Experiments with Classification of MMPI Profiles using Fuzzy Decision Trees

Krzysztof Pancerz
University of Rzeszów, Poland
Email: kpancerz@ur.edu.pl

Vitaly Levashenko, Elena Zaitseva
University of Zilina, Slovakia
Email: Vitaly.Levashenko@fri.uniza.sk
elena.zaitseva@fri.uniza.sk

Jerzy Gomuła
Cardinal Stefan Wyszyński University
in Warsaw, Poland
Email: jerzy.gomula@wp.pl

Abstract—The paper is devoted to classification of MMPI (Minnesota Multiphasic Personality Inventory) profiles using fuzzy decision trees generated by means of the algorithm that uses cumulative information estimations of the initial data proposed by V. Levashenko et al. All of the stages of the classification process (i.e., fuzzification of the input data, generation of the classifier, testing the classifier) are presented and the results are discussed. A special attention is focused on determination of the center points on the MMPI scales for the fuzzification process.

I. Introduction

OUR research, conducted for over eight years, is devoted to analysis and classification of data coming from the MMPI (Minnesota Multiphasic Personality Inventory) test (see some overview given in [1]). This test delivers psychometric data in the form of the so-called profiles (thirteen descriptive attributes corresponding to scales) used to assess patients in terms of personality traits and psychopathology. We have used several methodologies for classification of MMPI profiles which can be roughly grouped into the following categories: dissimilarity measure based classifiers, index based classifiers, classification functions, rule based classifiers, and decision tree based classifiers. A palette of classifiers has been extended by a classifier based on fuzzy decision trees. Firstly, preliminary results of research on application of fuzzy decision tress for classification of psychometric data presented in [2] are very promising in relation to results obtained for other kinds of classifiers (cf. [3]). Secondly, the character of MMPI data matches the idea of fuzzy set based approaches. For each scale included in the MMPI profile, we can define linguistic values (e.g., extremely low, very low, average, raised, high, very high, extremely high) which can be described by fuzzy sets. Fuzzy decision trees for creation of classifiers are generated by means of the algorithm based on cumulative information estimations of the initial data proposed by V. Levashenko et al. (see [4]). This algorithm is recalled in Section III-B. The main research problem, touched upon in this paper, is the determination of the intervals and/or the center points (shortly called the centers) on the MMPI scales for the fuzzification process. Therefore, we have tested different approaches which can be grouped into two categories. The first category includes approaches based on the expert knowledge used to determine intervals/centers for the fuzzification process. The second category includes approaches in which intervals/centers are induced from data.

II. Input Data

MMPI is a standardized psychometric test of adult personality and psychopathology (cf. [5]). The MMPI test delivers psychometric data in a form of the so-called profiles. Formally, a profile for the patient is a data vector consisting of values of thirteen descriptive attributes (corresponding to scales). The set of scales can be divided into two parts: the validity part (three scales: L - lying, F - infrequency, K - correction) and the clinical part (scales: $1.Hs$ - Hypochondriasis, $2.D$ - Depression, $3.Hy$ - Hysteria, $4.Pd$ - Psychopathic deviate, $5.Mf$ - Masculinity/Femininity, $6.Pa$ - Paranoia, $7.Pt$ - Psychasthenia, $8.Sc$ - Schizophrenia, $9.Ma$ - Hypomania, $0.It$ - Social introversion). In our research, we have used data coming from the WISKAD-MMPI test that is a Polish adaptation of the American test. The test originally was translated by M. Choynowski (as WIO) [6] and elaborated by Z. Płużek (as WISKAD) in 1950 [7]. The data set was collected for research by T. Kucharski and J. Gomuła in the Psychological Outpatient Clinic. It includes profiles of 1710 women. Before the profiles of women screened with the WISKAD-MMPI test formed a database for further experiments, first they had been sorted by the competent judges method - five specialists with many years of experience in the application and interpretation of the MMPI results/profiles. On the basis of these items, scores are calculated for both validity scales and clinical scales. Hence, values of descriptive attributes describing patients are expressed by the so-called T-scores [T]. The T-scores scale, which is traditionally attributed to MMPI, represents the following parameters: offset ranging from 0 to 100 T-scores, average equal to 50 T-scores, standard deviation equal to 10 T-scores. The scores are expressed as K-corrected T-Scores. The scales $1.Hs$, $4.Pd$, $7.Pt$, $8.Sc$, and $9.Ma$ are corrected by adding multiples of the scale K to them.

In our experiments, the patients' profiles are recorded in a tabular form which is formally called a decision table $DT = (U, Attr, Dec)$, where U - the set of cases (patients), $Attr = \{A_1, A_2, \ldots, A_{13}\}$ - the set of descriptive (condition) attributes corresponding to scales, $Dec = \{D\}$ - the set of decision attributes consisting of the attribute D assigning each patient from U to one of 20 classes such as the reference (*norm*) class and nosological types: neurosis (*neur*), psychopathy (*psych*), organic (*org*), schizophrenia

(*schiz*), delusion syndrome (*del.s*), reactive psychosis (*re.psy*), paranoia (*paran*), (sub)manic state (*man.st*), criminality (*crim*), alcoholism (*alcoh*), drug addiction (*drug*), simulation (*simu*), dissimulation (*dissimu*), and six deviational answering styles (*dev1*, *dev2*, *dev3*, *dev4*, *dev5*, *dev6*).

III. METHODS AND TOOLS

In this section, we present methods and tools used in experiments with classification of MMPI profiles by means of classifiers built on the basis of fuzzy decision trees.

A. Fuzzification

Fuzzification is the process that transforms the continuous value variables into linguistic variables whose domains contain linguistic values which can be described by fuzzy sets (their membership functions). Fuzzification is an important stage of the process of creation of a fuzzy decision tree based classifier. Many types of membership functions can be used to describe linguistic values, but triangular or trapezoidal shaped membership functions are the most common. In our approach, the fuzzification process consists of three stages. In the first stage, we determine intervals/centers (within the range $[0, 120]$) for each linguistic value assigned to a given descriptive attribute (scale). In the second stage, we define membership functions on the basis of centers determined in Stage 1 for each linguistic value assigned to a given descriptive attribute (scale). In the third stage, we calculate values of fuzzified descriptive attributes on the basis of membership functions defined in Stage 2. Determination of intervals/centers for the fuzzification process is one of the main research problems. In experiments, we have tested different approaches which can be grouped into two categories: approaches based on the expert knowledge used to determine intervals/centers (further, such approaches are called expert approaches) and approaches in which intervals/centers are induced from data (further, such approaches are called inductive approaches). The centers for four tested expert approaches are as follows:

- the Welsh's approach
  - all scales: 15.0, 35.0, 45.0, 55.0, 62.5, 67.5, 75.0, 85.0, 95.0, 110.0,
- the Płużek's (original) approach
  - L: 38.0, 43.0, 55.5, 75.5, 88.0,
  - F: 45.5, 60.5, 80.5, 100.5,
  - K: 36.0, 55.5, 74.5,
  - clinical scales: 34.5, 60.0, 75.5, 90.5, 105.5, 115.5,
- the Gomuła's (modified Płużek's) approach
  - validity scales: 15.0, 35.0, 42.5, 47.5, 55.0, 62.5, 67.5, 75.0, 82.5, 87.5, 100.0, 115.0,
  - clinical scales: 15.0, 37.5, 47.5, 52.5, 60.0, 67.5, 75.0, 90.0, 105.0, 115.0,
- the Gomuła's (original) approach:
  - all scales: 15.0, 35.0, 45.0, 55.0, 62.5, 67.5, 72.5, 77.5, 82.5, 87.5, 95.0, 105.0, 115.0.

The calculated centers for four tested inductive approaches are as follows:

- the K-means based approach [8]
  - L: 46.85, 55.69, 63.29, 79.58,
  - F: 56.43, 68.64, 80.82, 100.08,
  - K: 35.44, 48.60, 54.74, 66.47,
  - $1.Hs$: 53.14, 61.85, 69.29, 81.68,
  - $2.D$: 58.70, 70.22, 80.32, 92.67,
  - $3.Hy$: 55.97, 63.23, 69.61, 79.76,
  - $4.Pd$: 56.39, 65.07, 73.68, 87.68,
  - $6.Pa$: 57.45, 70.39, 83.23, 101.11,
    - $7.Pt$: 56.44, 67.28, 75.87, 95.50,
    - $8.Sc$: 58.06, 73.38, 85.22, 105.08,
    - $9.Ma$: 49.01, 58.11, 69.42, 85.47,
    - $0.It$: 51.26, 58.06, 62.39, 66.76,
- the equipotent interval approach
  - L: 46.25, 59.50, 76.25,
  - F: 55.75, 73.50, 94.75,
  - K: 38.75, 52.50, 68.75,
  - $1.Hs$: 41.25, 62.50, 89.25,
  - $2.D$: 48.75, 74.00, 99.25,
  - $3.Hy$: 42.25, 64.00, 89.75,
  - $4.Pd$: 42.25, 68.50, 95.75,
  - $6.Pa$: 44.25, 67.00, 96.25,
  - $7.Pt$: 41.75, 67.00, 88.75,
  - $8.Sc$: 45.75, 73.00, 98.75,
  - $9.Ma$: 37.25, 56.50, 83.75,
  - $0.It$: 42.25, 61.50, 75.25,
- the MDL based discretization approach (10 intervals)
  - L: 38.70, 44.10, 49.50, 54.90, 0.30, 65.70, 71.10, 76.50, 81.90, 87.30,
  - F: 47.65, 54.45, 60.75, 67.05, 73.35, 79.65, 85.95, 92.25, 98.55, 105.85,
  - K: 31.00, 37.50, 42.50, 47.50, 52.50, 57.50, 62.50, 67.50, 72.50, 79.00,
  - $1.Hs$: 36.50, 53.50, 60.50, 67.50, 74.50, 81.50, 88.50, 95.50, 102.50, 112.00,
  - $2.D$: 40.70, 57.10, 64.50, 71.90, 79.30, 86.70, 94.10, 101.50, 108.90, 116.30,
  - $3.Hy$: 37.35, 53.05, 57.75, 62.45, 67.15, 71.85, 76.55, 81.25, 85.95, 100.15,
  - $4.Pd$: 33.00, 49.50, 56.50, 63.50, 70.50, 77.50, 91.50, 84.50, 98.50, 110.50,
  - $6.Pa$: 37.65, 51.95, 59.25, 66.55, 73.85, 81.15, 88.45, 95.75, 103.05, 113.35,
  - $7.Pt$: 33.80, 50.90, 57.50, 64.10, 70.70, 77.30, 83.90, 90.50, 97.10, 103.70,
  - $8.Sc$: 35.85, 52.55, 60.25, 67.95, 75.65, 83.35, 91.05, 98.75, 106.45, 115.15,
  - $9.Ma$: 31.25, 44.75, 51.25, 57.75, 64.25, 70.75, 77.25, 83.75, 90.25, 100.75,
  - $0.It$: 31.75, 40.75, 45.25, 49.75, 54.25, 58.75, 63.25, 67.75, 72.25, 80.75,
- the MDL based discretization approach (min. 5 intervals)
  - L: 40.25, 49.50, 56.50, 62.00, 75.50, 87.75,
  - F: 52.75, 64.00, 68.00, 72.00, 78.00, 85.00, 89.50, 96.50, 106.25, 120.00,
  - K: 30.25, 36.50, 45.00, 52.00, 56.50, 64.25, 76.00, 120.00,
  - $1.Hs$: 38.25, 55.50, 58.50, 61.00, 63.50, 66.00, 70.50, 93.25, 120.00,
  - $2.D$: 44.75, 64.00, 67.50, 73.00, 85.00, 96.00, 109.75, 120.00,
  - $3.Hy$: 40.75, 58.50, 61.50, 66.00, 69.50, 73.50, 94.25, 120.00,
  - $4.Pd$: 39.75, 61.50, 65.00, 70.00, 74.50, 81.50, 103.25, 120.00,
  - $6.Pa$: 40.25, 57.50, 65.00, 70.00, 74.50, 78.50, 83.50, 97.00, 113.25, 120.00,
  - $7.Pt$: 37.75, 57.50, 61.00, 65.50, 70.50, 76.00, 85.50, 94.00, 101.75, 120.00,
  - $8.Sc$: 42.25, 62.50, 66.50, 70.00, 74.50, 81.50, 90.00, 102.50, 114.75, 120.00,
  - $9.Ma$: 33.25, 47.50, 52.50, 57.50, 61.50, 66.00, 70.50, 79.75, 97.50, 120.00,
  - $0.It$: 35.25, 53.50, 58.00, 61.00, 74.75.

In case of the K-means based approach, centers were generated for $k = 4$. In case of the equipotent interval approach, an unsupervised attribute discretization (the equal-frequency binning method) implemented in WEKA [9] was used. In this method, the same number of cases falls into each interval. Intervals are of different sizes. In case of the MDL based discretization approach, a supervised attribute discretization (the minimum length description method) implemented in WEKA was used. For more information, we refer readers to [10] and [11].

In each case, we have obtained a set (sequence) of centers located within the range $[0, 120]$. Formally, let $\{c_1, c_2, \ldots, c_{k_i}\}$ be a set of centers obtained for the i-th descriptive attribute. Triangular shaped membership functions are defined as follows.

1) For $j = 1$:

$$\mu_{c_j}(x) = \begin{cases} 1 & \text{if } x \geq 0 \text{ and } x \leq c_j, \\ 1 - \frac{x - c_j}{c_{j+1} - c_j} & \text{if } x > c_j \text{ and } x \leq c_{j+1}, \\ 0 & \text{otherwise.} \end{cases}$$

2) For $j > 1$ and $j < c_{k_i}$:

$$\mu_{c_j}(x) = \begin{cases} \frac{x - c_{j-1}}{c_j - c_{j-1}} & \text{if } x \geq c_{j-1} \text{ and } x \leq c_j, \\ 1 - \frac{x - c_j}{c_{j+1} - c_j} & \text{if } x > c_j \text{ and } x \leq c_{j+1}, \\ 0 & \text{otherwise.} \end{cases}$$

3) For $j = c_{k_i}$:

$$\mu_{c_j}(x) = \begin{cases} \frac{x - c_{j-1}}{c_j - c_{j-1}} & \text{if } x \geq c_{j-1} \text{ and } x \leq c_j, \\ 1 & \text{if } x > c_j \text{ and } x \leq 120, \\ 0 & \text{otherwise.} \end{cases}$$

Let $ls = c_j - 0.2(c_{j+1} - c_{j-1})$ and $rs = c_j + 0.2(c_{j+1} - c_{j-1})$, trapezoidal shaped membership functions are defined as follows.

1) For $j = 1$:

$$\mu_{c_j}(x) = \begin{cases} 1 & \text{if } x \geq 0 \text{ and } x \leq rs, \\ 1 - \frac{x - rs}{c_{j+1} - rs} & \text{if } x > rs \text{ and } x \leq c_{j+1}, \\ 0 & \text{otherwise.} \end{cases}$$

2) For $j > 1$ and $j < c_{k_i}$:

$$\mu_{c_j}(x) = \begin{cases} \frac{x - c_{j-1}}{ls - c_{j-1}} & \text{if } x \geq c_{j-1} \text{ and } x \leq ls, \\ 1 & \text{if } x > ls \text{ and } x < rs, \\ 1 - \frac{x - rs}{c_{j+1} - rs} & \text{if } x \geq rs \text{ and } x \leq c_{j+1}, \\ 0 & \text{otherwise.} \end{cases}$$

3) For $j = c_{k_i}$:

$$\mu_{c_j}(x) = \begin{cases} \frac{x - c_{j-1}}{ls - c_{j-1}} & \text{if } x \geq c_{j-1} \text{ and } x \leq ls, \\ 1 & \text{if } x > ls \text{ and } x \leq 120, \\ 0 & \text{otherwise.} \end{cases}$$

It is worth noting that, in each approach, intervals are disjoint. However, one can see that membership functions overlap.

Let $DT = (U, Attr, Dec)$ be a decision table containing MMPI data, where $Attr = \{A_1, A_2, \ldots, A_{13}\}$ and $Dec = \{D\}$. After fuzzification, for each descriptive attribute A_i, where $i = 1, 2, \ldots, 13$, we obtain k_i fuzzified attributes $A_i^1, A_i^2, \ldots, A_i^{k_i}$, where k_i is a number of linguistic values used for fuzzification of A_i. In case of the decision attribute D, we obtain 20 fuzzified attributes, each for one decision class, i.e., $D^{norm}, D^{neur}, \ldots, D^{dev6}$. However, values of the attributes $D^{norm}, D^{neur}, \ldots, D^{dev6}$ are binary. For example, $D^{norm}(u) = 1$ if $D(u) = norm$, and 0 otherwise, where $u \in U$.

B. Fuzzy Decision Trees

To build the classifier, we have used the algorithm for generation of fuzzy decision trees that uses cumulative information estimations of the initial data proposed by V. Levashenko et al. (see [4]). This algorithm was used by us in our preliminary experiments with the MMPI data (see [2]). The obtained results were very promising in relation to results obtained for other kinds of classifiers (cf. [3]). In general, the cumulative mutual information for a given attribute A_i, a sequence of attributes SFA, and the decision attribute D reflects the influence of the attribute A_i on the attribute D when the sequence SFA of attributes is known.

In this section, we briefly recall the algorithm used in our experiments. Let us assume the following notation: U - the set of cases, n - the number of cases, $lval(A)$ - the set of all linguistic values used for the fuzzification process of the attribute A, $cer(D^v)$ - the certainty of the decision class D^v of the attribute D, RA - the set of the remaining descriptive attributes, SFA - the set of the selected fuzzified attributes. The algorithm is recursive (see Procedure 1). There are two tuning parameters θ_{freq} and θ_{cer} used in the algorithm as the stop conditions. Expanding a tree branch is stopped when either the frequency of the branch is below θ_{freq} or when more than or equal to θ_{cer} percent of cases left in the branch has the same decision class label. Moreover, the natural stop condition is fulfilled if the set of the remaining descriptive attributes is empty (i.e., $RA = \emptyset$).

The cardinality measure of the set B of fuzzified attributes is defined as $card(B) = \sum_{u \in U} \prod_{B_i \in B} B_i(u)$. The certainty $cer(D^v)$ of the decision class D^v is calculated as $cer(D^v) = card(SFA \cup \{D^v\})$.

Procedure FDT

Data: A decision table $DT = (U, Attr, Dec)$

$RA \leftarrow Attr; SFA \leftarrow \emptyset;$

$E(D) \leftarrow n \log(n) - \sum_{v \in lval(D)} card(\{D^v\}) \log(card(\{D^v\}));$

foreach $A_i \in RA$ **do**

> $E(A_i) \leftarrow n \log(n) - \sum_{v \in lval(A_i)} card(SFA \cup \{A_i^v\}) \log(card(SFA \cup \{A_i^v\}));$
>
> $E(D, A_i) \leftarrow n \log(n) - \sum_{v \in lval(D), w \in lval(A_i)} card(SFA \cup \{D^v\} \cup \{A_i^w\}) \log(card(SFA \cup \{D^v\} \cup \{A_i^w\}));$
>
> $CMI(A_i) \leftarrow E(D) + E(A_i) - E(D, A_i);$

Select A_i from RA with the greatest $CMI(A_i)$;

> $RA \leftarrow RA - \{A_i\};$

foreach $v \in lval(A_i)$ **do**

> $SFA \leftarrow SFA \cup \{A_i^v\};$
>
> **if** $\max_{cer(D_v)} < \theta_{cer}$ *and* $\frac{card(SFA)}{n} \geq \theta_{freq}$ *and* $RA \neq \emptyset$ **then**
>
> > call FDT with $DT = (U, RA, Dec);$
>
> **else**
>
> > create a decision node;

C. The CLAPSS System

All of the stages of the classification process (fuzzification of the input data, generation of the classifier, testing the classifier) were performed using our software tool called CLAPSS (Classification and Prediction Software System) [12] that is a tool developed for solving different classification and prediction problems using, among others, some specialized approaches based mainly on fuzzy sets and rough sets. A new module added to CLAPSS consists of implementation of the selected methods based on fuzzy sets (especially to creation of classifiers based on fuzzy decision trees generated by means of the algorithm described in Section III-B). The user that creates classifiers based on fuzzy decision trees can select among others: a fuzzification process (*triangular, trapezoidal, Gaussian*), thresholds (*certainty* and *frequency*) to stop the process of fuzzy decision tree creation, *t*-norm

(*minimum*, *algebraic product*, *Lukasiewicz product*, *Einstein product*, *Hamacher product*, *drastic product*) for calculation of the certainty of rule antecedents, and a number of folds for the cross-validation procedure.

IV. RESULTS

Our experiments were performed on real-life data described in Section II using the CLAPSS software tool (see Section III-C). In each experiment, the $5.Mf$ scale was excluded. This scale is assumed by the experts to be the weakest one. In each case, the stratified 10-cross-validation approach was used to test the classifier. In the experiments, the following settings have been used: a shape of membership functions: *triangular* and *trapezoidal*, the certainty threshold: 0.999, the frequency threshold: 0.001, the t-norm for calculation of the certainty of rule antecedents: *algebraic product*. The results of the stratified 10-cross-validation tests for all approaches (both expert and inductive) are presented in Figure 1 (for triangular shaped membership functions) and Figure 2 (for trapezoidal shaped membership functions). The results showed that ex-
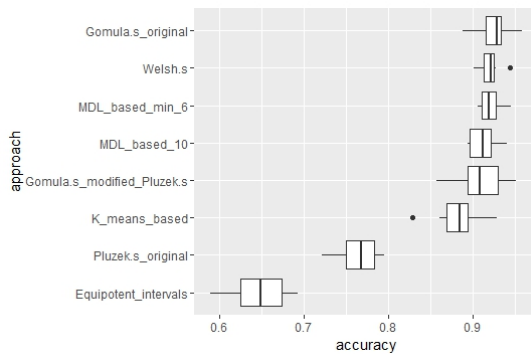

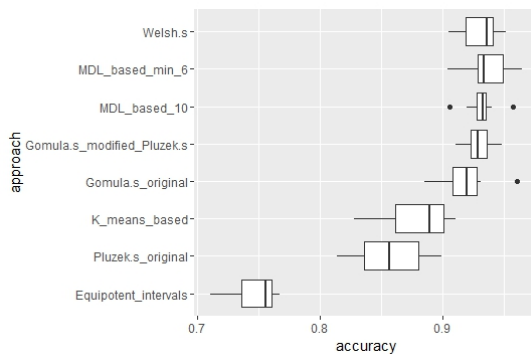Fig. 1. Results for triangular shaped membership functions.

Fig. 2. Results for trapezoidal shaped membership functions.

pert approaches like Gomuła's (original), Gomuła's (modified Płużek's), and Welsh's are suitable for the fuzzy decision tree based classification. Among inductive approaches, good results were obtained for MDL based multi-interval discretization. It is worth noting that easy and natural diagnostic interpretation of the obtained intervals becomes the advantage of the expert approaches. Weak results obtained for the original Płużek's

intervals do not seem to be surprising because this approach is recognized by the experts as rough. The approach based on equipotent intervals turned out to be inappropriate. On the basis of the results, one can see that classifiers based on fuzzy decision trees show a high effectiveness (accuracy noticeably greater than 0.9) in classification of the MMPI data. If we take into consideration solely the MMPI scales (without any additional indexes), only a few previously tested approaches are found to be such effective (cf. [3]).

V. CONCLUSIONS AND FURTHER WORK

In general, classifiers based on fuzzy decision trees showed a high effectiveness in classification of the MMPI data. The main challenge in the future is to propose the method for searching for optimal intervals used in the fuzzification process. Simultaneously, we need to take care of diagnostic interpretation of the obtained intervals. Therefore, automated searching for optimal intervals should be aided with the expert knowledge. This fact determines the main direction of our further research. Moreover, we plan to test application of some other shapes of membership functions and some other t-norms for calculation of the certainty of rule antecedents.

REFERENCES

[1] K. Pancerz, O. Mich, A. Burda, and J. Gomuła, "A tool for computer-aided diagnosis of psychological disorders based on the MMPI test: An overview," in *Applications of Computational Intelligence in Biomedical Technology*, ser. Studies in Computational Intelligence, R. Bris, J. Majernik, K. Pancerz, and E. Zaitseva, Eds. Cham: Springer International Publishing, 2016, vol. 606, pp. 201–213.

[2] V. Levashenko, E. Zaitseva, K. Pancerz, and J. Gomuła, "Fuzzy decision tree based classification of psychometric data," in *Position Papers of FedCSIS'2014*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. Maciaszek, and M. Paprzycki, Eds., vol. 3. Warsaw, Poland: PTI, 2014, pp. 37–41.

[3] O. Mich, A. Burda, K. Pancerz, and J. Gomuła, "The knowledge base for computer-aided diagnosis of mental disorders based on psychometric tests," in *Proceedings of DT'2014*, Zilina, Slovakia, 2014. doi: 10.1109/DT.2014.6868724 pp. 255–261.

[4] V. Levashenko and E. Zaitseva, "Usage of new information estimations for induction of fuzzy decision trees," in *Proceedings of IDEAL 2002*, ser. LNCS, H. Yin, N. Allinson, R. Freeman, J. Keane, and S. Hubbard, Eds., vol. 2412. Springer Berlin Heidelberg, 2002. doi: 10.1007/3-540-45675-9_74 pp. 493–499.

[5] D. Lachar, *The MMPI: Clinical assessment and automated interpretations.* Fate Angeles: Western Psychological Services, 1974.

[6] M. Choynowski, *Wielowymiarowy inwentarz osobowości (in Polish)*, Psychometry Laboratory, Polish Academy of Sciences, Warsaw, 1964.

[7] Z. Płużek, "Wartość diagnostyczna testu WISKAD-MMPI w zakresie nozologii psychiatrycznej (in Polish)," *Roczniki Filozoficzne / Annales de Philosophie / Annals of Philosophy*, vol. 17, no. 4, pp. 125–143, 1969.

[8] H.-M. Lee, C.-M. Chen, J.-M. Chen, and Y.-L. Jou, "An efficient fuzzy classifier with feature selection based on fuzzy entropy," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 31, no. 3, pp. 426–432, 2001. doi: 10.1109/3477.931536

[9] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques.* Morgan Kaufmann, 2005.

[10] U. Fayyad and K. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proceedings of IJCAI'1993*, Chambery, France, 1993, pp. 1022–1029.

[11] I. Kononenko, "On biases in estimating multi-valued attributes," in *Proceedings of IJCAI'1995*, Montreal, Quebec, Canada, 1995, pp. 1034–1040.

[12] K. Pancerz, "On selected functionality of the classification and prediction software system (CLAPSS)," in *Proceedings of IDT'2015*, Zilina, Slovakia, 2015. doi: 10.1109/DT.2015.7222984 pp. 278–285.