of these data, allowing the prediction of a preliminary diagnosis, at least of some frequent headaches. One way to do it is to develop and to validate a questionnaire (Lipton et al, 2003), whose answers could lead to a differential diagnosis, as defined by the International Classification of Headaches Disorders (Headache Classification Committee of the International Headache Society, 2004). Another alternative would be to use available data from existing structured anamnesis questionnaires in specialized clinics and assess, through computing, how such data can lead to a diagnosis.

An artificial neural network (ANN) is considered both as method and practice to solve problems in artificial intelligence. The most important characteristic of an artificial neural network is its ability to learn and improve their performance. It is based on real data, from which it derives a general model in an attempt to build patterns.

The results achieved with ANN have been compared with those obtained with traditional statistical methods, usually easier to use. Although controversial, there are several studies showing the advantages of ANN (Steiner et al, 2003; Nguyen and Cripps, 2001; Setiono, Leow and Zurada, 2002).

The network serves as a massively parallel and distributed processor. It consists of simple processing units, called artificial neurons. These are arranged in different layers, including an input layer through which data arrive, intermediate or hidden layers that help in processing and an output layer, which bring the result. The artificial neurons are connected by communication channels (connections), which usually are associated with numerical values (weights). These units are activated by mathematical functions, perform calculations and have large capacity in processing and storage of information (Haykin, 2001).

Hence, recognition of patterns is one of the main applications of ANN, including problems in areas such as trade, industry, services, education and science, among others. In the medical area, ANN has been used as support for both diagnosis and treatment of many different diseases (Malmgren, Borga and Niklasson, 2000).

The correct implementation of artificial neural network regarding the approach to headaches can have the following advantages:

• Guide for family doctors and professionals who act on primary health care.

• Facilitate and complement the work of specialist doctors.

• Less subjectivity in the diagnosis.

As the ANN learn through examples, details of how to recognize the disease are not necessary. Nevertheless, the training of an ANN requires a series of cases serving as representative examples of variations of the disease (Beale and Jackson, 1990).

With validation of the method and training of health workers, access from headache patients to health services can be much extended. From the resulting provisional diagnosis, patients can be referred to the appropriate health care level. Simple cases can be orientated by the health agent (Dowson et al, 2002). Cases that need medical advice can be attended by a family doctor or general practitioner. Complex cases would be referred to specialist or even hospitalization (De Hertogh et al, 2007; Ridsdale et al, 2008; De Klippel, Jansen and Carlos, 2008).

The purpose of this study is to verify the application of artificial neural networks as a technique of pattern recognition that can help in the medical diagnosis of the most frequent types of headache.

2 METHODS

The study data were extracted from the database of Neurological Clinic of Joinville / Brazil. A structured questionnaire was applied to 2177 consecutive headache patients from January 2002 to November 2006. The questions included form a body able to meet the necessary criteria for the more frequent diagnosis.

These diagnoses were made by a single neurologist, who uses in his daily practice the diagnostic criteria of the International Classification of Headaches Disorders. From the database, was chosen always the first diagnosis, made in the first interview. After 2004, it was used the criteria of the ICHD-2 (Headache Classification Committee of the International Headache Society, 2004).

To be used as training examples for artificial neural networks, we selected some prevalent and representative diseases. In that way, we could assess some of the most common causes of headaches and have a critical number of cases, so that the ANN training could be effective. Of the 2177 patients with headache, 1252 patients were diagnosed as tension-type headache, 307 as migraine without aura, 99 as migraine with aura and 100 patients as medication-overuse headache. 419 patients had other types of diagnosis.

The above-mentioned questionnaire consists of 14 multiple choice questions answered about the history of each patient: sex (male / female) and age of the patient (until 12 / 13 to 19 / 20 to 39 / above 40), onset of the pain (x days / weeks/ x months/ x years), localization (unilateral / bilateral / frontal / occipital / localized), intensity (mild / moderate / severe / very severe) and characteristics of the headache (pulsating / pressing / stabbing / atypical), episode pattern (episodic / chronic / cluster) and duration (seconds / minutes / hours/ days/ weeks), evolution (stable/ changing/ progressive) and frequency of pain (daily / more than 15 a month / 1 to 2 a week / 1 to 3 a month / some in the year), factors associated (nausea / vomit / photophobia / phonophobia / aura / dizziness / fever), precipitating (stress and mental tension / menstruation / food/ spirit/ sleep problems/ physical activity/ sex) and relieving (sleep/ relax / distraction / walk / pregnancy) and use of medications (never / rarely / as necessary / almost every day / daily / many a day). As multiple choice questions, they facilitate the dynamics of consultation and data processing.

ANNs are computational techniques able to recognize patterns. They must learn how to go to output data beginning from input data. We had already the input data (the questionnaire) and the output data (the diagnosis). Our job was to test which ANN configuration could better achieve this task.

There are multiple options for handling such data, seeking the best ANN performance.   We tested different treatments for both input and output data. The input data were treated in two ways, intending to perform different tests. In the first option the data were represented as a simple numerical scale, with a value for each multiple choice questionnaire option. We called it code A. In the second option the data were treated in a binary form, with the number "1" representing a positive response and the number "0", the negative response. This option was called code B.

Exemplifying, the variable "sex" presented in code A the value "1", if the patient was female and "2" if male. In code B, the vector (1, 0) was the representation for female patients and (0, 1) for males. Thus, the attribute "sex" had a binary vector with two coordinates. Other variables were treated similarly.

In code A, the input data matrix was 2177 x 14 (number of patients X number of questions about each patient).

In code B, we had an input vector with 66 binary coordinates, corresponding to the alternatives to 14 multiple-choice questions posed to each patient. The input data matrix was then 2177 x 66.

The output data were limited to the more frequent headache diagnoses. Five outputs were defined: one for tension-type headache (TTH), another for migraine without aura (MoA), one for migraine with aura (MWA) and a forth for medication-overuse headache (MOH). We represented all other diagnoses by a fifth output.

We could use several different ANN structures that could reach these five outputs. We chose to test two options: the first was to construct one ANN with five outputs. The second was five ANN with a single output each.

So we built two input options (codes A and B) and two output options (one or five ANN). We could prepare several different and independent tests.

The artificial neural networks of multiple layers type were created, trained and simulated using the software "MATLAB 7.0" and its statistics and data analysis "Neural Network Tollbox" (MATLAB®, Mathworks, Natick, MA, USA) (Palm, 2008). They enable the creation of different types of networks and offer freedom to change the parameters of the network. We used the back-propagation method for teaching the neural networks.

An ANN is composed of a number of artificial neurons connected by links (*figure 1*). Each link has a number associated as a weight. The network knowledge is acquired by the update of the weights. An artificial neuron has a number of inputs (x_1, $x_2 \ldots x_n$). Each input signal is multiplied by a weight (w_1, $w_2 \ldots w_n$), that can be negative or positive. Then it is performed the weighted sum of input signals applied to the artificial neuron. Next the bias is applied (θ), to increase the model freedom, enhancing the network capacity to adjust itself to the provided knowledge. The sum is processed by the activation function to produce the output neuron.

The ANN is based on widespread input pattern mapping to an area of output, minimizing the error between the produced output and the provided output (Rojas and Feldman, 1996). An algorithm was created to test different networks configurations, thus analyzing the influence of parameters in final simulation results.

The number of layers and the number of elements of each layer were set initially. Only one intermediate layer is necessary to approach a continuous function (Haykin, 2001). Thus, this work was composed of one input layer, one hidden layer and one output layer.

The activation functions used in this study were:

• Logistic sigmoid (logsig) both in hidden and output layers.

• Tangent sigmoid (tansig) in hidden layer and linear function in output layer.

• Tangent sigmoid function in hidden layer and logistic sigmoid function in output layer.

• Logistics sigmoid in hidden layer and linear function in output layer.

The network is generally initialized with "x" number of elements (called “artificial neurons” or simply “neurons”) in hidden layer, trained through "y" iterations. In our experiment error was calculated every 25 iterations. This procedure can be repeated "z" times. Then a new network is initialized with a different number of neurons in the hidden layer and so on. Finally, the algorithm should find the best resulting network, which had classified correctly the most number of patterns. Thus, the algorithm generated "z" initial different networks (with different sets of random weights), with the same number of neurons in the hidden layer (best topology). It trained them with "y" iterations and, at the end of the process, showed the weights that came closest to the desired outputs. The network was then tested to verify its ability to generalize.

The 2177 diagnoses of the sample were divided into two sets. Due to the large database, we used the holdout procedure to assess the spread of networks trained. This procedure separates 2/3 of data for network training and the rest 1/3 to test it (Lin and Shaw, 1997). The sample was stratified, ensuring that each class was represented in proportion in the two data sets (training and testing). In each three patients, two were selected for training and one for test.

The technique was implemented varying the activation functions and the number of neurons in input and output layers, to train the networks in three different tests:

• Test I: One ANN training, with 66 neurons in input layer (code B), and five neurons in output layer.

• Test II: Five ANN training, each with 66 neurons in input layer of and one neuron in output layer.

• Test III: One ANN training, with 14 neurons in input layer (code A) and five neurons in output layer.

There was also variation in the following parameters:

• Number of neurons in the hidden layer: In each test, the network was trained initially without the hidden layer and further tests using, respectively, 1, 2, 3, until 10 neurons in one hidden layer.

• Initial weights (five sets of randomly chosen were used in each test).

A critical decision is when stop. One of the most used criteria is the number of iterations (Cohen, 1995). In this study, the number of iterations was set in 1000 cycles.

The software showed the results for each configuration. We calculated the percentage of correct answers and selected the best results in each test. Finally we measured the performance of the test showing the best results among all.

3 RESULTS

In test I, the network was trained using the code B, with 66 binary inputs in the input network vector. The output layer was composed of 5 neurons. The results showed that the best architecture had 9 neurons in the hidden layer, using logistic sigmoid activation function in both hidden and output layers. The best simulations appointed 73.04% correct answers in 2177 cases.

In test II five networks were trained, with only one output each. Each network had one headache diagnosis, corresponding to one output:

• ANN I: TTH

• ANN II: MoA

• ANN III: MwA

• ANN IV: MOH

• ANN V: Other types of headache

The input vector was composed of 66 binary components, according to the code B. The result of the network output, with only one neuron, was calculated by approximation. The figures next to "1" (0.5-1) represented a type of headache and values close to "0" (0 to 0.5) the non occurrence of such headache.

The activation functions used in test II were logarithmic sigmoid both in hidden and in output layers in first instance. Then, we used tangent sigmoid function in hidden layer and logarithmic sigmoid function in output layer.

The highest correct percentage answers in test II showed 87.83% in output TTH, 94.99 in MoA, 98.53 in MwA, 95.73 in MOH and 88.25 in other types of headaches. All of them refer to the 2177 cases.

In test III networks were trained using code A, with 14 entries. The output layer was composed of 5 neurons. We used logarithmic sigmoid activation functions in both hidden and output layer, and then tangent sigmoid function in hidden layer and sigmoid logarithmic function in output layer.

The best architecture in this test presented 10 neurons in hidden layer, using tangent sigmoid activation function in hidden layer and logarithmic sigmoid activation function in output layer. The percentage of correct answers was 66.65% in 2177 cases.

Therefore, the best of all results for the prediction of the diagnosis of headache patients were obtained in test II, as shown in *Table 1*. It was found that five ANN, with only one output each, have better performance than one ANN with five outputs and that binary coordinates facilitates the network processing in both input and output. These results were then obtained using binary coordinate vectors as input variables (code B).

We measured the performance of test II results. *Table 2* shows performance measures obtained in these results, with regard to the chosen types of headaches. Sensitivity and specificity of artificial neural networks were respectively 0.93 and 0.91 for tension-type headache identification, 0.99 and 0.94 for migraine without aura, 1.0 and 0.98 for migraine with aura, 1.0 and 0.96 for medication-overuse headache and 0.95 and 0.87 for other types of headaches.

4 DISCUSSION

Our results show that a neural network using binary vectors as input variables and five neural networks as output can obtain the best results in diagnosing some frequents types of headache. We can see that these results could offer new possibilities for the headache patients with access problems to the Health System.

Despite the difficulties of epidemiological studies in headache, it is known that these diseases are highly prevalent, with large social costs (Boardman et al, 2003). Population studies vary widely, with some prevalence studies results showing over 90%. A recent study in southern Brazil shows annual prevalence of 80.8% (Queiroz, Barea and Blank, 2006).

The ideal way to establish a diagnosis of headache is through consultation with a neurologist who uses structured diagnostic criteria, such as the International Classification of Headache Disorders. The specialist has the ability to identify not only the most common causes of headaches, but also to evaluate the doubtful and mixed cases, secondary headaches and rare cases (Marks and Rapoport, 1997). We know, however, that access to neurologist is limited and expensive, and impossible to many patients.

For over 30 years the WHO advocates a model of increasing complexity to health care (WHO, 2008). In this period, a wide network of primary care has developed around the world, facilitating access to health system, especially for the low-income population.

To build a comprehensive system to headache patients in primary care level, it is necessary first of all to develop a questionnaire that could be applied to these patients by non-specialists. This questionnaire should be simple to implement and lead to an indication of the main diagnosis of headache (Khu, Siow. and Ho, 2008).

Various efforts have been made to develop questionnaires that might indicate the diagnosis of a primary headache (Sheftell et al, 2005; Hagen et al, 2000). Most of them are directed to the evaluation of patients suspected of migraine ([Lipton et al, 2003). Others expand their coverage to tension-type headache and trigeminal-autonomic headaches (Mongini et al, 2003; Yoon et al, 2008). Not all attempts were satisfactory ([Rasmussen, Jensen and Olesen 1991).

The artificial neural networks are systems that simulate human brain behavior, using trial and error in the process of new knowledge accumulation. Its use is broad and comprehensive, and its usefulness is recognized in information processing derived from public and private sectors. There are examples of its successful use in the health area, where they can be important instruments to physicians, administrators and other health professionals.

ANN is a tool that can learn and grow. They can learn from structured questionnaires used by experts. After properly trained, they can serve as a support mechanism for the generalist. Widely used in industry, commerce and services, this methodology has been applied successfully in several areas of medicine. Regarding headaches, there are papers using ANN, as those who seek graphic patterns in patients with headache (Bellotti et al, 2007), correlation of symptoms with environmental factors (Cathcart and Materazz, 1999) and even applications for classification of headaches (Aston et al, 1994).

In this study, ANN has been explored to solve the problem of predicting the most frequent diagnoses of headaches, which is one of the most common symptoms that afflict world population. We used the database from a specialized clinic that has in its system a structured headache questionnaire, which resulted in the data that munitioned the inputs of ANN.

All ANN were separately trained, in supervised learning. We varied the following parameters, with the goal of finding the best structure of network: the activation functions of input and hidden layers, number of neurons in input and output layers; number of neurons in the hidden layer.

The basic element of an artificial neural network is then called artificial neuron. The behavior of a "neuron" is determined by its associated functions and its input and output connections (Palma Neto and Nicoletti, 2005). It is through an activation function that the generated responses of the unity are calculated. It is a mathematical function and is very important for the behavior of an ANN, as it defines the output, and thus, the information pathway (Mitchell T., 1997). There are several types of activation functions. The most used are the Gaussian, sigmoid, sine, linear, logarithmic, and tangents functions. We can conduct various tests with different functions, to achieve the definition of the viability of a test.

The use of ANN proved to be adequate to treat the information from 2177 patients. The database was implemented from the data (symptoms) of each patient and the knowledge of expert (diagnosis).

ANN was shown to be a valuable tool for patterns recognition in medical diagnosis of most frequent types of headache. Thus, making use of the tool, health professionals could make predictive diagnosis in patients with tension-type headache, migraine without aura, migraine with aura and medication overuse headache.

The diagnosis is an imperfect process. Theoretically, it is preferable to have a test both highly sensitive and specific. However, this procedure is not usually possible. Many tests are actually based on a clinical measure that can take a series of values, in which case there is an inherent compromise between sensitivity and specificity.

The relationship between sensitivity and specificity can even be illustrated using a ROC curve. It plots the probability of a positive real - or the sensitivity of the test - versus the probability of a false positive result for a number of different cut points.

That is the reason of having only a sensitivity value for each test.

So there is good concordance between the gold standard and: MOH and others. For TTH, MoA and MwA there is excellent concordance with the gold standard

In our case we could use a kappa index to verify the agreement between two tests, where none of them provides an exact answer for a diagnosis (table 3)

It is also an instrument that can expand, by incorporating new signs and symptoms in input and new diagnoses in output. Thus, attaching some data of the physical and neurological examination to the questionnaire, ANN can be trained again, to assess their ability to assist in the diagnosis of other primary and secondary headaches, indicating particularly the cases of high risk.

Predicting some frequent diagnoses of headache, this tool could be used as screening, optimizing available resources in health system. Thus, in the gateway to the system, all patients registered for consultation would be asked about their chief complaint. Those whose complaints were headache would be referred for a health agent or a nurse trained to implement the questionnaire and obtain the diagnostic hypothesis. The only equipment required would be a simple computer. With proper supervision, a framework capable of guiding the patient in this primary care environment could be designed. The simplest cases (TTH, MoA and MwA) could be referred for follow-up by the family doctor. The more complex cases (MoH, other types) would be referred to a specialist and even to the hospital. Many unnecessary tests and consultations could be avoided, relieving the health system. Ultimately, the model could provide agility, economy and functionality to the system, with regards to the many patients with headache.

5 CONCLUSIONS

Artificial neural networks can be used as tools to support the diagnosis of common forms of headache and can therefore cooperate for greater access to the health system.

Acknowledgments: Appreciation is expressed to Dr. Glauco Adriano Westphal, Dr. Luiz Paulo de Queiroz and Dr. Luiz Carlos Daquino for their helpful comments concerning this manuscript and to Dr. Creusa Dal Bó for assistance with data management.

NO CONFLICT OF INTERESTS

6 REFERENCES

1. Aston M.L., Wener M.H., Thomas R.G., Hunder G.G. and Bloch D.A., 1994: Application of neural networks to the classification of giant cell arteritis. Arthritis Rheum. 37(5):760-70.
2. Beale R. and Jackson T., 1990: Neural Computing: An Introduction. Bristol: Institute of Physics Publishing.
3. Bellotti R., de Carlo F., de Tommaso M. and Lucente M, 2007: Migraine detection through spontaneous EEG analysis. Conf Proc IEEE Eng Med Biol Soc. 2007:1834-7
4. Berg J. and Ramadan N.M., 2006: Societal burden of the headache. In: Olesen J., Goadsby P.J., Ramadan N.M., Tfelt-Hansen P. and Welch K.M.A. (Eds) The Headaches. 3rd ed. Philadelphia: Williams and Wilkins; pp 35-42.
5. Boardman H.F., Thomas E., Croft P.R. and Millson D.S., 2003: Epidemiology of headache in an English district. Cephalalgia.23 (2):129-37.
6. Cathcart S. and Materazzo F., 1999: Headache interference as a function of affect and coping: an artificial neural network analysis. Headache 39(4):270-4.

7. Cohen, P. R , 1995: Empirical Methods for Artificial Intelligence. Cambridge MA: The MIT Press.

8. Dahlöf C.G.H. and Solomon G.D., 2006: Impact of headache on the individual and family. In: Olesen J., Goadsby P.J., Ramadan N.M., Tfelt-Hansen P. and Welch K.M.A. (Eds) The Headaches. 3rd ed. Philadelphia: Williams and Wilkins; pp 27-34.

9. De Hertogh W.J., Vaes P.H., Devroey D., Truijen S., Duquet W. and Oostendorp R., 2007: Management of headache disorders: design of a randomised clinical trial screening for prognostic patient characteristics. BMC Musculoskelet Disord 26;8:38.

10. De Klippel N., Jansen J.P. and Carlos J., 2008: Survey to evaluate diagnosis and management of headache in primary care: Headache Management Pattern programme. Curr Med Res Opin. 24(12):3413-22

11. Dowson A.J., Lipscombe S., Sender J., Rees T. and Watson D., 2002: Migraine In Primary Care Advisors: New guidelines for the management of migraine in primary care. Curr Med Res Opin. 18(7):414-39.

12. Hagen K., Zwart J.A., Vatten L., Stovner L.J. and Bovim G., 2000: Head-HUNT. validity and reliability of a headache questionnaire in a large population-based study in Norway. Cephalalgia 20(4):244-51.

13. Haykin, S. 2001: Redes neurais. Porto Alegre: Bookman. Portuguese.

14. Headache Classification Committee of the International Headache Society, 2004: ICHD-II Abbreviated Pocket version. Cephalalgia. 24 (Suppl 1):1-160.

15. Khu J.V., Siow H.C. and Ho K.H., 2008: Headache diagnosis, management and morbidity in the Singapore primary care setting: findings from a general practice survey. Singapore Med J. 49(10):774-9.

16. Lin F.R. and Shaw M.J., 1997: Active training of backpropagation neural networks using the learning by experimentation methodology. Annals of Operations Research 75:105-122.

17. Lipton R.B., Dodick D., Sadovsky R., Kolodner K., Endicott J., Hettiarachchi J. and Harrison W., 2003: A self-administered screener for migraine in primary care: The ID Migraine validation study. Neurology 12;61(3):375-82.

18. Lipton R.B., Scher A.I., Kolodner K., Liberman J., Steiner T.J. and Stewart W.F., 2002: Migraine in the United States: epidemiology and patterns of health care use. Neurology 26;58 (6):885-94.

19. Lipton R.B., Stewart W.F. and Simon D., 1998: Medical consultation for migraine: results from the American Migraine Study. Headache 38(2):87-96.

20. Lipton R.B. and Bigal M.E., 2007: Ten lessons on epidemiology of migraine. Headache Suppl 1:S2-9.

21. Malmgren H., Borga M. and Niklasson L., 2000: Artificial Neural Networks in Medicine and Biology: Proceedings of the ANNIMAB-1 Conference. New York: Springer.

22. Marks D.R. and Rapoport A.M., 1997: Practical evaluation and diagnosis of headache. Semin Neurol. 17(4):307-12.

23. Mitchell T., 1997: Machine Learning. Columbus: Mcgraw-Hill; 81-108.

24. Nguyen N. and Cripps, A. 2001: Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks. The Journal of Real Estate Research, 22(3): 313-36.

25. Mongini F., Deregibus A., Raviola F. and Mongini T., 2003: Confirmation of the distinction between chronic migraine and chronic tension-type headache by the McGill Pain Questionnaire. Headache 43(8):867-77.

26. Palm W., 2008: A Concise Introduction to MATLAB. Columbus: McGraw-Hill.

27. Palma Neto L.G. and Nicoletti M.C., 2005: Introdução às Redes Neurais Construtivas. São Carlos: Editora da Universidade Federal de São Carlos. Portuguese

28. Pan American Health Organization, 2007: Renewing primary health care in the Americas. Washington DC: Pan American Health Organization.

29. Queiroz L.P., Barea L.M. and Blank N., 2006: An epidemiological study of headache in Florianopolis, Brazil. Cephalalgia 26(2):122-7.

30. Rasmussen B.K., Jensen R. and Olesen J. 1991: Questionnaire versus clinical interview in the diagnosis of headache. Headache 31(5):290-5.

31. Ridsdale L., Doherty J., McCrone P. and Seed P., 2008: A new GP with special interest headache service: observational study. Br J Gen Pract. 58(552):478-83.

32. Rojas R. and Feldman J., 1996: Neural Networks: A Systematic Introduction. New York, Springer: 149-180.
33. Setiono, R.; Leow, W. K.; Zurada, J. M. Extraction of Rules from Artificial Neural Networks for Nonlinear Regression. IEEE Transactions on Neural Networks, v. 13, n. 3, 2002.
34. Sheftell F.D., Cady R.K., Borchert L.D., Spalding W. and Hart C.C., 2005: Optimizing the diagnosis and treatment of migraine. J Am Acad Nurse Pract. 17(8):309-17.
35. Steiner, M. T. A.; Soma, N. Y.; Shimizu, T.; Nievola, J. C.; Steiner Neto, P. J. Influence of Exploratory Data Analysis. International Journal of Operations and Quantitative Management (IJOQM), v. 12, n. 1, p. 73-83, 2006.
36. World Health Organization, 1978: Primary health care: report of the International Conference on Primary Health Care, Alma-Ata, USSR, 6–12 September (1978) Geneva: World Health Organization.
37. World Health Organization, 2008: The World Health Report 2008: Primary Health Care now more than ever. Geneva: World Health Organization.
38. Yoon M.S., Obermann M., Fritsche G., Slomke M., Dommes P., Schilf C., Diener H.C. and Katsarava Z., 2008: Population-based validation of a German-language self-administered headache questionnaire. Cephalalgia 28(6):605-8.

Table

Table 1. - Tests bests results

	Neurons						Correct	% in
Test	Input	Occult	ANN	Headache	Code	Activ. Function	answers	2177
I	66	9	5		B	logsig/logsig	1590	73.04
II	66	10	1	TTH	B	tansig/logsig	1912	87.83
	66	2	1	MoA	B	tansig/logsig	2058	94.99
	66	8	1	MwA	B	logsig/logsig	2084	98.53
	66	7	1	MOH	B	tansig/logsig	2145	95.73
	66	8	1	Others	B	tansig/logsig	1921	88.24
III	14	10	5		A	tansig/logsig	1451	66.65

Table

Table 2. – Test II best results performance measures

Diagnosis	Provided	True positive	False positive	False negative	True negative	Sensitivity	Specificity	PPV	NPV
TTH	1252	1156	169	96	756	0.92	0.82	0.87	0.89
MoA	307	306	108	1	1762	1.00	0.94	0.74	1.00
MwA	99	99	32	0	2046	1.00	0.98	0.76	1.00
MOH	100	100	93	0	1984	1.00	0.96	0.52	1.00
Others	419	399	236	20	1522	0.95	0.87	0.63	0.99

* PPV= positive predictive values

† NPV= negative predictive values

Table

Table 3 - Kappa índex

Diagnosis	Kappa	p
TTH	0,75	< 0,001
MoA	0,82	< 0,001
MwA	0,85	< 0,001
MOH	0,66	< 0,001
Others	0,68	< 0,001

Figure
Click here to download high resolution image