Cent. Eur. J. Med. • 8(2) • 2013 • 157-165
DOI: 10.2478/s11536-012-0098-5

VERSITA

Central European Journal of **Medicine**

Automatic diagnosis of primary headaches by machine learning methods

Research Article

Bartosz Krawczyk*[1], Dragan Simić[2], Svetlana Simić[3], Michał Woźniak[1]

1 Wrocław University of Technology, Department of Systems and Computer Networks, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

2 University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

3 University of Novi Sad, Faculty of Medicine, Hajduk Veljkova 1-9, 21000 Novi Sad, Serbia

Received 1 June 2012; Accepted 26 August 2012

Abstract: Primary headaches are common disease of the modern society and it has high negative impact on the productivity and the life quality of the affected person. Unfortunately, the precise diagnosis of the headache type is hard and usually imprecise, thus methods of headache diagnosis are still the focus of intense research. The paper introduces the problem of the primary headache diagnosis and presents its current taxonomy. The considered problem is simplified into the three class classification task which is solved using advanced machine learning techniques. Experiments, carried out on the large dataset collected by authors, confirmed that computer decision support systems can achieve high recognition accuracy and therefore be a useful tool in an everyday physician practice. This is the starting point for the future research on automation of the primary headache diagnosis.

Keywords: *Clinical decision support • Feature selection • Headache • Machine learning • Medical informatics*

1. Introduction

Headaches are classified commonly as primary or secondary headaches. The primary headaches are disease without underlying organic etiology. Secondary headaches are caused by an underlying organic process and are considered a symptom of the underlying disease.

Primary headaches are common disease present among the working people with a significant impact on the quality of life of the affected person as well as his/her working capability and productivity [1,2].

Working population is the carrier and the backbone of every society. This population is reproductively an active part of any society as well. Having these aspects in mind, it can be said that they are the most important group in a society. Nevertheless, there is still a discrepancy between the society's expectations and investments when it comes to the working population. The common practice is to pay attention to the health of this population when it comes to the risky occupations and perform regular medical checkups. The care about employees who suffer from a primary headache begins only when they themselves go to see a doctor [3].

Headache disorders are the most prevalent of all the neurological conditions and they are among the most frequent of medical complains seen in a general practice [4]. Half of the general population experience a headache during any given year, and more than 90% report a lifetime history of head pain [5,6].

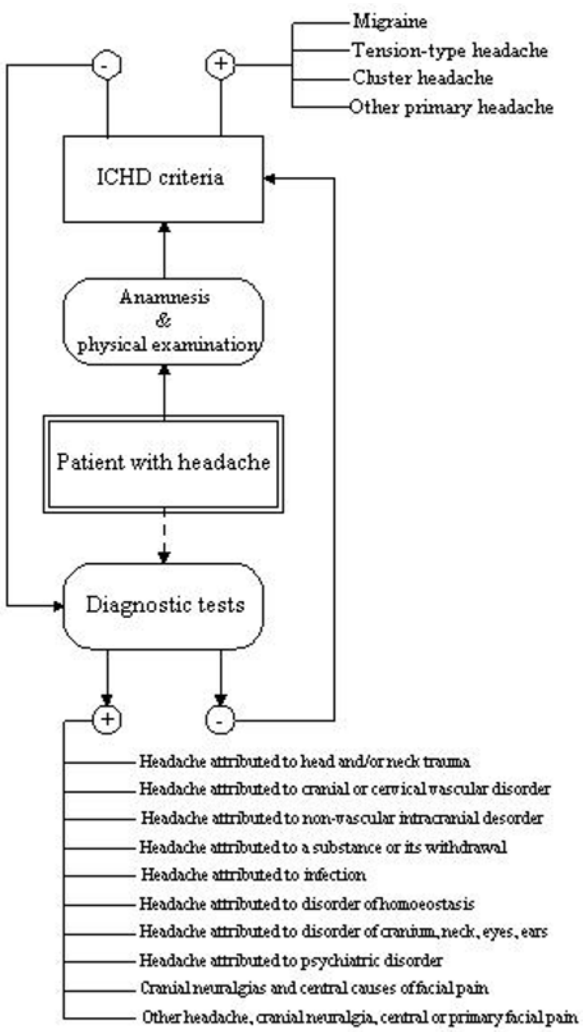
According to the International Classification of Headache Disorders – the First Edition (ICHD-1) and The International Classification of Headache Disorders – the

* *E-mail: bartosz.krawczyk@pwr.wroc.pl*

Second Edition (ICHD-2) we can established an uniform terminology and consistent operational diagnostic criteria for the wide range of the headache disorders around the world. The ICHD-2 provides a hierarchy of diagnoses with varying degrees of specificity. Headache disorders are identified with three or sometimes four digit codes. The first digit specifies the major diagnostic categories (i.e. migraine). The second digit indicates a disorder within the category (i.e. migraine without aura). Subsequent digits permit more specific diagnosis for some headache types [7-9].

The diagnostic criteria developed by the International Headache Society have been widely validated for helping physicians make diagnoses, and we have made our own questionnaire based on the International Headache Society (IHS) criteria. Figure 1 represents the methodological approach to diagnosis of certain headache types – The "Diagnosis and differential diagnosis cycle".

Figure 1. The "Diagnosis and differential diagnosis cycle" – methodological approach to diagnosing headache types.



The basic aim of the researches devoted to the headache diagnosis is to establish the most precise diagnosis of headache or/and headaches present in patients, which largely indicates the way and the efficiency of the treatment. The starting point of an accurate diagnosis is differentiating primary headaches, without organic cause, from secondary headaches, where etiological cause can be determined.

Headaches diagnosing is usually based on: anamnesis, clinical examination and additional examinations. The advantage is usually given to some of those elements, depending on the physicians' clinical orientation. Everyday clinical reality is such where there are two types of physicians in all social systems and all medical branches:

A. The ones who have the tendency to base their diagnosis on anamnesis and clinical examinations, and afterwards on additional examinations–misunderstood as the "old-school physicians".

1. And those who show the tendency to base their diagnosis primarily on additional examinations – misunderstood as the "contemporary physicians".

Physicians who are more concerned about the detailed anamnesis and clinical examinations, in the first contact with patients, considerably apply ICHD-2 criteria and can easily establish the primary headache diagnosis. If the criteria are not satisfied, the physicians will have to suggest an additional examination of a patient, according to which the diagnosis of adequate secondary headache will be established, in correlation with the findings and ICHD-2 criteria. Detailed anamnesis leads to well-directed diagnostic procedures.

When a patient suffering from headache first refers to the physicians who have the tendency to base their diagnosis on additional examinations, secondary headache diagnosis can be easily established if short anamnesis and status have already been well-focused. However, ICHD-2 criteria for the secondary headache have to be considered as well. If the additional examinations do not discover the etiological cause, physicians will have to apply the ICHD-2 criteria for establishing a primary headaches diagnosis.

Migraine is a chronic disease characterized by occasional headache attacks, mostly on one side, pulsating with moderate or severe pain which is aggravated by physical activity, can be preceded by aura period, and often followed by nausea, photo and phonophobia [7,8]. Migraine is divided into six major categories, the two most important of which are: migraine without and migraine with aura. Migraine affects about 12% of the population, and is the three times more women suffer from migraine than men do [10-12].

Let's present the ICHD-2 diagnostic criteria for migraine without aura:

B. At least five attacks fulfilling criteria B-D

C. Headache attacks lasting 4-72h and occurring on < 15 days/month (untreated or unsuccessfully treated)

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

1. Unilateral location
2. Pulsating quality
3. Moderate to sever pain intensity
4. Aggravation by or causing avoidance of routine physical activity (i.e. walking or climbing stairs)

E. During headache at least one of the following:

1. Nausea and/or vomiting
2. Photophobia and phonophobia

F. Not attributed to another disorder

Tension-type headache (TTH) is the most common primary headache. It is a non specific headache that does not have vascular causes nor is associated with the organic damage [13]. Tension-type headache typically causes pain that spreads like a band, on both sides of the head, starting at forehead and progressing towards the occiput. It often radiates towards the neck muscles, and it can even radiate towards the trapezius muscles, muscles of the shoulder girdle of scapular and interscapular region [14]. TTH pain is of mild or changing intensity, and it is described as tension, pressure or dull pain. Anamneses gives distinctive pain description – the feeling that the head is "pressed as in a vice", "inability to think clearly", "and the numbness and tingling in the head", feeling as if there was "a casque on the head". When the headache is holocephalic, the patients describe the accentuated sensitivity of the vertex of the head while combing [15]. Migraine like pain in one side of the head, pulsating pain, nausea, vomiting and photophobia is not usually present.

Let's present the ICHD-2 diagnostic criteria for tension-type headache:

A. At least ten attacks fulfilling criteria B-E

Headache < 1 day/month (episodic infrequent); 1-14 days/month (episodic frequent) or ≥ 15 days/month (chronic)

B. Headache lasting from 30 min to 7 days

C. At least two of the following characteristics:

1. Bilateral location
2. Pressing or tightening (non-pulsating) quality
3. Mild or moderate intensity (may inhibit but does not prohibit activities)
4. No aggravation by walking stairs or similar routine physical activity

D. Both of the following:

1. No nausea or vomiting (anorexia may occur)
2. Photophobia and phonophobia are absent, or one but not the other may be present

E. Not attributed to another disorder

2.X.1. Associated with pericranial tenderness

2.X.2. Not Associated with pericranial tenderness

It may be difficult to differentiate the episodic TTH from the migraine without aura or probable migraine without aura.

Tension-type headache starts a bit later than a migraine, in the second half of the third life decade and it gets less frequent as for elderly people [16].

A proper recognition of the headache type is an important but not trivial medical task. Due to the high number of every-day examinations this diagnosis is often affected by physical conditions such as weariness and routine. Therefore a high-quality computer software which can aid the physician in patient examination would be much desired. Machine learning (ML) methods are an attractive solution for such a task as they offer fast and precise intelligent analysis of multidimensional data. Such algorithms are widely used for clinical decision support [17] and are applied by authors to the tasks as the hypertension diagnosis [18], drug discovery [19], nephropathy detection among new-borns [20], or abdominal pain diagnosis [31].

This paper introduces a proposition of an automatic medical decision support system for headache's classification. Arbitrary chosen classification methods originating from the ML area are utilised to analyse the multi-dimensional data extracted from our medical investigation carried on a large group of patients. Additionally we assume that not all information from the survey yield similar discriminative power i.e., contribute in the same amount to the automatic classification of headaches. We use the statistical feature selection algorithms to select the most relevant features, what leads to the simplification of classification model, reduction of computational complexity of the classifiers and to the increase of classification accuracy. To verify the statistical selection methods we compare their results with the expert physician choice.

2. Materials and methods

2.1. Data acquisition

As has been mentioned in the previous section, it may be difficult for a physician to differentiate the episodic TTH from the migraine without aura or probable migraine without aura. That was one of the reasons, why this model was simplified in this study.

When creating any model that is this complex and which contains a large number of very close linguistic descriptions it is extremely difficult to accurately define and set clear boundaries between the studied groups–classes. This is the reason why only three clearly defined headache classes were chosen for our first study: migraine, tension-type headache and other headaches. Even this simple medical classification does not have to be very clear and precise, nor for the physicians, and neither for the automatic classification systems.

The research has been conducted on a sample of 1022 employees of both sexes and between 20-65 years of age in the Novi Sad area (Republic of Serbia) that have adequately filled in the questionnaire, and returned them the following day(s). Our research results indicate that headache is present in 579 (56.65%) subjects. Migraine prevalence in our working population is 16.53% (169 subjects), prevalence for tension type headache is 21.91% (224 subjects), and for other headache types it is 18.21% (186 subjects).

If pay attention only on subjects which suffer headaches – 579 persons, 169 subjects present 29.18 % for migraine, 224 subjects 38.60% for TTH, and 186 subjects present 32.12% other headache types.

Here other headache does not mean other primary headache but all the other headaches.

The following four questionnaires have been used as research instruments: :

- I. General Questionnaire – which contains general questions, and questions related to gender, age, company status, marital status, family status, level of education, overtime work, smoking, headache in relatives, presence of chronic disease in the examinees. The last two questions of the general questioner are related to the presence of headaches in the last year and the last month. The examinees that have answered affirmatively to the question whether they have had a headache within the last year and/or last month have become part of the next research.
- II. Questionnaire on Headache Characteristics – which contains questions related to: the year of life when the first headache occurred, the frequency of headache attacks, localization, the intensity and quality of pain, associated symptoms, the presence of prodrome and aura, headache triggers. The last two questions in this questioner were for female examinees only, and they were related to the connection between the menstrual period and/or menstrual cycle and headaches. Following the ICHD-2 criteria, the questions in the questioner have been selected in such a fashion that examinees' answers help establish the diagnoses of the headache type.

Additionally questionnaire on effects of a headache on life and work and short questionnaire on examination and treatment were also used.

2.2. Machine learning and pattern recognition

The aim of the pattern recognition is to classify a given object to the one of the predefined categories, on the basis of observation of the features describing it [21]. The object and its attributes are presented as a feature vector $x \in X$.

The pattern recognition algorithm maps the feature space X to the set of class labels M.

$$\Psi : X \rightarrow M \tag{1}$$

The mapping (1) is established on the basis of examples included in a learning set or rules given by experts. The learning set consists of learning examples i.e., observation of features described object and its correct classification.

2.3. Classification algorithms

In this study six arbitrary chosen machine learning algorithms were tested. First three of them are single-model approaches and the remaining are multiple classifier systems i.e., they use a collective decision of group of classifiers. We chose such algorithms to assess if the complexity of used methods can have a positive impact on the their quality,

Naïïve Bayes is a probabilistic classifier, rooted in Bayes' theorem with strong assumptions regarding the independence of the features [21]. It requires the prior class probability and the feature probability distribution. Those parameters are estimated with the usage of maximum likelihood estimator. Those parameters are used to assume the posterior probabilities. The Naïve Bayes classifies a new object into this class, for which the posterior probability is the highest

C4.5 is based on the 'Top Down Induction of Decision Tree' (TDIDT) [22]. The central idea of the TDIDT algorithm is the selection of 'the best' attribute, i.e. which attribute to test at each node in the tree. The family of algorithms based on the ID3 method (e.g. C4.5) uses the information gain that measures how well the given attribute separates the training examples according to the target classification. The future implementations of decision tree induction algorithms use measures based on the previously defined information gain (e.g. information ratio).

Support Vector Machine (SVM) is a deterministic binary linear classifier [23]. It constructs a set of hyperplanes in a high dimensional feature space, which are used for separating the data. A hyperplane achieves a good separation that if it has the largest distance to the nearest training data points of any class (so-called functional margin). In general the larger the margin means the lower generalization error of the classifier.

In case of non-linearly separable data the original finite dimensional space is mapped into a higher dimensional space, which makes the linear separation possible. SVM uses a fast mapping into a larger space so that cross products may be computed efficiently for the variables in the original space. This cross products in the higher dimensional space is defined as a kernel function, carefully selected to suit the problem. The hyperplanes in the large space are defined as the set of points whose cross product with a vector in that space is constant.

Bagging (or bootstrap aggregating) is an ensemble meta-algorithm developed by Breiman [24]. It is based on creating a set of new bootstrap object samples from the original dataset and training one classifier on each of them. This assures that each of the classifiers was created on a diverse, heterogeneous dataset.

Boosting is a family of algorithms dedicated to, similarly as bagging, improving the quality of weak predictors [25]. Boosting algorithms consist of iteratively learning weak classifiers and adding them to a committee. Then they are weighted according to their individual accuracies. After each iteration the data is reweighted: the weight of misclassified examples is increased, while the weight of correctly labelled samples decreases. In this paper we use an AdaBoost.M1 algorithm [25].

Random Forest approach was introduced by Breiman [26]. It is similar to the previously presented Bagging algorithm as it also uses new subsets to create heterogeneous classifiers. Yet in this case the subspaces do not only consist of bootstrap samples of objects, but also of randomly chosen features. Whereas in Bagging each classifier is trained on the same features but on different objects, in Random Forest each classifier (decision tree) is trained on different objects and different features.

2.4. Feature selection algorithms

As we mentioned above we used some preprocessing methods to assure the lower complexity and higher classification quality. Let's present the used feature selection approach shortly.

Consistency measure filter is based on the idea that the selected subset of features should be consistent and self-contained, i.e. there should be no conflicts between the objects described by similar features [27]. An dataset, described by a subset of features, is considered inconsistent if there exists at least two objects belonging to it such that they are similar except their class labels. match all but their class labels.

ReliefF is based on attribute estimation [28]. It evaluates the discriminate power of an feature by repeatedly sampling an example and considering the value of the given attribute for the nearest instance of the same and different class. This is also a filter approach

Genetic algorithm wrapper is a nature-inspired algorithm, which uses an evolutionary optimization for finding the best subset of features [29]. Feature space is encoded binary, where 1 stands for a selected feature and 0 stands for a discarded feature. As a criterion the overall classification error is selected. This approach starts with generating some random population representing the solution space (feature subsets) and through operations of cross-over (responsible for exchanging information between population members) and mutation (responsible for introducing random diversity into population) searches for an optimal solution. This is a wrapper approach i.e., it is dependent on the chosen classification model.

3. Statistical methods and experimental procedures

For automatic diagnosis of primary headaches we have applied six different classification algorithms, described in the previous section. Additionally we have incorporated a feature selection step to reduce the dimensionality of the data. From the vast number of information available an experience physician also selects the most valuable ones. Therefore after the interview with medical specialists we have established a subset of features to which the physicians pay the highest attention.

- to compare the performance of several state-of-the-art machine learning techniques to establish their usefulness in clinical decision support system aimed at headache diagnosis;
- to compare the automatic feature selection methods with the heuristic based on the physicians experience.

All simulations where done in R environment [30], with tested algorithms taken from dedicated packages, thus ensuring that results achieved the best possible efficiency and that their performance is not decreased by the wrong implementation.

SVM was trained with the Sequential Minimal Optimization algorithm and utilized the polynomial kernel. C4.5 used the information gain split criterion and post-pruning method to reduce the complexity of the trained model. Both bagging and AdaBoost.M1 used C4.5 as a base classifier and were run for ten iterations. Random Forest consisted of 120 decision trees with maximum depth constraint set to five.

Two types of ReliefF algorithms were tested – with greedy stepwise stop criterion and with heuristic selection of top ten features. For genetic algorithm we have used an population of 50 offsprings, two-point cross-over with the probability set to 0.7 and a single-point mutation with probability set to 0.3. For selection of population members a tournament scheme was applied.

All tests were concluded using a ten-fold cross validation. Feature selection step was repeated for each of the folds.

Additionally a 10CV t-test for confidence 5% [21] was used to establish if the differences between classifiers were statistically significant. As a test score the probability of rejecting the null hypothesis is adopted i.e., that classifiers have the same error rates. As an alternative hypothesis it is conjectured that tested classifiers have different error rates. A small difference in the error rate implies that the different algorithms construct two similar classifiers with similar error rates; thus the hypothesis should not be rejected. For a large difference, the classifiers have different error rates, and the hypothesis should be rejected. Therefore, two classifiers differ in a statistically significant way if the null hypothesis considering them is rejected.

4. Results

The results of the experiments are presented in the Table 1. Rows stands for different classification models and columns for used feature selection approaches (all features, four tested models and the features selected by the physician). Bolded results indicates which classification model for a given feature selection method was statistically better (in columns), while results in italics indicates which feature selection method for a given classifier model was statistically better (in rows).

Achieved results show that machine learning methods can achieve a high quality of classification of primary headaches. The best results were returned by compound machine learning methods such as Random Forest, Bagging and Boosting – as for the most of the cases they have outperformed the single classification models. These results are satisfactory from the medical point of view as well. Therefore it is possible to introduce a fully automatic decision support system for this task that achieve the diagnosis accuracy level comparable, or even slightly better, to an experienced physician.

Additionally one may see that used automatic feature selection methods returned comparable, or even slightly better, results than models built on the basis of features suggested by an expert physician. In Table 2 we present the comparison of features selected by automatic methods and by the expert physician. Pluses denotes selected feature by the considered method. In case of the physician selection double pluses indicates a feature denoted as a very important (major) by the expert and a single plus indicates features that are used as an additional support in harder cases (minor). Additionally one feature according to the physician is important for recognition of TTH class.

As we can see the automatic feature selection methods have always chosen all the major features and most of the minor ones. This indicates that those questions from the survey have at the same time a medical importance and the statistical significance for the discriminative process. Additionally two features, not used by standard medical tests, were pointed out by feature selection methods as important ones. Therefore that additional information should be studied closer form

Table 1. Classification accuracy [%] with the respect to the chosen predictor model and feature selection approach.

	All	Consistency	ReliefF top10	ReliefF Greedy	Genetic	Physician choice
Naïve bayes	72.02±4.21	77.21±3.23	76.17±4.54	77.89±3.32	77.03±5.78	79.97±5.12
SVM	76.34±1.76	77.37±1.32	78.41±1.45	78.23±1.43	77.37±1.99	78.58±1.86
C4.5	76.51±3.04	75.48±2.89	**80.14±4.05**	77.89±3.21	74.61±4.44	77.37±4.11
Random Forrest	**79.97±3.13**	**79.45±3.13**	**81.00±2.67**	**81.02±2.45**	**80.14±3.61**	**80.14±3.21**
AdaBoost.M1	76.68±2.43	76.01±2.76	77.89±1.84	76.51±2.80	76.33±3.05	75.13±2.05
Bagging	**78.24±2.98**	**79.41±2.55**	**80.14±2.01**	79.27±2.65	**79.10±3.70**	**79.27±2.98**

Table 2. Comparison between selected features by an expert physician and by automatic machine learning methods.

		Consistency	ReliefF Greedy	ReliefF top10	Genetic	Physician
	Sex					
1.	1. How old were you when the headache occurred for the first time ?					
2.	2. How often do you have headache attacks ?	+				
3.	3. My headaches last	+	+	+	+	++
4.	4. My headaches are located:	+	+		+	++
5.	5. The intensity of the pain is:	+	+	+	+	++
6.	6. The quality of the pain you experience is:	+	+	+	+	++
7.	7. Do your headaches worsen after physical activities such as walking or staircase climbing?	+	+	+	+	++
8.	8. Do you avoid routine physical activities because you are afraid they might trigger your headache?		+	+		++
9.	9.a) Are the headaches accompanied by? a) Nausea	+	+	+	+	+
10.	9.b) Are the headaches accompanied by? b) Vomiting		+	+		+
11.	9.c) Are the headaches accompanied by? c) Photophobia (light sensitivity)	+	+	+	+	+
12.	9.d)Are the headaches accompanied by? d) Phonophobia (noise sensitivity)		+	+		+
13. 10	10. Do you have temporary visual, sensory or speech disturbance?					
14.	11. Do you, during a headache attack, have tension and/ or heightened tenderness of head or neck muscles?	+	+	+	+	+ TTH
15.	12. Do you have any body numbness or weakness?					
16.	13. Do you have any indications of oncoming headache?	+	+		+	
17.	14. Headache is usually triggered by: Menstrual periods	+				
18.	15. In the half or my visual field, lasting 5 minutes to an hour, along with the headache attack or an hour before					
19.	16. Along with the headache attack or an hour before one I have following sensory symptoms					

the medical point of view and included in the clinical decision support system.

5. Conclusions

Primary headaches are the most prevalent among professionally and reproductively active population. Over the past two decades significant advances have been made in the data collection and understanding of the pathophysiology, pharmacology, genetics and epidemiology of headache, although there are still numerous issues that need further clarification.

The paper introduced a decision support system for automatic classification of primary headaches. It combined classification models with feature selection algorithms. The classifiers allowed for a fast and accurate decision making, while feature selection methods both improved the recognition step and delivered valuable information about the significance of certain parts of our diagnostic survey.

The established specificities and differences between the primary headache types, which are discussed in this paper – migraine, tension-type headache, and other headaches, among working population suggest the need for further more extensive research in this filed.

New developments in each of these fields according with different classification methods as well as discussed, will contribute to a more comprehensive understanding of this significant medical problem.

This research is a small step which towards physician's attention to the automatic decision support system for primary headaches. The high level of congruence results obtained by clinical physicians and using machine learning methods offers the possibility to further improve the diagnosis accuracy and makes daily clinical practice easier. Physicians may benefit from application of machine learning methods in diagnostics primary headaches, making their judgment faster and more precise. Given the high prevalence of primary headaches on the one hand, and frequent neglect of the other automatic diagnosis is a highly required tool in this field.

We should point out that results presented in this paper are preliminary investigations of a wide problem of the automatic headache classification. Our future works will concentrate on a full, not simplified, diagnostic case. We would like to introduce complex classification methods, such as class decomposition, one-class classification and hierarchical classifiers, to improve the quality of our clinical decision support system.

Acknkowledgement

This work is supported in part by The Polish National Science Center under the grant which is being realized in years 2011-2014.

References

[1] Pop P.H.M., Gierveld C.M., Karis H.A.M., Tiedink H.G.M., Epidemiological aspects of headache in a workplace setting and the impact on the economic loss. EFNS European Journal of Neurology 2002; 9:171-174

[2] Lipton R.B., Stewart W.F., Diamond S., Diamond M.L., Reed M.L., Prevalence and burden of migraine in the United States: Data from the American Migraine Study II. Headache 2001; 41:646-657

[3] Simić S., Simić D., Cvijanović M., Clinical and socio-demographic characteristics of tension type headache in working population, HealthMED 2012; 6(4): 1341-1347

[4] Devenport R., Acute headache in the emergency department. J Neurol Neurosurg Pshychiatry 2002; 72:33-37

[5] Kececi H., Dener S., Epidemiological and Clinical Characteristics of Migraine in Sivas, Turkey. Headache 2002; 42:275-280

[6] Hagen K., Zwart J-A., Vatten L., Stovner L.J., Bovin G., Prevalence of migraine and non-migrainous headache – head-HUNT, a large population-based study. Cephalalgia 2000; 20: 900-906

[7] Olesan J., Lipton R.B., Headache classification update 2004. Current Opinion in Neurology 2004; 17:257-282

[8] Headache Classification Subcommittee of the International Headache Society, The International Classification of Headache Disorders. Second Edition. Cephalalgia 2004; 24 (Suppl. 1):1-160

[9] Lipton R.B., Bigal M.E., Steiner T.J., Silberstein S.D., Olesen J., Classification of primary headaches. Neurology 2004; 63:427-435

[10] MacGregor E.A., Menstruation, sex hormones, and migraine. Neurol Clin 1997; 15:125-141

[11] Simić S., Simić D., Relationship between sociodemographic characteristics and migraine in working women. HealthMED 2010; 4(1):21-28

[12] Kaniecki R.G., Migraine and tension-type headache. Neurology 2002; 58(Suppl 6):S15-S20

[13] Mathew N.T., Tension type headache. Curr Neurol Neurosci Rep 2006; 6(2):100-105

[14] Cano Garcia F.J., Rodriguez Franco L., The validity of the International Headache Society criteria and the modifications put forward in 2002 in the diagnosis of migraine and tension type headaches. Rev Neurol. 2003; 16-30; 36(8): 710-714

[15] Jensen R., Olesen J., Diener H.C., Tension-type headache. In: Brandt T., Caplan L.R., Dichgans J., Diener H.C., Kennard C. (eds). Neurological Disorders – Course and Treatment. Second edition, Amsterdam, Academic Press, 2003, 23-30

[16] Schwartz B.S., Stewart W.F., Simon D., Lipton R.B., Epidemiology of tension-type headache. JAMA 1998; 279:381-383

[17] Burduk R., Wozniak M., Different Decision Tree Induction Strategies for a Medical Decision Problem, Central European Journal of Medicine , 2012, Volume 7, Number 2, 183-193

[18] Wozniak M., Two-Stage Classifier for Diagnosis of Hypertension Type, Lecture Notes in Bioinformatics, vol. 4345, Springer-Verlag, Berlin Heidelberg New York, 2006, s. 433-440

[19] Krawczyk B., Pattern recognition approach to classifying cyp 2c19 isoform, Central European Journal of Medicine, 2012, vol. 7, no. 1, pp. 38–44

[20] Krawczyk B., Classifier committee based on feature selection method for obstructive nephropathy diagnosis, in Semantic Methods for Knowledge Management and Communication, ser. Studies in Computational Intelligence, vol. 381, 2011, pp. 115–125

[21] Duda R.O., Hart P.E. Stork D.G., Pattern Classification, Wiley-Interscience 2001

[22] Quinlan JR. Induction of decision trees. Mach Learning, 1986, 1(1):81-106

[23] Vapnik V., Statistical Learning Theory, Willey 1998
[24] Breiman L. Bagging predictors. Mach Learning 1996;24(2):123-40
[25] Freund, Y., Schapire, R. Experiments with a new boosting algorithm. In Proceedings of the Thirteenth International Conference on Machine Learning (ICML) 1996, 148-156
[26] Breiman L. Random forests. Mach Learning 2001;45(1):5-32
[27] Arauzo-Azofra, A., Benitez, J. M. and Castro, J. L. Consistency measures for feature selection. J. Intell. Inf. Syst. 30, 3, 2008, 273-292
[28] Kenji Kira, Larry A. Rendell: A Practical Approach to Feature Selection. In: Ninth International Workshop on Machine Learning, 1992, 249-256
[29] Guyon I., Gunn S., Nikravesh M. and Zadeh L., Feature extraction, foundations and applications, Springer, 2006
[30] webpage of R project: http://www.r-project.org/
[31] Burduk R., Classification error in Bayes multistage recognition task with fuzzy observations. Pattern Anal. Appl. 13(1): 85-91,2010