Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2017) 25: 4200 – 4210

doi:10.3906/elk-1612-178

An ant colony optimization algorithm-based classification for the diagnosis of primary headaches using a website questionnaire expert system

Ufuk ÇELİK[1,*], **Nilüfer YURTAY**[2]

[1]Department of Computer Programming, Gönen Vocational School, Bandırma Onyedi Eylül University, Balıkesir, Turkey

[2]Department of Computer Engineering, Faculty of Computer and Information Sciences, Sakarya University, Adapazarı, Turkey

Received: 15.12.2016 • **Accepted/Published Online:** 27.03.2017 • **Final Version:** 05.10.2017

Abstract: The purpose of this research was to evaluate the classification accuracy of the ant colony optimization algorithm for the diagnosis of primary headaches using a website questionnaire expert system that was completed by patients. This cross-sectional study was conducted in 850 headache patients who randomly applied to hospital from three cities in Turkey with the assistance of a neurologist in each city. The patients filled in a detailed web-based headache questionnaire. Finally, neurologists' diagnosis results were compared with the classification results of an ant colony optimization-based classification algorithm. The ant colony algorithm for diagnosis classified patients with 96.9412% overall accuracy. Diagnosis accuracies of migraine, tension-type, and cluster headaches were 98.2%, 92.4%, and 98.2% respectively. The ant colony optimization-based algorithm has a successful classification potential on headache diagnosis. On the other hand, headache diagnosis using a website-based algorithm will be useful for neurologists in order to gather quick and precise results as well as tracking patients for their headache symptoms and medication usage by using electronic records from the Internet.

Key words: Data mining, electronic health records and systems, clinical decision support, headache disorders, ant colony optimization algorithm

1. Introduction

Headache diagnosis using computer sciences attracted computer scientists' attention after the definition of headache criteria by the International Headache Society (IHS) in 1988. It became more popular especially after the publication of The International Classification of Headache Disorders Second Edition (IHCD-2) [1] in 2004. According to the criteria, primary headaches are diseases without any underlying organic etiology, which distinguishes them from secondary headaches. They were classified in four types as migraine, tension-type, cluster-type, and other primary types of headache. Each type of headache has different subgroups of classification. The criteria of subgroups show similarities with each other, which seems complicated and puzzling. Therefore, doctors may experience misdiagnosis due to the limited consultation time for each patient at hospitals. In Turkey, 54% of the doctors indicated that error in the diagnosis of migraine is because of the density of the patients in hospitals [2]. There have been various computerized headache diagnosis studies [3–8] based on headache criteria so far by using different machine learning algorithms. However, the unique point of the present study is that it makes use of the ant colony optimization (ACO)-based classification algorithm,

*Correspondence: ucelik@bandirma.edu.tr

which has been a crucial heuristic algorithm recently, and it has not been used for headache classification before. The ACO algorithm was implemented first by Marco Dorigo [9] and it was inspired by ants' foraging process of determining the shortest way to the food source. The ACO is a well-known optimization algorithm. However, it can be used for the optimization of attribute selection for the creation of classification rules. Hence, basically the classification with the ant colony algorithm aims to extract the best rules for a class among all possible rules by imitating ants' search for the shortest path to food.

2. Related work

There have been several studies of computerized headache diagnosis like statistically clinical record program [10] or headache record software by using Microsoft Access database [11]. However, while those studies did not use any artificial algorithms, our study evaluated a web-based headache diagnosis expert system using the ant colony algorithm for the first time that also provides medical records of patients for neurologists. While Kopec et al. [3] and Pryse-Philips et al. [12] implemented algorithms only for the migraine diagnosis by using a rule-based algorithm and decision-tree algorithm, respectively, we evaluated an artificial intelligence system for all primary headaches. Maizel and Wolfe [13] correctly identified episodic migraine, tension type, and cluster headache. However, Sarchielli et al. [14] diagnosed specifically primary chronic headaches. Although some studies claim that they successfully classified all primary headaches, they obviously have problems related to sample size and accuracy. For example, Krawcyk et al. mentioned 80% consistency in their research [4] using various machine learning methods. As another example, Simić [5] evaluated a rule-based fuzzy logic diagnosis system for only 80 patients in order to classify migraine and tension type but not for cluster headaches. Simić et al. also presented another solution about case-based and rule-based decision support in headache disorder [15]. Another research explained a guideline-based headache diagnosis and the system gathered more than 90% of accuracy results [16] and discussed the validation of the system in another study [17]. Another approach for the diagnosis of headache was to use interval-valued intuitionist fuzzy sets and aggregate operator [18,19]. Zimming et al. developed a decision support system headache diagnosis but only for probable migraine and probable tension-type headaches [20]. Another study showed the comparison of machine learning classifiers for a headache dataset [21].

There is a specific study with the same subject and nearly the same method [22]. This method uses the ACO algorithm but it classifies headaches into three classes by a clustering approach. They obtained 89.2%, 84.3%, and 85.7% accuracy for migraine, tension-type headache, and trigeminal autonomic cephalalgias, respectively. Walters et al. developed a four-item migraine screening algorithm among a nonclinical sample that had 1829 participants and they achieved 93% area under curve [23]. Finally, two similar studies from the same authors using the artificial immune system algorithm achieved a success continuum ranging from 95% to 99% [24,25].

3. Objectives

Computer-based decision support systems have been widely used in many areas of medical science [26,27] as well as headache diagnosis [28]. This kind of application provides a better solution for data collection. Moreover, they make a prediction of disorders by using machine learning methods that are quite powerful for classification of the diseases. In addition, several attempts have been made to diagnose headache disorders by using computer-assisted software but the IHS has not approved any system so far. There has not been any software that is fully capable of diagnosing headaches according to IHS criteria.

In this study, we aimed to collect information of patients who suffer from primary headache disorders using a website-based survey. We designed a website with a calendar entry for patients' symptoms and medication usage during their headache attacks. After collecting the information, patients' headache types were predicted by using the ACO-based classification algorithm and the results were obtained.

4. Methods

We performed this research in 850 patients who went to the doctor because of headaches. Informed consent was obtained from all patients via registration for the website questionnaire and the requirement for written informed consent was waived by the investigational review board of Balıkesir University Medical Faculty.

The patients were of both sexes (70% female, 30% male) and they were from three cities in Turkey with ages between 15 and 65. The participants responded to a questionnaire hosted on the www.migbase.com website with the assistance of doctors in the consultation by using a tablet or computer. Doctors are supposed to enter their personal opinions for diagnosis in the web system following the patient's completion of the questionnaire. The questionnaire contains headache-related questions such as duration, severity, frequency, localization, aggravation, characterization, nausea, vomiting, phonophobia, photophobia, and aura symptoms as well as general questions like sex, age, and smoking. There are 40 attributes gathered from questions for the classification. This system was set up using the MySQL database and PHP programming language and a general view of the system is shown in Figure 1.

Doctors' classifications of headaches were compared with the results of the ACO-based classification algorithm. The neurologists' diagnosis results are provided in Table 1.

Table 1. Headache diagnosis result from neurologist examination.

Headache types	Number of patients	Percentage
Migraine	609	71.65%
Tension type	185	21.76%
Cluster	56	6.59%

According to the results, migraine (71.65%) seems to be the most commonly suffered headache type, whereas cluster-type of headache (6.59%) takes third place in terms of primary headache prevalence and there is only one patient who has no headache. The ant colony algorithm implemented classification through a learning stage by using a training dataset derived from research data using 10-fold cross validation [29]. This validation process provides variability in the training dataset crucial for a successful classification.

The ACO-based classification algorithm process is basically discovering of classification rules adjusted in the form of "if–then" rules like "IF <condition> THEN <class>". Also neurologists diagnose headaches similar to the "if–then" context according to ICHD-2 criteria. The ACO classification algorithm generates rules by simulating ants' foraging. Ants make a trail along the path (Figure 2a) between the nest and food by leaving a chemical residue called pheromone on the ground. This trail leads other ants to the food by following the path. During this travel the pheromone level decreases if the path is long. The ants that found the shortest way leave more pheromone on the route (Figure 2b). Finally, ants are forced to use the shortest way with the guidance of pheromone level (Figure 2c) [30].

The first application of ACO-based data mining for classification rules was implemented by Parpinelli et al. [31]. Each ant creates a rule by using an attribute term $<V_i = \text{Value}_k>$ from the data instances. This rule is based on the pheromone function (τ (t)) and the heuristic value (η) for each term. The ant continues

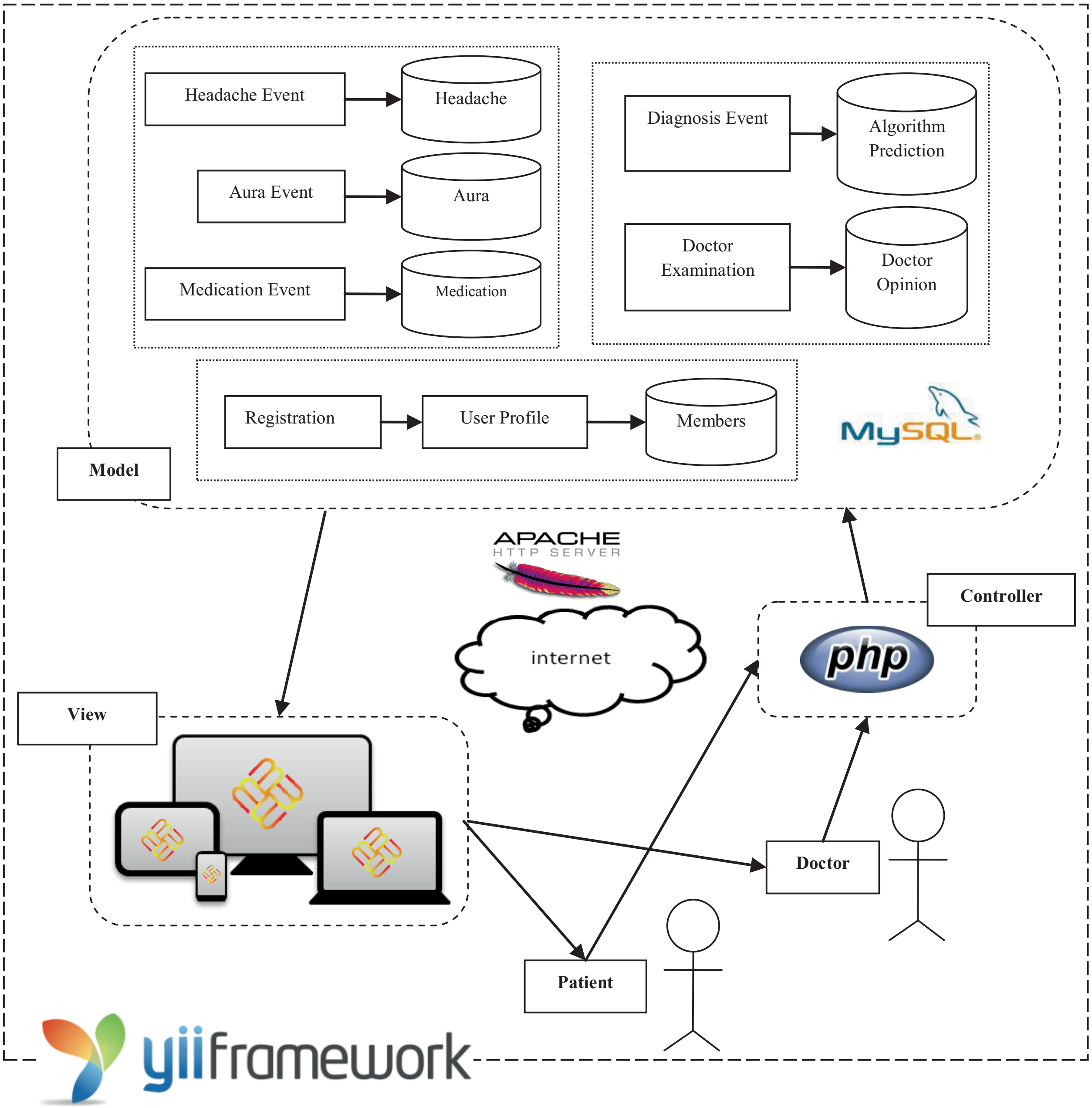

Figure 1. Headache tracking and diagnosis system.

Table 2. Confusion matrix of classification for ACO algorithm.

		Patients according to physician's examination			Total
		Migraine	Cluster	Tension	
Patients according to software diagnosis classification	Migraine	598	1	7	606
	Cluster	3	55	7	65
	Tension	8	0	171	179
	Total	609	56	185	

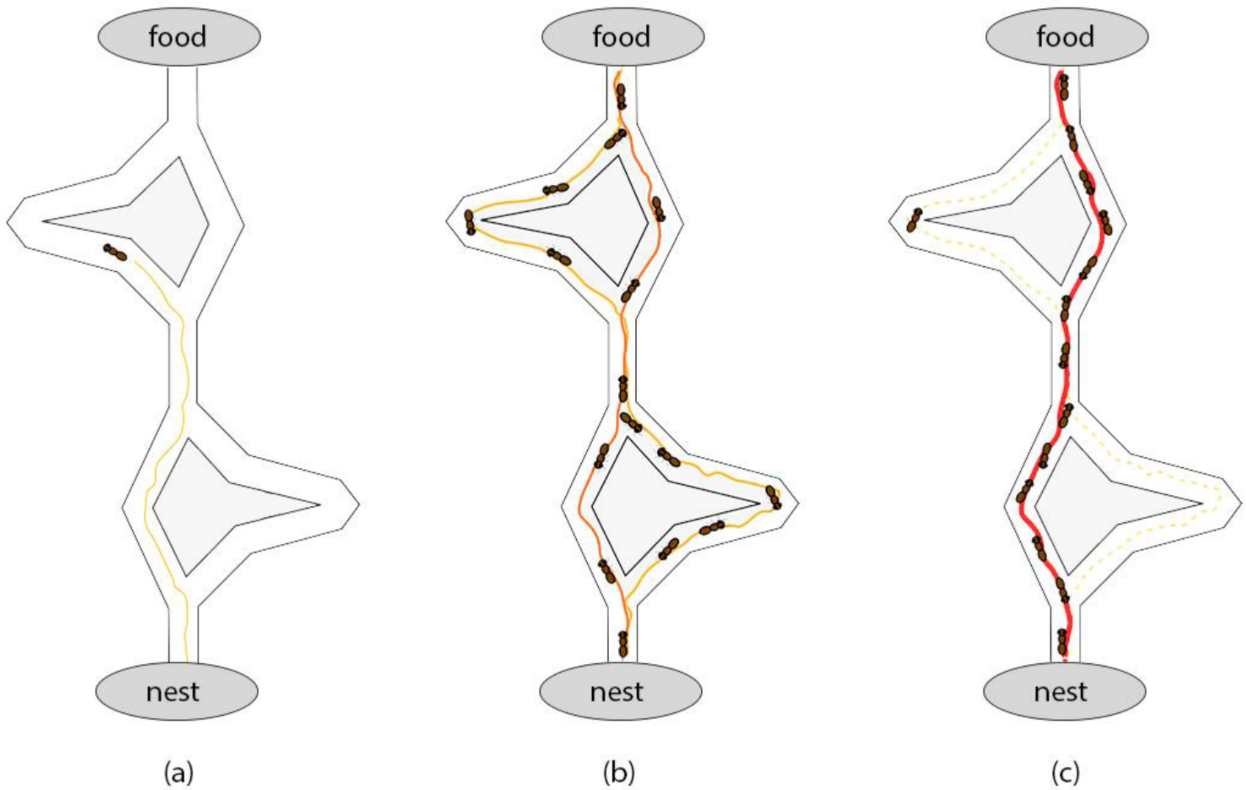

Figure 2. Foraging process of ants.

to add one term at a time unless all attributes are not used or the number of rules for convergence is smaller than the user-defined minimum value. The created rule is pruned and pheromone levels are updated because of evaporation or increasing on the trails followed by ants. Another ant starts to create a new rule according to new pheromone amount. This process is repeated unless the number of rules is greater than the user-defined threshold value or ants converge on the same rule like previous ones. The best rule within the created rules is stored in the discovered rules array and the training cases matched by this rule are eliminated from the training set. All this process continues until the number of uncovered cases is lower than the training set [32]. The flow diagram of the ACO-based algorithm [31] for classification rules is presented in Figure 3.

The ACO-based classification algorithm process can be represented as a construction graph, which is shown in Figure 4, illustrating the example of migraine diagnosis.

As shown in Figure 4, an ant tries all possibilities of the attributes for the headache class and calculates pheromone levels between the values of attributes according to which the best classification is determined.

5. Results

In this study, we did not examine the subgroups for the main class levels of headaches (migraine, tension-type, and cluster-type). We changed numeric attributes such as duration to nominal. Therefore, we collected duration information as intervals. We utilized 10-fold cross validation for classifications and we gathered the results shown in Tables 2 and 3 below after the evaluation of the ant colony algorithm. We obtained the best results in the ant colony algorithm by employing 50 ants, adjusting 4 for the parameter of minimum cases per rule, 10 for maximum uncovered cases, 20 for convergence of rules, and implementing 500 iterations.

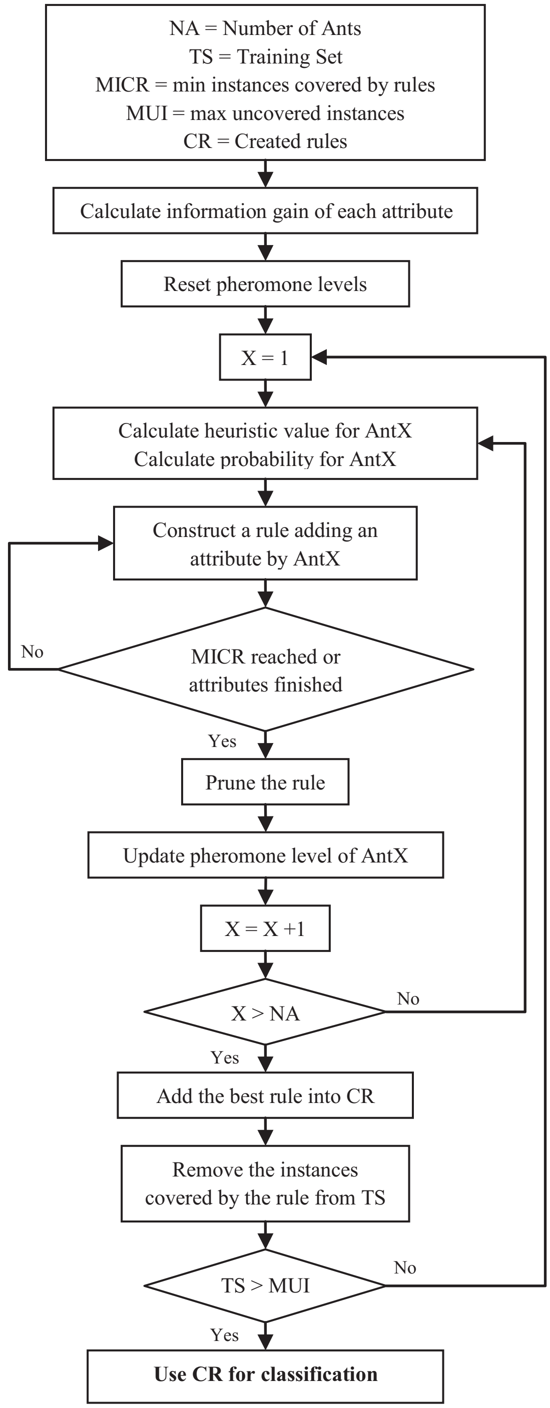

Figure 3. Flow diagram of ACO-based classification algorithm.

In this classification, 26 patients were misdiagnosed by the ant colony classification for headache diagnosis, leading to 96.9412% accuracy (Table 4). As can be seen from the ROC analysis [33] results, the headache diagnosis system that was developed using the ant colony algorithm showed a successful performance in classifying headache diagnosis. The rules created by the ant colony algorithm for headache diagnosis:

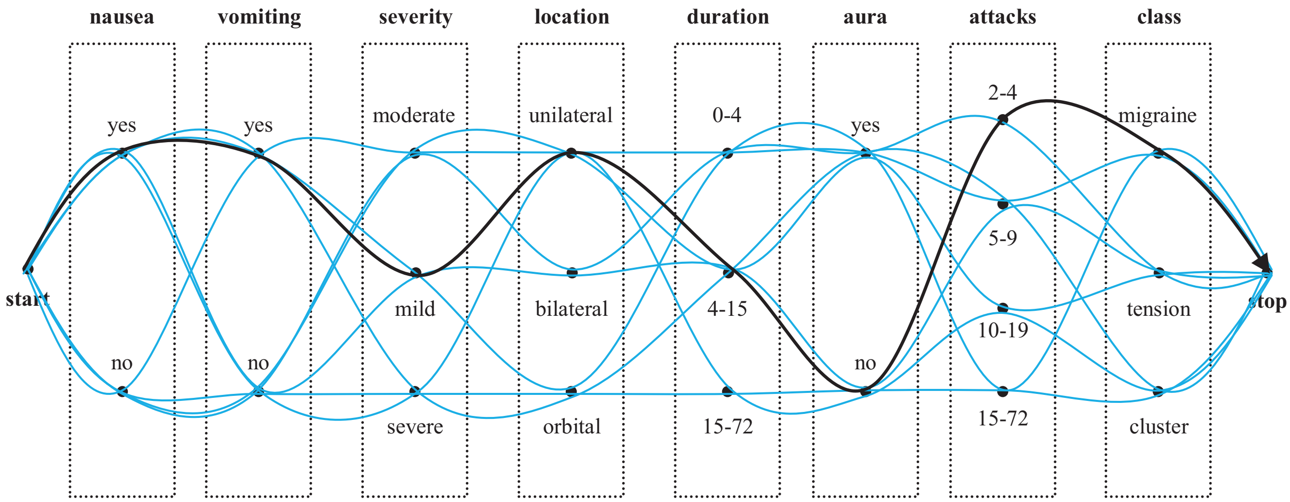

Figure 4. Example path of migraine type headache diagnosis described by an ant.

Table 3. Detailed accuracy of classification by class for ACO algorithm.

Class	True positive rate	False positive rate	Precision	f-measure	ROC area	Accuracy (%)
Migraine	0.982	0.033	0.987	0.984	0.974	98.2
Cluster	0.982	0.013	0.846	0.909	0.985	98.2
Tension	0.924	0.012	0.955	0.940	0.956	92.4

Table 4. Overall classification performance of ACO algorithm.

	Patients	Patient ratio (%)
True diagnosed	824	96.9412
False diagnosed	26	3.0588

- IF characterization = 'pulsating' AND nasal_congestion = 'no' THEN 'migraine'
- IF severity = 'severe' AND aggravation = 'no' AND hemiplegic = 'no' THEN 'cluster'
- IF location = 'bilateral' AND photophobia = 'no' THEN 'tension'
- IF severity = 'moderate' AND nausea = 'no' AND agitation = 'no' THEN 'tension'
- IF photophobia = 'yes' AND phonophobia='yes' THEN 'migraine'
- Default rule: migraine

6. Discussion

Medical applications with the support of computer algorithms have increased rapidly and gained a serious reputation because of remarkable results. There have been many studies of headache diagnosis as shown in Table 5.

Some of these systems did not explain the number of attributes or patients [5,6,8,11,15]. Moreover, the accuracy rate is not clear in some of them [5,8,10,11,19]. There are significant studies that achieved great

Table 5. The performance comparison of different headache diagnosis systems.

Authors	Patient	Attribute	Headache types	Methods	Accuracy
Kopec et al. [3]	6	10	Migraine	CLIPS language	75–97
Krawczyk et al. [4]	1022	19	Migraine Tension-Type Other	Naive Bayes, C4.5, Support Vector Machine, Bagging, Boosting, Random Forest	72.2, 76.34, 76.51, 79.97, 76.68, 78.24
Simic et al. [5]	80	?	Migraine, Tension-Type, Other	Rule Based Fuzzy Logic	?
Andrew et al. [6]	68 + 54	?	Migraine, Tension-Type, Other	Structured Headache Diagnosis Interview	91
Hasan et al. [8]	?	40	Migraine, Tension-Type, Cluster, Other	Decision Tree	?
Gallai et al. [10]	500	15	Migraine, Tension-Type, Cluster	No algorithm. Just Structured Recording	?
Simone et al. [11]	?	?	?	Microsoft Access Based Recording system	?
Pryse-Philips et al. [12]	461	45	Migraine	Classification and Regression Trees	91
Maizels and Wolfe [13]	135	?	Episodic Migraine, Chronic Tension-Type, Episodic Tension-Type, Episodic Cluster	Initial branch points determined by headache frequency and duration	75–100
Sarchicelli et al. [14]	200	?	Chronic migraine, Probable chronic migraine, Chronic tension type, Probable chronic tension type, Medication overuse	Computerized Structured Record System	?
Dong et al. [17]	282	?	Migraine, Tension-Type, Cluster, Chronic daily headache	Rule Based Decision Clinical Decision Support System	89–97
Ahn et al. [19]	?	20 15 13	Migraine, Tension-Type, Cluster	Fuzzy	?
Yin et al. [16,20]	676	24	Probable Migraine, Probable tension-type	Case based reasoning, Genetic algorithms and K-Nearest neighbors	93.14, 89.36
Aljaaf et al. [21]	900	8	Migraine, Tension-Type, Cluster, Other	Naive Bayes, Artificial Neural Network, Decision Tree, Zero R Classifier, Support Vector Machines, k-Nearest Neighbors, Logistic Regression	92.67, 96.11, 97, 66.67, 96, 96.22, 95.33
Wu and Duan [22]	375	13	Migraine, Tension-Type, Trigeminal Autonomic Cephalalgias	Ant Colony Optimization Clustering Algorithm	89.2, 84.3, 85.7
Walters [23]	1829	4	Migraine	Four Item Migraine Screening Algorithm	93
Celik et al. [24,25]	850	40	Migraine, Tension Type, Cluster	Immunos-1, Immunos-2, Immunos-99, AIRS1, AIRS2, AIRS2-Parallel, ClonalG, CSCA	94.47, 71.65, 95.65, 99.29, 98.82, 99.65, 98.71, 99.18

classification performance [20,21,24,25], but all these applications assisted researchers in headache diagnosis and they showed how to apply new algorithms to diagnose headaches.

Specifically, web-based medical expert systems are more popular due to the fact that they can be accessed by many people. Therefore, we created a web-based headache diagnosis system and collected patients' data from different hospitals. In addition, we analyzed these data for headache categorization by using ACO which is a popular heuristic algorithm. Although heuristic algorithms are specifically implemented for optimization problems, they have significant performance in classification.

7. Conclusion

Patients who suffer from headaches need a proper tracking system for their attacks that also supports diagnosis. This kind of system helps doctors to understand and evaluate patients' symptoms, because the patients who use the headache diagnosis system provide sufficient and precise information for the doctors. A computer-based headache tracking and diagnosis system will be very common among the patients and doctors in the future. Moreover, these applications will be much more useful if they are accessible via mobile devices.

To sum up, in this study, which can be claimed to be unique in employing a heuristic algorithm for headache diagnosis for the first time, we classified primary headaches according to IHCD-2 criteria by using the ACO-based classification algorithm among 850 participants and obtained 96.9412% accuracy. The results indicated that the algorithm in this paper represents a new and successful solution to the problem of headache classification especially in terms of predictive accuracy and simplicity in generating rules. Although our system achieved a significant result, its performance will be improved as far as consultation records of new patients are examined. It will be possible to classify subgroups of migraine, tension type, or cluster headaches with the ant colony algorithm.

Acknowledgments

The authors wish to thank medical doctors Emine Rabia Koç and Nermin Tepe from Balıkesir University Neurology Department, Halil Güllüoğlu from İzmir Medical Park Hospital Neurology Department, and Mustafa Ertaş from İstanbul University Neurology Department, who assisted in the field of neurology for the medical information of primary headaches and supplying patients' survey data.

References

[1] ICHD-II classification: Parts 1–3: primary, secondary and other. Cephalalgia 2004; 24(1 suppl): 23-136.

[2] Karlı N, Zarifoğlu M, Ertaş M, Saip S, Öztürk V, Bıçakçı Ş, Boz C, Selçuki D, Oğuzhanoğlu A, Neyal M et al. Economic impact of primary headaches in Turkey: a university hospital based study: part II. J Headache Pain 2006; 7: 75-82.

[3] Kopec D, Shagas G, Selman J, Reinharth D, Tamang S. Development of an expert system for aiding migraine diagnosis. J Inform Tech Health 2004; 2: 355-364.

[4] Krawczyk B, Simić D, Simić S, Woźniak M. Automatic diagnosis of primary headaches by machine learning methods. CentEur J Med 2013; 8: 157-165.

[5] Simić S, Simić D, Slankamenac P, Simić-Ivkov M. Computer-assisted diagnosis of primary headaches. In: Corchado E, Abraham A, Pedrycz W, editors. Hybrid Artificial Intelligence Systems. Lecture Notes in Computer Science. 5271. Berlin, Germany: Springer, 2008. pp. 314-321.

[6] Andrew ME, Penzien DB, Rains JC, Knowlton GE, McAnulty RD. Development of a computer application for headache diagnosis: the headache diagnostic system. Int J Biomed Comput 1992; 31: 17-24.

[7] Francis JH, Pennal BE, Wadsworth W. Development of a computer-assisted headache diagnostic and treatment system. Headache 1984; 24: 35-38.

[8] Hasan MR, Hasan MS, Siraj F, editors. An expert system based headache solution. Computer Applications and Industrial Electronics (ISCAIE), 2012 IEEE Symposium; 3–4 December 2012; Kota Kinabalu, Sabah. Malaysia, pp. 271-276.

[9] Dorigo M. Optimization, learning and natural algorithms. PhD Thesis, Politecnico di Milano, Italy. 1992.

[10] Gallai V, Sarchielli P, Alberti A, Pedini M, Gallai B, Rossi C, Cittadini E, The Collaborative Group for the Application of IHSCotISftSoH. Application of the 1988 International Headache Society diagnostic criteria in nine Italian Headache Centers using a computerized structured record. Headache 2002; 42: 1016-1024.

[11] Simone R, Marano E, Bonavita V. Towards the computerisation of ANIRCEF Headache Centers. presentation of AIDA CEFALEE, a computer assisted diagnosis database for the management of headache patients. Neurol Sci 2004; 25: 218-222.

[12] Pryse-Phillips W, Aubé M, Gawel M, Nelson R, Purdy A, Wilson K. A headache diagnosis project. Headache 2002; 42: 728-737.

[13] Maizels M, Wolfe WJ. An expert system for headache diagnosis: the Computerized Headache Assessment Tool (CHAT). Headache 2008; 48: 72-78.

[14] Sarchielli P, Pedini M, Coppola F, Rossi C, Baldi A, Corbelli I, Mancini ML, Calabresi P. Application of the ICHD-II criteria to the diagnosis of primary chronic headaches via a computerized structured record. Headache 2007; 47: 38-44.

[15] Simic S, Simic A, Slankamenac P. Integrating case-based and rule-based decision support in headache disorder. In: ICCMSE 2006, International Conference of Computational Methods in Sciences and Engineering, 27 October–01 November 2006; Chania, Crete, Greece. pp. 507-510.

[16] Yin Z, Dong Z, Yu S, Lu X, Feng G, Duan H. A guideline-based decision support system for headache diagnosis. Stud Health Technol Inform 2013; 192: 1022.

[17] Dong Z, Yin ZM, He MW, Chen XY, Lv XD, Yu SY. Validation of a guideline-based decision support system for the diagnosis of primary headache disorders based on ICHD-3 beta. J Headache Pain 2014; 15.

[18] Ahn JY, Han KS, Oh SY, Lee CD. An application of interval-valued intuitionistic fuzzy sets for medical diagnosis of headache. Int J Innovative Comput Appl 2011; 7: 2755-2762.

[19] Ahn JY, Choi KH, Park JH. A headache diagnosis method using an aggregate operator. Commun Stat Appl Methods 2012; 19: 359-365.

[20] Yin Z, Dong Z, Lu X, Yu S, Chen X, Duan H. A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case-based reasoning. J Headache Pain 2015; 16: 29.

[21] Aljaaf AJ, Al-Jumeily D, Hussain AJ, Fergus P, Al-Jumaily M, Radi N. A systematic comparison and evaluation of supervised machine learning classifiers using headache dataset. In: Huang DS, Han K, editors. Advanced Intelligent Computing Theories and Applications, Icic 2015, Pt Iii. Lecture Notes in Artificial Intelligence. 92272015. pp. 101-108.

[22] Wu Y, Duan H. An application of ant colony optimization clustering approach for primary headache diagnosis. In: Park JJ, Pan Y, Chao HC, Yi G, editors. Ubiquitous Computing Application and Wireless Sensor: UCAWSN-14. Dordrecht, Netherlands: Springer; 2015. pp. 643-648.

[23] Walters AB, Smitherman TA. Development and validation of a four-item migraine screening algorithm among a nonclinical sample: The Migraine-4. Headache 2016; 56: 86-94.

[24] Celik U, Yurtay N, Koc ER, Tepe N, Gulluoglu H, Ertas M. Diagnostic accuracy comparison of artificial immune algorithms for primary headaches. Comput Math Methods Med 2015; 2015: 8.

[25] Celik U, Yurtay N, Koc ER, Tepe N, Gulluoglu H, Ertas M. Migraine, tension-type and cluster-type of headaches classification by using immunos algorithms. J Med Imaging Health Inform 2016; 6: 1173-1177.

[26] Keleş A. Expert doctor verdis: integrated medical expert system. Turk J Elec Eng & Comp Sci 2014; 22: 1032-1043.

[27] Keleş A, Keleş A. Extracting fuzzy rules for the diagnosis of breast cancer. Turk J Elec Eng & Comp Sci 2013; 21: 1495-1503.

[28] Taşdelen B, Helvaci S, Kaleağasi H, Özge A. Artificial neural network analysis for prediction of headache prognosis in elderly patients. Turk J Med Sci 2009; 39: 5-12.

[29] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI'95 Proceedings of the 14th International Joint Conference on Artificial Intelligence; 20–25 August 1995; Montreal, Quebec, Canada.

[30] Michelakos I, Mallios N, Papageorgiou E, Vassilakopoulos M. Ant colony optimization and data mining. In: Bessis N, Xhafa F, editors. Next Generation Data Technologies for Collective Computational Intelligence. Studies in Computational Intelligence. 352. Berlin, Germany: Springer; 2011. pp. 31-60.

[31] Parpinelli RS, Lopes HS, Freitas AA. Data mining with an ant colony optimization algorithm. IEEE T Evol Comput 2002; 6: 321-332.

[32] Martens D, Backer M, Haesen R, Baesens B, Holvoet T. Ants constructing rule-based classifiers. In: Abraham A, Grosan C, Ramos V, editors. Swarm Intelligence in Data Mining. Studies in Computational Intelligence. 34: Berlin, Germany: Springer; 2006. pp. 21-43.

[33] Gribskov M, Robinson NL. Use of receiver operating characteristic (ROC) analysis to evaluate sequence matching. Comput Chem 1996; 20: 25-33.