Original Article

Diagnosis of Common Headaches Using Hybrid Expert-Based Systems

Abstract

Background: Headache is one of the most common forms of medical complaints with numerous underlying causes and many patterns of presentation. The first step for starting the treatment is the recognition stage. In this article, the problem of primary and secondary headache diagnosis is considered, and we evaluate the use of intelligence techniques and soft computing in order to predict the diagnosis of common headaches. **Methods:** A fuzzy expert-based system for the diagnosis of common headaches by Learning-From-Examples (LFE) algorithm is presented, in which Mamdani model was used in fuzzy inference engine using Max–Min as Or–And operators, and the Centroid method was used as defuzzification technique. In addition, this article has analyzed common headache using two classification techniques, and headache diagnosis based on a support vector machine (SVM) and multilayer perceptron (MLP)-based method has been proposed. The classifiers were used to recognize the four types of common headache, namely migraine, tension, headaches as a result of infection, and headaches as a result of increased intra cranial presser. **Results:** By using a dataset obtained from 190 patients, suffering from primary and secondary headaches, who were enrolled from a medical center located in Mashhad, the diagnostic fuzzy system was trained by LFE algorithm, and on an average, 123 pieces of If-Then rules were produced for fuzzy system, and it was observed that the system had the ability of correct recognition by a rate of 85%. Using the headache diagnostic system by MLP- and SVM-based decision support system, the accuracy of classification into four types improved by 88% when using the MLP and by 90% with the SVM classifier. The performance of all methods is evaluated using classification accuracy, precision, sensitivity, and specificity. **Conclusion:** As the linguistic rules may be incomplete when human experts express their knowledge, and according to the proximity of common headache symptoms and importance of early diagnosis, the LFE training algorithm is more effective than human expert system. Favorable results obtained by the implementation and evaluation of the suggested medical decision support system based on the MLP and SVM show that intelligence techniques can be very useful for the recognition of common headaches with similar symptoms.

Keywords: *Headache, Learning-From-Example algorithm, multilayer perceptron, recognition, support vector machines*

Monire Khayamnia[1],
Mohammadreza
Yazdchi[2],
Aghile Heidari[3],
Mohsen
Foroughipour[4]

[1]*Department of Mathematics, Payame Noor University, Tehran,* [2]*Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan,* [3]*Department of Mathematics, School of Mathematics, Mashhad Payame Noor University,* [4]*Department of Neurology, Faculty of Medicine, Neurology School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran*

Introduction

The need for improved efficiency in the use of diagnostic systems has long been documented. This need for selectivity has been identified clearly for illnesses with close symptoms. Headache is one of the most common reasons for neurological consultation and has many causes and symptoms. There are two categories of headache: primary and secondary headaches. Primary headache is not associated with other diseases and those diseases that are not due to other underlying health problems, such as migraine and tension headache. Secondary headache is caused by associated diseases, which have a separate cause

such as headaches as a result of infection and headaches as a result of increase in intracranial pressure (ICP).[1-6] The recognition of headache would be possible through its signs and symptoms, but due to similar symptoms and markers, it could lead to some mistakes in the recognition process of low-experienced medicines or doctors. In addition to pain specification checking, most of the doctors would precede clinical palpation, which is comprised of blood pleasure and pulse checking, complete central nervous system checking, and ophthalmoscopy. In addition to the mentioned actions, sometimes, it is necessary to proceed some specific recognition approaches, such as computed tomography scan, magnetic resonance

Address for correspondence:
Dr. Monire Khayamnia,
Payame Noor
University, Tehran, Iran.
E-mail: mmmkhayam@gmail.
com

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.JMSS_47_18

Quick Response Code:

www.SID.ir

Archive of SID

imaging, electroencephalography (EEG) (EEG registration), visibility scrutinizing, and blood testing for recognizing of pain reasons. These sorts of experiments would be prescribed for putting away the other reasons of the pain.[6-8] For recognition of headaches, fuzzy system could play a valuable role.[9-15] For attaining fuzzy If-then rules, two possible approaches are in access. The first one is receiving from human expert and the second one is fuzzy system production by automatic learning approaches. It could be possible which the received If-then rules from human expert were not complete or these human experiences could not cover the whole items, and due to this reason, the systems which were based on human expert experience are assumed to be of low level of validity. In recent years, for producing fuzzy If-then rules, many approaches have been offered from training data.[16-18] One of the existing approaches is the Learning-From-Example (LFE) algorithm.[18]

The use of classifier systems in medical decision support systems is increasing gradually and has shown great potential in medical diagnosis. Classification systems can minimize errors in the recognition of disease in shorter time and improve the diagnostic accuracy. Multilayer perceptron (MLP) neural networks and support vector machines (SVM) are being used in medical diagnosis.[19,20] Artificial neural networks (ANNs) have been used in the recognition of urological dysfunctions, heart disease, and psychiatric disorders.[21-23] In urology, prostate cancer serves as a good example for working out with ANNs.[24] SVM has been used successfully as an learning method for classification and for the diagnosis of erythemato-squamous diseases and breast cancer diagnosis.[25,26]

The objective of this study was to evaluate the sensitivity and specificity of three different diagnostic strategies: a fuzzy expert system which the If-Then rules have been reached from the LFE algorithm, MLP neural networks classifier, and SVM classifier. In addition, we also compared the ability of these strategies to achieve an absolute diagnostic test accuracy of >88%.

The rest of the article is organized as follows. Section 2 describes the methods; summarizes the related work on common headache diagnosis; describes LFE algorithm, basic MLP neural networks concepts, and the method of SVM. The experimental results are reported in Section 3. The methodology and experiments are discussed in Section 4. Finally, the conclusions are summarized in Section 5.

Methods

We reviewed our database of patients who are suffering from headache and presented to Ghaem Hospital in Mashhad city, Iran, from April 2014 to January 2016. In fact, the data were clearly collected through clinical data that were presented from patients' medical records. Patients eligible for inclusion were consecutive adults (>10 years) with symptoms based on headache. All patients were evaluated and screened for study eligibility by the first author (ENE) prior to the study entry. This was a convenience sample of patients with headache; the patients were enrolled when the first author was present in the physicians' office. According to our primary study, it was estimated that the system could have the diagnostic ability of 76% for a sample by thirty members; in other words, the estimated "*p*" factor is 0.76. Then, sample society volume was calculated using Cochran formulation (due to unknown society volume) and finally obtained $n = 190$. The Cochran formulation (on the time which the society volume is unknown) is $n = \dfrac{z_{\alpha/2}^2 \, p\,(1-p)}{d^2}$, in which the $z_{\alpha/2}^2$ depends on the current distance of certainty and the error of α. Whenever the error level is considered 0.05, the certainty level would be 0.95, and consequently the $z_{\alpha/2}^2 = 1.96$ and $d = 0.05$ could be gained.

Related work on common headache diagnosis

Ambiguity and uncertainty in medicine knowledge is explicit and clear, which is related to the modeling of medicine knowledge. The fuzzy logic has a good power for describing the enigmatic and imprecise aspects and due to this reason, this tool could be used for the system modeling.

A fuzzy expert system with 12 input parameters and 4 output parameters [Table 1] and 237 If-Then rules from human expert was used for the recognition of common headaches, Mamdani model was used in fuzzy inference

Table 1: Used diagnostic and linguistic parameters in regulations

Diagnosis parameter	Linguistic parameter
x_1=Fever	Yes/no (1/0)
x_2=Diplopia	Yes/no (1/0)
x_3=Convulsion	Yes/no (1/0)
x_4=Vomiting	Yes/no (1/0)
x_5=Aura	Yes/no (1/0)
x_6=Severing by special smell	Yes/no (1/0)
x_7=Improving with inhalation	Yes/no (1/0)
x_8=Headache site	One sided, both sides, entire head
x_9=Headache quality	Throbbing, not throbbing
x_{10}=Headache intensity	Low and moderate, severe, and highly severe
x_{11}=Headache duration	Fewer than 4 h, from 4 h to 72 h, from 72 h to 4 weeks, from 4 weeks to months
x_{12}=Headache history	Some days, some weeks, some months, some years
y=Types of headache	Migraine, tension, headaches as a result of infection, headaches as a result of IICP

IICP – Increased intracranial pressure

www.SID.ir

engine using Max–Min as OR–AND operators, and Centroid method was used as a defuzzification technique. Accuracy of the fuzzy expert system was 80% by using data obtained from 190 patients.[12] It is very likely that an expert may not be able to express his or her knowledge explicitly and accurately. Another approach is to generate fuzzy rules through a machine-learning process. In recent years, for producing fuzzy if-then regulation, many approaches have been offered from training data.[16-18] One of the existing approaches is the LFE algorithm.[18]

Learning-From-Example algorithm

Wang and Mendel proposed a five-step algorithm for generating fuzzy rules by learning from examples.[18]

In LFE algorithm, Step 1 divides the input and output spaces into fuzzy regions; Step 2 generates fuzzy rules from the given desired input–output data pairs; Step 3 assigns a degree to each generated rule; Step 4 forms the combined fuzzy rule base; and Step 5 presents a defuzzifying procedure for obtaining a mapping based on the combined fuzzy rule base.

Nearly 90% of our current existing data is assumed as our training data and 10% of this data is considered as testing data. By training data, the system could be taught and then the If-Then rules could be produced. By testing data, the ability of extension to fuzzy deduction would be evaluated.

It would be assumed that the *A* complex includes the below members:

$$A = \{a_1, a_2, \ldots, a_m\} \tag{1}$$

Inside this complex, each a_i member shows 12 input parameters and 4 output parameters as shown below:

$$a_i = ([x_{i1}, x_{i2}, \ldots x_{i12}], y_i) \tag{2}$$

At first, the degree of each parameter $x_{i1}, x_{i2}, \ldots x_{i12}, y_i$ in the different fuzzy groups should be calculated, and then the system would dedicate a group by the maximum degree to each parameter and by this approach, one rule for each pair of input–output data would be achieved. For example, the membership function of severity of headache is shown in Figure 1.

For the next stage, for each rule, one degree would be considered according to the below formulations:

$$D \text{ (rule } [i]) = \mu(x_{i1}).\ \mu(x_{i2})\ldots \mu(x_{i12}).\ \mu(y_i) \tag{3}$$

If there is more than one rule, in which their assumption part is the same but they are different at their results, the rule which has the higher degree would be considered and then a compressed group of rules would be reached.

Multilayer perceptron neural networks

MLP is one of the most commonly used neural network architectures in medical decision support systems because of its features such as the ability to learn, ease of
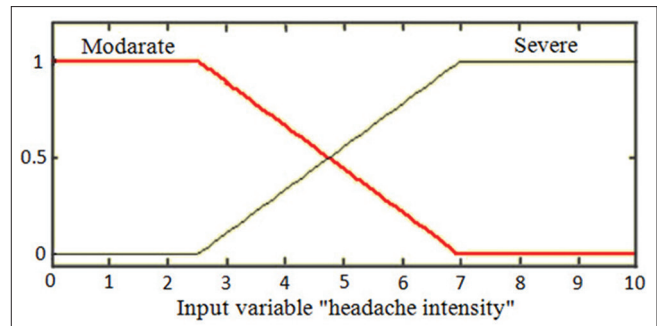

Figure 1: Membership function of severity of headache

implementation, fast operation, and it belongs to the class of supervised neural networks.[27]

A MLP network consists of three or more layers of neurons: an input layer that receives external inputs, one or more hidden layers, and an output layer which generates the classification results [Figure 2]. A jth neuron in a first hidden layer by which MLP computes the weighted sum of the inputs and adds a "bias" term (θj) and transforms this sum through a suitable mathematical "transfer function," and transfers the result to neurons in the next layer. The whole process is defined as follows:

$$\sum_{i=1}^{p} w_{ji}x_i + \theta_j = v_j \tag{4}$$

$$f_j(v_j) = y_j$$

Where $x_1, x_2, \ldots x_p$ are inputs, θ_j is the bias, and w_{ji} is the connection weight between the input x_i and the jth neuron, and f_j is the transfer function of the j th neuron, and y_j is the output. Various transfer functions are available; however, the most common choice of the transfer function is the sigmoid one, as defined in Eq. 5.

$$f(x) = \frac{1}{1+e^{-x}} \tag{5}$$

The learning used is back propagation (BP) algorithm with the adaptive learning rate and the momentum constant. BP is one of the simplest methods for the supervised training of MLP.[27,28] The basic BP algorithm runs as follows:

1. All the connection weights *w* are initialized with small random values from a pseudorandom sequence generator
2. While the error *E* is below a preset value or until the gradient $\partial E(t)/\partial W$ is smaller than a preset value, the following three basic steps are repeated, until to converge:
   - The update are computed by $\Delta W(t) = -\eta \frac{\partial E(t)}{\partial W}$
   - The weights are updated using
     $W(t+1) = W(t) + \Delta W(t)$
   - Compute the error $E(t+1)$.

Where *t* is the iteration number, *W* is the connection weight, and η is the learning rate. The error E can be chosen as the

Archive of SID

mean square error function between the actual output y_j and the desired output d_j:

$$E = \frac{1}{2}\sum_{j=1}^{n_j}(d_j - y_j)^2 \qquad (6)$$

To achieve faster learning and avoid oscillation problems during the search for the minimum value on the error surface, an additional term is used and hence Eq. 7 becomes:

$$\Delta W(t) = -\eta\frac{\partial E(t)}{\partial W} + \alpha\Delta W(t-1) \qquad (7)$$

Where α is the "momentum" term and $0 < \alpha < 1$. The learning rate can be updated using $\eta(t) = \eta_0 + \lambda E(t-1)$, where η_0 is a preset learning rate and $\lambda > 0$

Support vector machine

Vapnik proposed the SVM that has been studied extensively for classification.[29] The SVM algorithm represents an example of a binary classifier which can be generalized to a multiclass classifier. Assume that the classification of the training vectors belongs to two linearly separable classes,

$$(x_i, y_i), x_i \in R^n, y_i \in \{+1, -1\}, i = 1, \ldots, n \qquad (8)$$

Where x_i is a n-dimensional input vector and y_i is a label that determines the class of x_i. SVM is mainly based on a separating hyperplane which is represented by Eq. 9.

$$W.X + b = 0 \qquad (9)$$

Where

W is an orthogonal vector and represents the weight vector

X represents the input vector

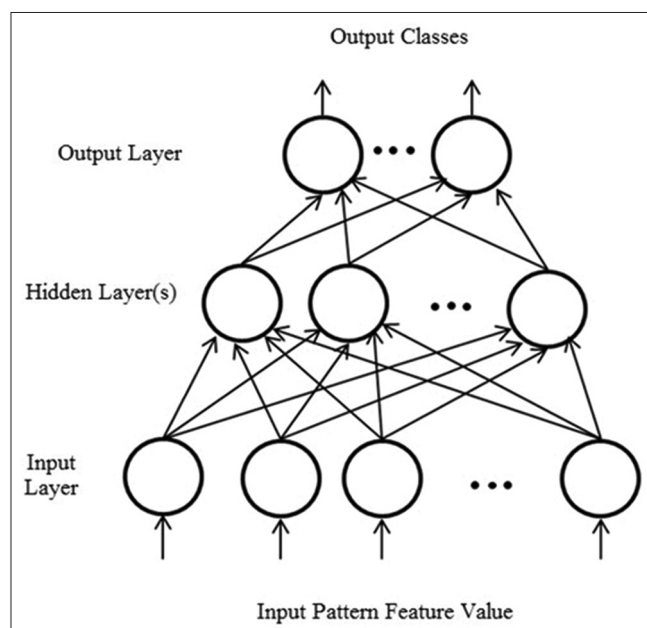b represents a bias or a threshold

The parameters W and b are constrained by

$$\min_i |w.x_i + b| \geq 1 \qquad (10)$$

that in canonical form must satisfy the following relations,

$$y_i(w.x_i + b) \geq 1, i = 1, 2, \ldots, n \qquad (11)$$

A separating hyperplane with a large margin is defined by (12)

$$M = \frac{w}{\|W\|}(x_2 - x_1) = \frac{2}{\|W\|} \qquad (12)$$

Where x_1 and x_2 represents two support vectors. The maximization process of the margin is equivalent with the minimization of the $\|W\|^2$

Hence, if we can separate the data perfectly, then we can optimize the following:

Minimize $\|W\|^2$ $\qquad (13)$

subject to $y_i(wx_i + b) \geq 1$

To deal with the nonseparable case, one can rewrite the problem as:

$$Minimize\ w^2 + C\sum_{i=1}^{n}\xi_i \qquad (14)$$

subject to $y_i.(w.x_i + b) \geq 1 - \xi_i, i = 1, 2, \ldots, n, \xi_i \geq 0$

with a user-defined positive finite constant C.

Results

Between April 2014 and January 2016, 213 patients were assessed for initial eligibility and invited to participate. Figure 3 shows the flow of patients through the study. In total, 190 patients completed the study, with a completion rate of 89%. Of the 190 patients with headache, 133 had migraine headache, 19 had tension headache, 12 had headache as a result of infection, and 26 had headache as a result of increased ICP (IICP).
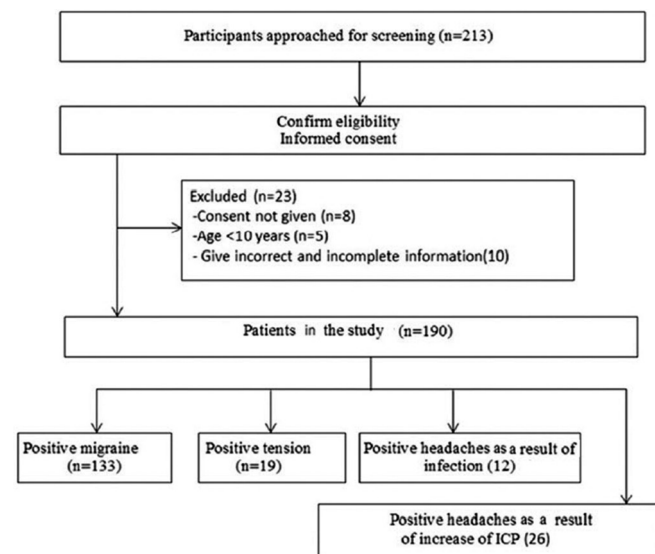

Figure 2: Architecture of a multilayer perceptron network

Figure 3: Flow diagram from a study

www.SID.ir

Archive of SID

The median age of the participants was 32 years (range: 17–65), and 112 participants (58.9%) were female.

Baseline demographic and clinical characteristics of the study participants are shown in Table 2.

The designed fuzzy system by LFE algorithm has been studied and evaluated on 190 medical records collected from patients suffering from four headache diseases, have been used to learn and test the system by 10-fold cross-validation and on an average, 123 rules of fuzzy if-then were attained for the system and showed 82% good agreement. For example, two rules that obtained are: if fever is no, diplopia is no, convulsion is no, vomiting is no, aura is yes, severing by special smell is yes, improving with inhalation is no, headache site is both sides, headache quality is not throbbing, headache intensity is severe and high severe, headache duration is from 4 h to 72 h, and headache history is some month, then the type of headache is migraine.

If fever is yes, diplopia is no, convulsion is no, vomiting is no, aura is no, severing by special smell is yes, improving with inhalation is yes, headache site is both sides, headache quality is throbbing, headache intensity is severe and high severe, headache duration is from 72 h to 4 weeks, and headache history is some days, then the type of headache is headache as a result of infection.

The accuracy, precision, sensitivity, and specificity of this system are presented in Table 3. Overall, the value of Macro F-Score was 0.88.

The MLP used in this study has been designed through the WEKA (at the University of Wyokota in New Zealand) software, which consists of three layers including an input layer, a hidden layer, and an output layer with 12 input variables, 15 hidden neurons, 4 outputs, and the activation function was the sigmoidal function. The BP training parameters in the system are shown in Table 4.

The SVM is a binary classifier which can be extended into a multiclass classifier. For four classes (A–D), four classifiers are necessary; for example, one SVM classifies A from B, C, and D and a second SVM classifies B from A, C, and D. The multiclass classifier output codes for A, B, C, and D have codes namely $(1, -1, -1, -1)$, $(-1, 1, -1, -1)$, $(-1, -1, 1, -1)$, and $(-1, -1, -1, 1)$, respectively.

The training algorithm of the SVMs is based on quadratic programming. The quadratic programming problem in the SVMs was solved by using the MATLAB optimization toolbox.

In this study, the classifiers proposed for medical decision-making were MLP and SVM. To comparatively evaluate the performance of the two classifiers, both of them were trained by the same training data set and tested with the evaluation data set. The data set (190 patients) was divided into two separate data sets – the training data set and the testing data set. We have used the most common measures to evaluate the effectiveness of our method. These measures are classification accuracy, precision, and sensitivity. The cross-validation estimated the accuracy of each MLP test run, and the mean accuracy based on the 10-fold cross-validation method is listed in Tables 5 and 6. The accuracy of the proposed MLP and SVM-based decision support system was 0.88 and 0.90, respectively. Overall, the value of Macro F-Score for MLP and SVM was 0.76 and 0.81, respectively.

Discussion

Headache is almost a universal experience that most of us have some kind of headache at some time in our lives. In fact, headache could be the symptom of many illnesses,

Table 2: Baseline demographic and clinical characteristics of the study	
	n (%)
Patient characteristics	
Median age (range) (in years)	32 (17-65)
Women	112 (58.9)
Presenting symptoms	
Fever	9 (4.73)
Diplopia	2 (1.05)
Convulsion	4 (2.1)
Vomiting	73 (38.42)
Aura	6 (315)
Severing by special smell	70 (36.84)
Improving with inhalation	10 (5.26)
Headache site	One-sided: 66 (34.74), both side: 75 (39.47), entire head: 49 (25.26)
Headache quality	Throbbing: 99 (52.11), not throbbing: 90 (43.37)
Headache intensity	Low and moderate: 42 (22.10), severe and highly severe: 145 (76.31)
Headache duration	Fewer than 4 h: 98 (51.58), from 4 h to 72 h: 63 (32.81), from 72 h to 4 weeks: 22 (11.58), from 4 weeks to months: 7 (3.68)
Headache history	Some days: 31 (16.31), some weeks: 13 (6.84), some months: 16 (8.42), some years: 127 (66.84)

www.SID.ir

Archive of SID

Table 3: Effectiveness of Learning-From-Examples method

	Accuracy	Precision	Sensitivity	Specificity
Migraine	0.92	0.90	0.94	0.81
Tension	0.74	0.92	0.84	0.94
Headaches as a result of infection	0.98	0.97	0.74	0.98
Headaches as a result of ICP	0.96	0.96	0.79	0.98

ICP – Intracranial pressure

Table 4: The training parameters

The training parameters	Value
Momentum constant	0.2
Preset learning rate	0.3
Training time	500
Validation threshold	20

Table 5: Effectiveness of multilayer perceptron method

	Accuracy	Precision	Sensitivity	Specificity
Migraine	0.92	0.93	0.96	0.81
Tension	0.92	0.58	0.58	0.95
Headaches as a result of infection	0.98	1	0.5	1
Headaches as a result of ICP	0.95	0.81	0.81	0.97

ICP – Intracranial pressure

Table 6: Effectiveness of support vector machine method

	Accuracy	Precision	Sensitivity	Specificity
Migraine	1	0.99	1	0.99
Tension	0.93	0.80	0.65	0.98
Headaches as a result of infection	0.97	0.82	0.80	0.98
Headaches as a result of ICP	0.90	0.70	0.75	0.93

ICP – Intracranial pressure

and it could only be assumed to be a disease and due to these reasons, the headaches could be so varied and populated. Due to the similar symptoms of headaches, it can cause some faults and mistakes for low-experienced doctors. Because of these faults and mistakes in the recognition system, the systems which are using the current knowledge can support the doctors' duties, which would be so important and valuable. In this article, we have designed a fuzzy system by using the LFE algorithm, and we have presented a medical decision support system based on the MLP neural network architecture and multiclass SVM for common headache diagnosis. By noticing the importance of on-time recognition and gaining pleasant results, the results of these systems are compared. A headache diseases database consisting of 190 cases has been used in this study and 10-fold cross-validation has been applied to assess the

generalization of these systems. Our two main hypotheses were that all quality variables such as headache intensity, fever, and severing by special smell can change to quantity variables, and with 12 features, it has been attempted to design intelligent systems to the recognition of common headaches.

Conclusions

This study aimed at recognizing, predicting, and diagnosing prevalent headaches by soft-computation approaches and designing intellectual systems. Unlike hard-computation approaches, in soft-computation style, the contrived plans which have been produced from humankind would be dedicated for solving the problems.

In this study, in order to classification and diagnosis of the four types (migraine, tension, headaches as a result of infection, and headaches as a result of IICP) of headache by the usage of features that attain from patients' medical records, were introduced a fuzzy system without the usage of genius person's knowledge and through the LFE algorithm and the validity amount resulted to 82%. In addition, two types of classifiers (MLP and SVM) were implemented. We have applied cross-validation methods to assess the generalization of the system. It was observed that the accuracy of system was 88% and 90% for MLP and SVM, respectively.

The results show that the multiclass SVM and the MLP neural network can achieve very high diagnostic accuracy of 90% and 88%, respectively, and demonstrate that the multiclass SVM and MLP can be used in the diagnosis of the common headaches by taking into consideration the misclassification rates and prove its usefulness and ability in medical decision support systems.

By consideration of on-time diagnoses and existence of some deficiencies for received linguistic rules from experts, LFE system as a decision-making helper would be used for diagnosing headache and headache sorts in hospitals because better results have been obtained by the application of LFE learning algorithm, and the operation of SVM model has been in better condition rather than that of MLP algorithm.

Financial support and sponsorship

Conflicts of interest

There are no conflicts of interest.

References

1. Ravishankar K. The art of history-taking in a headache patient. Ann Indian Acad Neurol 2012;15:S7-14.
2. Steiner TJ, Birbeck GL, Jensen RH, Katsarava Z, Stovner LJ, Martelletti P, *et al.* Headache disorders are third cause of disability worldwide. J Headache Pain 2015;16:58.
3. William C, Shiel Jr, MD, FACP, FACR, Headache Symptoms &

www.SID.ir

Signs, 2015.

4. Cafasso J. Migraine Symptoms, Medically Reviewed by Steven Kim. Health Line, MD; 20 May, 2015.

5. Forsyth PA, Posner JB. Headaches in patients with brain tumors: A study of 111 patients. Neurology 1993;43:1678-83.

6. Dalessio DJ. Diagnosing the severe headache. Neurology 1994;44:S6-12.

7. Evans RW. Diagnostic testing for headache. Med Clin North Am 2001;85:865.

8. Godwin A, Villa J. Acute headache in the ED: Evidence-based evaluation and treatment options. Emerg Med Pract 2001;3:1-32.

9. Sikchi SS, Sikchi S, Ali MS. Fuzzy expert systems (FES) for medical diagnosis. Int J Comput Appl (0975-8887) 2013;63.

10. Yao JF, Yao JS. Fuzzy decision making for medical diagnosis based on fuzzy number and compositional rule of inference. Fuzzy Sets Syst 2001;120:351-66.

11. Sanchez E. Truth-qualification and fuzzy relations in natural languages, application to medical diagnosis. Fuzzy Sets Syst 1996;84:155-67.

12. Khayamnia M, Yazdchi M, Heidari A, Foroughipour M. Fuzzy expert system for diagnosis of common headaches with similar symptoms. MJMS 2017;5:1680-4.

13. Kim YH, Kim SK, Oh SY, Ahn JY. A fuzzy differential diagnosis of headache. J Korean Data Inf Sci Soc 2007;18:429-38.

14. Ahn JY, Kim YH, Kim SK. A fuzzy differential diagnosis of headache applying linear regression method and fuzzy classification. IEICE Trans Inf Syst 2003;E86-D:2790-3.

15. Ahn JY, Mun KS, Kim YH, Oh SY, Han BS. A fuzzy method for medical diagnosis of headache. IEICE Trans Inf Syst 2008;E91:1215-7.

16. Nomura H, Hayashi I, Wakami N. A learning method of fuzzy inference rule\ by descent method. IEEE International Conference on Fuzzy Systems. San Diego: S~wcw.s; 1992. p. 203-10.

17. Burkhardt DG, Bonissone PP. Automated fuzzy knowledge base generation and tuning, IEEE International Conference on Fuzzy Systems. San Diego: CA, USA; 1992. p. 179-88.

18. Wang LX, Mendel JM. Generating fuzzy rules by learning from examples. IEEE Trtrrls Syst Morl CJhrrI7VI 1992;22:1414-27.

19. Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A, Havel J. Artificial neural networks in medical diagnosis. J Appl Biomed 2013;11:47-58.

20. Bo G, Xianwu H. SVM multi-class classification. J Data Acquis Process 2006;3:47-52.

21. Gil D, Johnsson M, Chamizo JM, Paya AS, Fernandez DR. Application of artificial neural networks in the diagnosis of urological dysfunctions. Expert Syst Appl 2009;36:5754-60.

22. Yan H, Jiang Y, Zheng J, Peng C, Li Q. A multilayer perceptron-based medical decision support system for heart disease diagnosis. Expert Syst Appl 2006;30:272-81.

23. Peled A. Plasticity imbalance in mental disorders the neuroscience of psychiatry: Implications for diagnosis and research. Med Hypotheses 2005;65:947-52.

24. Batuello JT, Gamito EJ, Crawford ED, Han M, Partin AW, McLeod DG, *et al.* Artificial neural network model for the assessment of lymph node spread in patients with clinically localized prostate cancer. Urology 2001;57:481-5.

25. Übeyli ED. Multiclass support vector machines for diagnosis of erythemato-squamous diseases. Expert Syst Appl 2008;35:1733-40.

26. Akay MF. Support vector machines combined with feature selection for breast cancer diagnosis. Expert Syst Appl 2009;36:3240-7.

27. Bishop CM. Neural Networks for Pattern Recognition. Oxford: Oxford University Press; 1995.

28. Duda RO, Hart PE, Strok DG. Pattern Classification. New York: Wiley; 2001.

29. Vapnik V. The Nature of Statistical Learning Theory. New York: Springer-Verlag; 1995.

BIOGRAPHIES

Monire Khayamnia Ph.D. of Applied Mathematics, Tehran Payame Noor University, Tehran, Iran

Email: mmmkhayam@gmail.com

Aghile Heidari Profesor, Department of Mathematics, School of Mathematics, Mashhad Payame Noor University, Mashhad, Iran

Email: aghileheydari@yahoo.com

Mohammadreza Yazdchi Associate Professor, Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran

Email: yazdchi@eng.ui.ac.ir

Mohsen Foroughipour Associate Professor, Neurology school of medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Email: Foroughipourm@mums.ac.ir

www.SID.ir