Pareto filter in the process of multi-label classifier synthesis in medical diagnostics support algorithms

A. AMELJAŃCZYK

andrzej.ameljanczyk@wat.edu.pl

Military University of Technology, Faculty of Cybernetics
Institute of Computer and Information Systems
Kaliskiego Str. 2, 00-908 Warsaw, Poland

The paper presents the possibility of using multi-criteria optimization methods for simple classifiers fusion in a more precise and reliable classifiers complex. There are defined simple classifiers (one label) in the form of classifier committees and sample the synthesis relations of classifiers allow to obtain classifiers filed with improved properties.

Keywords: Pareto filter, medical diagnostics, the committee classifiers, classifiers synthesis, not dominated diagnosis.

1. Introduction

Contempratory algorithms for computer-supported medical diagnostics require increasingly accurate classifiers enabling the performance of the conclusion process on the basis of complicated multimedia, frequently incomplete and uncertain medical data on the patient's health condition [11, 18]. Available literature in this field presents many papers on the structure (fusion, synthesis) of the integrated classifier in the context of striving to obtain more certain and precise medical diagnostic classifications [6, 7, 8, 9, 15, 19, 21, 22, 28]. Diagnostic information on the patient's health is contained in data describing the disease symptoms, risk factors and results of specialist laboratory tests (usually in multimedia form). The construction of classifiers making comprehensive use of such complex, complicated and diversified data is a difficult task. Development of simple classifiers e.g. domain and in particular binary is much easier [8, 9, 15, 16]. This leads to a problem of fusion of the acquired diagnostics information, reduced in most cases to so-called simple classifier synthesis. The result of synthesis (fusion) of simple classifiers is so-called complex classifiers. These are a "certain function" of simple classifiers. The purpose of synthesis is to obtain a classifier of better classifying (diagnostics) properties and in particular of lower classification error. The specific nature of medical diagnostics, due to uncertainty and incompleteness of medical data and due to the fact that the patient may suffer not from a single disease but from two or more, gives preferences to the multi-label (multi-class) classifiers [6, 7, 8, 13, 15]. The specific nature of the medical diagnostics processes, due to common uncertainty and incompleteness of data and possibility of presence of concomitant diseases, practically excludes the support algorithms using the single-label classifiers.

2. Classifiers applying the ranking functions

Let X determine the finite set of medical diagnostic data sets (observations, instances, test results), called an observation space.

Let $\mathcal{L} = \left\{ l_1, ..., l_m, ..., l_M \right\}$ – a set (repository) of labels (objects) of disease units, numbered with the $m \in \mathcal{M} = \left\{ 1, ..., M \right\}$ index.

A single-label classifier will be the function

$$C : X \to \mathcal{L} \qquad (1)$$

Each observation (instance) $x \in X$ is "associated" with the single label $l_m \in \mathcal{L}$

$$C(x) \in \mathcal{L}, \; x \in X \qquad (2)$$

An preimage of label l_m shall be a set of observations leading to the same label (decision class) l_m.

$$C^{-1}\left(\left\{ l_m \right\} \right) = \left\{ x \in X \mid C(x) = l_m \right\} \subset X \quad (3)$$

Function $C(x)$ does not need to be an injection.

A multi-label classifier will be the function

$$C : X \to 2^{\mathcal{L}} \qquad (4)$$

such that

$$C(x) \subset \mathcal{L}, \ x \in X \qquad (5)$$

The relation of correlations (associations) of observation x with the set of classification results shall be the relation of the following type

$$R_p \subset X \times 2^{\mathcal{L}} \qquad (6)$$

Thus the element R_p shall be the pairs:

$$(x, \mathcal{L}_x), \ \text{where} \ \mathcal{L}_x \subset \mathcal{L}, \ x \in X \qquad (7)$$

The pair of type (7) shall be called an indication generated by the result (observation) x.

For each observation $x \in X$ we may define a relation of ranking preferences R_x in a way that $(l_i, l_j) \in R_x$ when and only when for the observation $x \in X$, label l_i is more preferred (is "better", more "fitted") than label l_j. The symbol $r(R_x)$ will determine the ranking generated by relation R_x [2, 5].

It is willingly assumed for the relations R_x that these should be linear order relations [20]. Ranking $r(R_x)$, determined by relation R_x is in such a case permutation of set $\mathcal{L}$. The number of such rankings is $|\mathcal{L}|! = M!$ This assumption is however frequently difficult to meet. In practice, the models (relations) R_x are frequently defined (determined) with the use of ranking (scoring) functions,

$$f_x : \mathcal{L} \to \mathcal{R}^1 \qquad (8)$$

values of which are in general determined on the basis of different similarity ("fitting", "distance") models, for example: Tversky, Bayes, Jaccard, Hamming, Dice, Sokal, Russel, Lance and others observation x to disease unit labeled $l \in \mathcal{L}$ [6, 8, 24, 28].

$$R_x = \left\{ (l_i, l_j) \in \mathcal{L}^2 \mid f_x(l_i) \geq f_x(l_j) \right\} \qquad (9)$$

Functions $f_x(l)$, $l \in \mathcal{L}$ are sometimes called the utility functions or similarity or fitting ratios. In this paper, we will further assume that the functions $f_x(l)$ are normalized in the range of $[0,1] \subset \mathcal{R}^1$ [2, 15, 22, 23].

Thus, if for any $l_i, l_j \in \mathcal{L}$ it is true that $f_x(l_i) \geq f_x(l_j)$, from the perspective of result (observation) $x \in X$, l_i is placed in the ranking before the label l_j, which means that l_i is better fitted to the result x than l_j. Definition

(9) uses purposefully the sign "$\geq$", which results mostly from the fact that the ranking functions are generally not injective functions (this feature results usually from the properties of a model used for definition), [2, 6]. Such an assumption results in that the relations of preference R_x are not antisymmetric [20] which means that these determine only the so-called quasi-order [2, 4, 20]. The rankings acquired in this way are not permutations of set $\mathcal{L}$ (are not linear rankings). Adopting the weaker assumptions is implied, however, by the "practice" of defining the ranking functions, which – as already mentioned – are usually not injective functions. The symbol $r(f_x(l))$ shall determine the sequence ($\mathcal{L}$ set ranking) acquired with the use of function $f_x(l)$ [5]. Such ranking functions are frequently used for classifier development. Let $f_x(l)$, $l \in \mathcal{L}$ be a certain ranking function determined on set $\mathcal{L}$.
This function determines the classifier:

$$C : X \to 2^{\mathcal{L}}$$

according to the following formula:

$$C(x) = \arg \max_{l \in \mathcal{L}} f_x(l), \ x \in X \qquad (10)$$

Formula (10) may be presented as follows:

$$C(x) = \arg \max_{l \in \mathcal{L}} f_x(l) = \left\{ \overset{*}{l} \in \mathcal{L} \mid f_x\left(\overset{*}{l} \right) = \max_{l \in \mathcal{L}} f_x(l) \right\} \qquad (11)$$

The formula (11) presents the association between the classifier $C(x)$ and the ranking function $f_x(l)$. The classifier so structured is in general a multi-label classifier [13].

If for each $x \in X$ it is true that:

$$\left| \arg \max_{l \in \mathcal{L}} f_x(l) \right| = 1 \qquad (12)$$

the classifier developed on the basis of such ranking function is the single-label classifier.
If there is any $x \in X$ such that

$$\left| \arg \max_{l \in \mathcal{L}} f_x(l) \right| > 1 \qquad (13)$$

the classifier developed on the basis of such ranking function is the multi-label classifier. The classifier of type (10) shall be simple (ranking). Classifiers developed on the basis of ranking functions generally fail to meet the condition (12). There are also cases, in which the ranking function is not an injective one and

the condition (12) is met only for a certain subset $x \in \overline{X} \subset X$.

3. Classifier synthesis in medical diagnostics

Let further be $\mathcal{N} = \{1,...,n,...,N\}$ – set of numbers of ranking functions of type (8) [10, 24],

$$f_x^n(l), \ n \in \mathcal{N} \ - n - \text{ranking function} \qquad (14)$$

These functions generate the set (committee) of simple (ranking) classifiers

$$C(x) = \{C_1(x),...,C_n(x),...,C_N(x)\} \qquad (15)$$

Where

$$C_n(x) = \arg \max_{l \in \mathcal{L}} f_x^n(l) \qquad (16)$$

n – simple classifier.

As already mentioned, functions $f_x^n(l)$ may be defined on the basis of different similarity (fitting) model. For example, in medical diagnostics these may include the similarity indices in the area of diagnosed disease symptoms, risk factors or results of specialist diagnostic (e.g. laboratory) tests [6, 7], defined as metric [6], graphic or binary similarity indices [24].

The set of simple classifiers (15) may form a basis for obtaining, as a result of fusion (synthesis), complex classifiers, more precisely and reliably classifying the observations $x \in X$ [1, 6, 25, 28]. The selection of specific classifiers for the synthesis of a complex classifier (selection of classifier "committee") should ensure satisfying a series of conditions and expectations concerning, among others: heterogeneity, independence, no correlation and for the most low classification error. Let further be:

$\mathcal{L} = \{l_1,...,l_m,...,l_M\}$ – a finite label set (repository) and the vector ranking function of the following type:

$$f_x : \mathcal{L} \to \mathcal{R}^N \text{, such that}$$

$$f_x(l) = \left(f_x^1(l),...,f_x^n(l),...,f_x^N\right) \in \mathcal{R}^N, \qquad (17)$$

Set Y_x shall be the ranking image of set $\mathcal{L}$ for the observation $x \in X$, given by the function f_x.

$$Y_x = f_x(\mathcal{L}) = \{y = f_x(l) \in \mathcal{R}^N | l \in \mathcal{L}\} \qquad (18)$$

Element $y \in f_x(\mathcal{L})$ is an image of label l in the meaning of its assessment by all ranking functions $f_x^n(l)$, understood as multi-objective

level of similarity (fitting) of the observation $x \in X$ to the disease unit labeled $l \in \mathcal{L}$.

Thus the simplified formula shall further be:

$$y = (y_1,...,y_n,...,y_N) = \left(f_x^1(l),...,f_x^n(l),...,f_x^N(l)\right) \in \mathcal{R}^N$$

where $y_n = f_x^n(l)$ – ranking value of label $l \in \mathcal{L}$ in the meaning of n-ranking function associated with observation $x \in X$.

For each $x \in X$, $Y_x \subset Y = [0,1] \times ... \times [0,1]$.

Set Y shall mean the classifier synthesis area.
The synthesis relation (or relation of preferences of classifier committee) shall be the following relation

$$R \subset f_x(\mathcal{L}) \times f_x(\mathcal{L}) = Y_x \times Y_x$$

defined as follows:

$$R = \{(y, z) \in Y_x \times Y_x | \text{"committee prefers y then z"}\} \qquad (19)$$

The relation induced by relation R shall be the relation $\overline{R} \subset \mathcal{L} \times \mathcal{L}$, defined as follows:

$$\overline{R} = \{(l_k, l_m) \in \mathcal{L} \times \mathcal{L} | (f_x(l_k), f_x(l_m)) \in R\} \quad (20)$$

The synthesis relation R plays a key role in the process of simple classifier synthesis, since it induces a relevant preference relation $\overline{R}$ in label set L. Properties of induced relation $\overline{R}$, in particular so called ordering properties [20], depend primarily on the ordering properties of relation R and properties (e.g. injective properties) of the ranking function. The synthesis relation R expresses the principle of preferences of committee in the area of deciding whether the label l_k is "better fitted" to observation $x \in X$ compared to label l_m. There are many known preferences applicable to such synthesis. The most typical principle is the Pareto principle (relation, filter). It states that label l_k is more preferred (better fitted to observation x) than label l_m, provided that l_k is at least at the same position (or higher) as label l_m in the ranking of each committee member [2, 3]. This means that the following must be true:

$$f_x^n(l_k) \geq f_x^n(l_m), \ n \in \mathcal{N} \qquad (21)$$

The Pareto Filter (PF) is an algorithm enabling determination from any set of elements the set of elements of the highest quality in this set (in the meaning of Pareto relation) [2, 3, 6]. The effect (result) of applying the Pareto filter on set Y is so-called 'Pareto front' (set of nondominated

(minimum)) elements in the meaning of Pareto relation Y_N^R defined as follows:

$$Y_N^R = \left\{ \begin{array}{l} y \in Y \,\big|\, does\ not\ exists \\ z \in Y - \{y\},\ such\ (z,y) \in R \end{array} \right\} \qquad (22)$$

Therefore, the result of the filtration process is decisive for the adopted preferences (filtration) relation R (in more detail – its properties). So, such a relation is frequently called a preference filter or briefly: filter. The general reflection of the Pareto filter is a cone filter (CF), in which the filtration reaction is generated by a cone [3, 4, 26, 27].

The other known preference principle is the lexicographic principle [3] (considering the importance hierarchy (quality, competences) of the "committee members"). Its basis is formed by the set of permutations of set N. Each lexicographic relation leads to the ordering of a linear set $\mathcal{L}$ [2, 4, 20]. Other synthesis relations (filters) may be the Hurwicz relations [3, 11] including in particular pessimistic relation R^P and optimistic relation R^O.

$$R^P = \left\{ (l_k, l_m) \in \mathcal{L} \times \mathcal{L} \,\Big|\, \min_{n \in \mathcal{N}} f_x^n(l_k) \geq \min_{n \in \mathcal{N}} f_x^n(l_m) \right\} \qquad (23)$$

$$R^O = \left\{ (l_k, l_m) \in \mathcal{L} \times \mathcal{L} \,\Big|\, \max_{n \in \mathcal{N}} f_x^n(l_k) \geq \max_{n \in \mathcal{N}} f_x^n(l_m) \right\} \qquad (24)$$

The pessimistic relation R^P (pessimistic filter) means that the "committee" prefers label l_k, even if in the least advantageous ranking for label l_k it achieves at least the same value as label l_m, [3]. It is analogical in the case of optimistic relation R^O [4]. An interesting property of the discussed synthesis relations is the fact that the Pareto relation is a subset of each of them which results in the fact that classifications obtained in effect of applying these relations have nonempty intersection with the Pareto classification [3, 4].

Let $y \in f_x(\mathcal{L})$. The symbol $f_x^{-1}(\{y\})$ shall determine the preimage of element y

$$f_x^{-1}(\{y\}) = \{l \in \mathcal{L} \,|\, f_x(l) = y\} \qquad (25)$$

The fact that $(y, z) \in R$ means that the "classifier committee" prefers the labels from set $f_x^{-1}(\{y\})$ compared to labels from set $f_x^{-1}(\{z\})$.

Therefore the following is true:

$$f_x^{-1}(\{y\}) \times f_x^{-1}(\{z\}) \subset \overline{R}$$

The *CCS* task – complex (integrated) classifier synthesis – may be defined as multi-objective optimization problem [2, 3, 27] in the form:

$$CCS = (\mathcal{L}, f_x, R) \qquad (26)$$

which may be abbreviated (see (18)) to the pair:

$$(Y_x, R) \qquad (27)$$

The synthesis relation R may be used to develop a complex, multi-label classifier (meta-classifier) and meta-ranking (committee ranking), being a "specific synthesis" of component rankings determined by the ranking functions $f_x^n(l)$, $n \in \mathcal{N}$.

The solution of the task (26) is thus an preimage of the filtration task solution (27), i.e. subset of labels, from which there are no "better" labels in the set $\mathcal{L}$ (better fitted) to the observation $x \in X$.

$$\mathcal{L}_x^{RN} = f_x^{-1}(Y_x^{RN}) \qquad (28)$$

where

$$Y_x^{RN} = \left\{ \begin{array}{l} y \in Y_x \,\big|\, does\ not\ exists \\ z \in Y_x - \{y\},\ such\ (z,y) \in R \end{array} \right\}, \qquad (29)$$

thus

$$\mathcal{L}_x^{RN} = f_x^{-1}(Y_x^{RN}) = \{l \in \mathcal{L} \,|\, f_x(l) \in Y_x^{RN}\}, \qquad (30)$$

Set $\mathcal{L}_x^{RN}$ is called a nondominated label set [2,3,2627]. This is a subset of these labels from the set $\mathcal{L}$, from which there are no better "fitted" labels to the observation $x \in X$. This is the effect of filtration of set Y_x, using relation R.

The integrated classifier in the meaning of relation R (meta-classifier) is the complex classifier:

$$C_R(x) = f_x^{-1}(Y_x^{RN}) \subset \mathcal{L} \qquad (31)$$

This is in general the multi-label classifier, which assigns to each observation (instance) $x \in X$ the "optimum" subset of nondominated labels $\mathcal{L}_x^{RN}$ in the meaning of relation R.

In medical diagnostics, this diagnosis is considered the "best fitted" diagnosis corresponding to observation $x \in X$. This proposal is the most important and the most frequently applied diagnostic reference in the process of computer diagnosing support [5, 6, 7].

Set of classifications $\mathcal{L}_x^{RN}$ (result of operation of classifier (31)) has many interesting properties. These include among others:

a) in set $\mathcal{L}_x^{RN}$ there is no label l_k which would be better (more fitted to observation

$x \in X$) from any other label from set $\mathcal{L}_x^{RN}$ (labels in set $\mathcal{L}_x^{RN}$ do not dominate each other);

b)   for each label $l_k \notin \mathcal{L}_x^{RN}$ there is in set $\mathcal{L}_x^{RN}$ a label l_m better than this label in the meaning of R (more fitted);

c)   committee preference relation (synthesis relation) R divides the set Y_x (and thus the label set $\mathcal{L}$) into nondominated ranking clusters [5]

$$Y_x^{RN}(k) = \left(Y_x \setminus \bigcup_{i=0}^{k-1} Y_x^{RN}(i) \right)_N^R \text{ dla } k = 1, 2, \dots \tag{32}$$

and, respectively, the sets:

$$\mathcal{L}_x^{RN}(k) = f^{-1}\left(Y_x^{RN}(k) \right), \ k = 1, 2, \dots \tag{33}$$

i.e. to label equivalency clusters, dividing the entire label set $\mathcal{L}$ and thus creating the equivalency cluster ranking [5]

$$r(\mathcal{L}, R) = \left\langle \mathcal{L}_x^{RN}(1), \dots, \mathcal{L}_x^{RN}(k), \dots, \mathcal{L}_x^{RN}(K) \right\rangle.$$

The formula (32) is thus in a certain sense a generalization of the a conventional ranking of set Y [2, 5].

4.   Conclusions

The paper presents the method of simple classifier synthesis using the methodology of multi-objective optimization in the form of Pareto filtration. The synthesis of the new classifiers applies the most frequently used in the multi-objective optimization Pareto relation. The other meta-classifiers may be acquired using the other synthesis relation (see (26)). Such relations include the Hurwicz relations (23) and (24) and many others. The complex classifiers developed in such a manner (generally multi-label ones) of type $C_R(x) = f_x^{-1}\left(Y_x^{RN} \right) \subset \mathcal{L}$ have many interesting properties. Formally, these are the preimages of sets of nondominated elements (or dominating [3, 4]) of set

$$Y_x = f_x(\mathcal{L}) = \left\{ y = f_x(l) \in \mathcal{R}^N \big| l \in \mathcal{L} \right\} \quad \text{i.e. the}$$

ranking set of label repository obtained on the basis of observation x. Properties of this set depend primarily on the property of the adopted synthesis relation R [2, 3, 5, 26].

5.   Bibliography

[1]   Acid S., Campos L.M., "A comparison of learning algorithms for Bayesian Networks: a case study based on data from an emergency medical service", *Artificial Intelligence in Medicine*, 30, 215–232 (2004).

[2]   Ameljańczyk A., „Metryki Minkowskiego w tworzeniu uniwersalnych algorytmów rankingowych", *Biuletyn WAT*, Vol. LXIII, Nr 2, 324–336 (2014).

[3]   Ameljańczyk A., *Multipe optimization*, WAT, Warszawa, 1986.

[4]   Ameljańczyk A., *Optymalizacja wielo-kryterialna w problemach sterowania i zarządzania*, Ossolineum, Wrocław, 1984.

[5]   Ameljańczyk A., „Metoda podziału zbioru obiektów na wielokryterialne klastry jakościowe", *Biuletyn Instytutu Systemów Informatycznych*, Nr 12, 1–7 (2013).

[6]   Ameljańczyk A., "Multicriteria similarity models in medical diagnostics support algorithms", *Bio-Algorithms and Medical Systems*, Vol. 21, No.1, 33–39 (2013).

[7]   Ameljańczyk A., „Wiarygodność komputerowych systemów wspomagania diagnostyki medycznej", w: *Problemy modelowania i projektowania opartych na wiedzy systemów informatycznych na potrzeby bezpieczeństwa narodowego*, 23–39, WAT, Warszawa, 2014.

[8]   Balcan M.F., Bansal N., Beygelzimer A., Coppersmith D., Langford J., and Sorkin G.B., "Robust reductions from ranking to classification", *Machine Learning*, 72(1–2), 139–153 (2008).

[9]   Bouyssou D., Marchant T., "An axiomatic approach to noncompensatory sorting methods in MCDM, I: The case of two categories", *EJOR*, 178(1), 217–245 (2007).

[10] Brans J.P.,Vincke Ph., "A preference ranking organization method: The PROMETHEE method for Multiple Criteria Decision-Making", *Management Science*, Vol. 31, No. 6, 647–656 (1985).

[11] Courtney H., Kirkland J., Viguerie P., „Strategia w warunkach niepewności", w: *Zarządzanie w warunkach niepewności*, Harvard Busines Review, 316–324, Helion, Gliwice, 2006.

[12] Dieterich G.T., "Machine learning research: four current directions", *AI Magazine*, 18, 97–136 (1997).

[13] Furnkranz J., Hullermeier E., Mencia E., and Brinker K., "Multilabel classification

via calibrated label ranking", *Machine Learning*, 73, 133–153 (2008).

[14] Kahneman D., Tversky A., "Prospect theory: An Analysis of Decision Under Risk", *Econometrica*, Vol. 47, No. 2, 67–72 (1978).

[15] Kuncheva L.I., Whitaker C.J., "Measures of diversity in classifier ensembles", *Machine Learning*, Vol. 51, No. 2, 181–207 (2003).

[16] Larose D., *Metody i modele eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa, 2008.

[17] Pawlak Z., "Rough Sets", *International Journal of Computer and Information Sciences*, Vol. 11, 341–356 (1965).

[18] Pawlak Z., *Systemy informacyjne – podstawy teoretyczne*, WNT, Warszawa, 1983.

[19] Perl J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, Los Angeles, California, 1988.

[20] Rasiowa H., *Wstęp do matematyki współczesnej*, PWN, Warszawa, 2005.

[21] Ruiz C., *Illustration of the K2 Algorithm for Learning Bayes Net Structures*, Department of Computer Science WPI, Bayesian Network Power Constructor, Worcester, MA, 2009.

[22] Saaty T.L., "Rank from comparisons and from ratings in the analytic hierarchy/network processes", *EJOR*, 168(2), 557–570 (2006).

[23] Seo F., Sakawa M., *Multiple Criteria Decision Analysis in Regional Planning*, D. Reidel-Kluwer, Dordrecht–Boston––Lancaster–Tokyo, 1988.

[24] Seung-Seok Ch., Sung-Hyuk Ch., Tappert Ch.C., *A Survey of Binary Similarity and Distance Measures*, Pace University, New York, 2006.

[25] Shipp C.A., Kuncheva L.I., "Relationships between combination methods and measures of diversity in combining classifiers", *Information Fusion*, Vol. 3, No. 2, 135–148 (2002).

[26] Yu P.L., Leitmann G., "Compromise solutions, domination structures and Salukwadze's solution", *JOTA*, Vol. 13, 14–21 (1974).

[27] Yu P.L., Leitmann G., "Nondominated decision and cone convexity in dynamic multicriteria decision problems", *JOTA*, Vol. 14, 195–203 (1974).

[28] Walczak A., Bieniek B., Różyk-Jahnz K., Paluchowska E., „Fuzja klasyfikatorów w diagnostyce chorób skóry", w: *Problemy modelowania i projektowania opartych na wiedzy systemów informatycznych na potrzeby bezpieczeństwa narodowego* (red. T. Nowicki i Z. Tarapata), 143–152, WAT, Warszawa, 2014.

Filtr Pareto w procesie syntezy klasyfikatorów wieloetykietowych w algorytmach wspomagania diagnostyki medycznej

A. AMELJAŃCZYK

W pracy przedstawiono możliwość wykorzystania metod optymalizacji wielokryterialnej w procesie fuzji klasyfikatorów prostych w bardziej precyzyjne i wiarygodne klasyfikatory złożone. Przedstawiono proste (jednoetykietowe) klasyfikatory w postaci komitetów klasyfikatorów, pozwalające uzyskiwać klasyfikatory złożone o lepszych własnościach.

Słowa kluczowe: filtr Pareto, diagnostyka medyczna**,** komitet klasyfikatorów, synteza klasyfikatorów, diagnoza niezdominowana.