Simultaneous Differential Diagnoses Basing on MMPI Inventory
Using Neural Networks and Decision Trees Methods

Norbert Jankowski and Jerzy Gomuła

Department of Computer Methods & Psychology Outpatient Clinic
Nicholas Copernicus University

ul. Grudziądzka 5, 87–100 Toruń, phone: +48 56 6113307,
{norbert,gomula}@phys.uni.torun.pl, http://www.phys.uni.torun.pl/˜norbert

Problem description: Our research concerns psychometric diagnoses basing on the Minnesota Multiphasic Personality Inventory (MMPI) test [1, 3, 4]. The goal of MMPI test is to count the personality–psychometric dimensions which help in differential diagnosis of given person and to assess psychotherapy results. Finally given person should be assigned to one (or two if given case is placed on a border) of personality types. Unfortunately even after the psychometric (MMPI) dimensions are obtained the diagnosis is still so complex for many cases. Our goal is to build a model using neural networks and machine learning technics which could be able to diagnose nosological type and will reduct above difficulties of MMPI test. It is also important to evaluate the influence of each scale in a given patient diagnosis, present alternative classes and estimate the accuracy of such diagnosis.

MMPI test consists of 556 questions. Some of them repeat oneself. About 200 questions of MMPI test do not have any clinical value, but are helpful in constructing control scales (dimensions). Using the answers from each MMPI test 13 numerical factors are computed (by some arithmetic operations) forming the intermediate basis (scales, not the final hypothesis) for the diagnosis. The first 3 scales (L, F, K) form the control part and next 10 dimensions form the clinical part (hypochondria–Hp, depression–D, hysteria–Hy, psychopathy–Ps, masculinity–Mf, paranoia–Pa, psychasthenia–Pt, schizophrenia–Sc, manic–Ma, social introversion–It/Si). Unfortunately clinical scales count the similarities to given nosological types in differential diagnoses than the exact value of hypochondria, depression, etc. This is the reason of difficulties in the final classification of nosological type.

Personality types split the space over norm$^{(1)}$, neurosis$^{(2)}$, psychopathy$^{(3)}$, organic$^{(4)}$, schizophrenia$^{(5)}$, syndrome delusion$^{(6)}$, reactive psychosis$^{(7)}$, paranoia$^{(8)}$, manic state$^{(9)}$, criminality$^{(10)}$, alcoholism$^{(11)}$, drag induction$^{(12)}$, simulation$^{(13)}$, dissimulation$^{(14)}$, deviational answering style 1$^{(15)}$, deviational answering style 2$^{(16)}$, deviational answering style 3$^{(17)}$, deviational answering style 4$^{(18)}$, deviational answering style 5$^{(19)}$, deviational answering style 6$^{(20)}$. Polish version of MMPI was designed by M. Choynowski [4] and Z. Płużek [8, 9]. The polish version differ in cultural questions which were removed and substituted.

Data sets – learning population: 4 data sets were collected and classified by psychologists. Each case from these data sets may be assigned to one of 20 classes for women and 27/28 (used in result comparison only)for women and men respectively. Each case is assigned to one of presented above nosological type. Data sets consist of 1027 and up to 1711 examples. Population of each data set were collected very carefully to satisfy several conditions which concern the MMPI test. Every case classified by three clinic psychologists and only noncontroversial patients were chosen. Database cases were collected from several psychiatric hospitals and clinics. Below the distribution of one data set is presented — 20 classes set (women set). See [2] for more details. Data sets were collected by T. Kucharski and J. Gomuła from Psychological Outpatient Clinic.

Distribution of patient population over nosological type — women data set

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
269	218	42	78	459	21	37	32	35	81	23	52	203	43	41	26	14	13	12	12

Differential diagnoses using classification models. We used two different classification systems: Incremental Network (IncNet) [6, 5, 7] and decision tree C 4.5 for rule extraction [10].

Structure of incremental neural network is controlled by growing and pruning to match the complexity of training data. Extended Kalman Filter algorithm and its fast version is used as learning algorithm. Bi-radial transfer functions, more flexible than other functions commonly used in artificial neural networks, are used. See [6, 5, 7] for more datails. Table 1 compares generalization performance of IncNet and C 4.5 for to data sets: 27– and 28–class in cross-validation tests. Results clearly show the higher performance of IncNet model.

Probabilistic confidence intervals answer the question: *How* the probabilities of the *winner* and *alternative* classes change as a function of attribute values for different input dimensions. Displaying such probabilities the *probabilistic intervals of confidence* (PIC) are obtained.

Assuming that the values of other features are held constant for a given case (patient) $\mathbf{x} = [x_1, x_2, \ldots, x_N]$ three probabilities for each feature r are important and will be visualized in analysis of a given case (cf. Fig. 2 and 3).

Model	27–classes				28–classes			
	10%		5%		10%		5%	
	TRS	TES	TRS	TES	TRS	TES	TRS	TES
IncNet (neural network)	**99.03**	**93.14**	**98.77**	**96.08**	**98.95**	**93.10**	**98.29**	**94.83**
C 4.5 (decision trees)	93.22	83.70			93.13	78.90		

Table 1: Comparison of IncNet and C 4.5 for two data sets. TRS – training set, TES – testing set.

First probability (solid curve) is the probability of the **winning class** defined by $p(C(\mathbf{x})|\bar{\mathbf{x}}; M)$. Probability $p(C^i|\mathbf{x}; M)$ is computed directly from IncNet networks for each class. Note that such probability changes for different values of $\bar{\mathbf{x}} = [x_1, \ldots, x_{r-1}, \bar{x}, x_{r+1}, \ldots, x_N]$.

The next probability displayed (a dotted curve) is the probability $p(C^{k_2}|\bar{\mathbf{x}})$ of the most probable **alternative class**, where class index k_2 is defined by $k_2 = \arg\max_i \{p(C^i|\mathbf{x}; M),\ C^i \neq C(\mathbf{x})\}$The k_2 class is determined for the point $\mathbf{x}$ only.

The third probability (dashed line) presents the probability $p(C^{k_M}|\bar{\mathbf{x}})$of the most probable ***variable*** **alternative class** at the point $\bar{\mathbf{x}}$. The index k_M is defined by $k_M = \arg\max_i \{p(C^i|\bar{\mathbf{x}}),\ C^i \neq C(\mathbf{x})\}$and may change, while index k_2 does not change.

These three probabilities carry all information about the case given for analysis, showing the stability of classification against perturbation of each feature and the importance of alternative classes in the neighborhood of the input $\mathbf{x}$. Probabilistic confidence intervals in action are shown in the next section. In other words the PIC present dimension–sensitivity in diagnosis of given patient $\mathbf{x}$.

Another advantage of PIC is that movie animations may be constructed to present the evolution of therapy for given person (and/or disease). Browsing the animations smooth change of diagnosis and PIC may be observed.

Information on winner and alternative classes is continuous and very precise in uncertainty estimation. Confidence interval shows neighboring alternative classes (if they exist). The distance from the considered case to decision borders may be analyzed in this way. Analysis of complex cases, which often lie on the decision border, is much more reliable using probabilistic confidence intervals than logical rules. It is very easy to find which features are important and which may be omitted in investigation.

Logical rules from C 4.5 The final set or rules extracted from C 4.5 for 20 classes (women) data set tree consists of 53 rules. Comparison on training and testing sets were presented in Tab. 1. Below a few selected examples of extracted rules are presented (note $\wedge$ means *AND*):

$$R9: \quad if \quad L \in [36,67] \wedge K \in [27,59] \wedge Hp \in [23,66] \wedge Hy \in [24,63] \wedge Pd \in [20,67] \wedge Mf \in [49,95] \wedge Pt \in [55,107] \wedge Sc \in [21,63] \wedge Ma \in [25,65] \wedge It \in [50,87] \quad then \quad \text{Norm}$$

$$R10: \quad if \quad L \in [36,59] \wedge Hp \in [23,66] \wedge Hy \in [24,63] \wedge Mf \in [49,95] \wedge Sc \in [23,63] \wedge It \in [50,61] \quad then \quad \text{Norm}$$

$$R20: \quad if \quad Hy \in [61,112] \wedge Mf \in [20,48] \wedge Sc \in [23,63] \quad then \quad \text{Nneurosis}$$

$$R49: \quad if \quad Hp \in [23,64] \wedge Pd \in [20,72] \wedge Pa \in [27,79] \wedge Pt \in [20,78] \wedge Sc \in [70,111] \wedge Ma \in [56,108] \wedge It \in [56,87] \quad then \quad \text{Schizophrenia}$$

$$R52: \quad if \quad Hp \in [23,64] \wedge Pd \in [20,68] \wedge Pt \in [20,78] \wedge Sc \in [70,111] \wedge It \in [56,87] \quad then \quad \text{Schizophrenia}$$

Diagnosis example for selected patient. Nontrivial example of women (J.K., 28 age, cf. [9]) schizophrenia has been chosen as nosological diagnosis example. This person was tested twice: before therapy (during first schizophrenia detection) and two years later (after therapy and schizophrenia come back). After collecting answers from MMPI test and computing scales this person can be characterized by (all values are in T-scores):

Profile J.K._1: L=57, F=60, K=53, H=63, D=65, Hy=66, Pd=60, Mf=56, Pa=48, Pt=65, Sc=74, Ma=50, It=63
Profile J.K._2: L=53, F=67, K=50, H=49, D=67, Hy=62, Pd=66, Mf=44, Pa=60, Pt=75, Sc=78, Ma=60, It=65

We made two simultaneous differential diagnoses. First using IncNet and PIC model and second using logical rules based on C 4.5 decision tree. All systems previously were learned using 20 classes data base designed for women. In Fig. 2 PIC of first examination was presented. Note that small square picture is the patient position on each scale. Browsing described figure it is clear that this person was assigned to schizophrenia class with so high probability (0.95), but reader should pay attention to sub-figure 11 which describes the schizophrenia scale. This sub–figure shows that the patient (small square) is not *very* far from the left border of this scale, but decreasing value of this scale could change diagnose to *unsafe* region (nearly unreal cases or not covered by data base — probability on axis OY is too small for all classes). This may mean the beginning of psychotic process (schizophrenisation, see schizophrenia scale). The rest of scales look safe and have no crucial influence to the final nosological type. Next figure (3) presents PIC for the same patient after two year break. Again the final diagnosis is schizophrenia with probability 0.97. On sub–figure 11 the alternative class norm can be seen, but the patient is far from the dashed line which describes (this time) the norm. Also on hypochondria scale the alternative class can be seen — organic disorders — so far from considered patient. Sub–figure 12 (manic scale) has an alternative class too, but this one is closer than previous. But even with mentioned last alternative class the diagnosis looks faithfull and is (a bit) more probable than the previous one. Decreased hypochondria in second examination may designate the chronic schizophrenia.

Figure 1 shows two diagrams. The left one presents first examination result with a logical rule which covers considered patient. This rule belong to schizophrenia. Rule is presented by vertical intervals (thick line). Each column describe one scale. Patient position is shown by line with circles. The right diagram present results collected two years latter. Rule presented on this diagram also belong to schizophrenia.

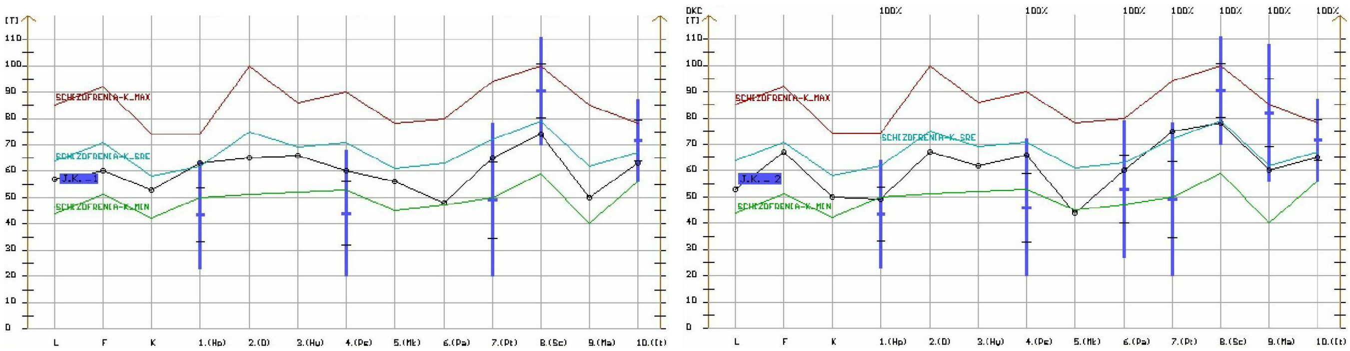

Figure 1: Diagnosis using logical rules: left first examination (rule accuracy 98.4%, rule cover 3.5% – R49), right second two years later (rule accuracy 99.1%, rule cover 60.4% – R52). All cases are assigned to schizophrenia.

Possible extensions and complementary support. We plan to extend the diagnose by another type of confidence intervals partially similar to logical rule but with continuous belong function.

The diagnosis using the IncNet with PIC and logical rules will be complement by several classificational and descriptional factors (cf. [2]) which for example illustrate neuroticism, psychoticism, anxiety, types of neurosis, disturbances of character, control of impulses and many others.

Conclusions. Probabilistic confidence intervals defined above are a new tool which may be very useful in the process of diagnosis. The most important is that probabilistic confidence intervals are constructed (on-line) for a given case basing on previously estimated model — in contrary to logical rule. Information on winner and alternative classes is continuous and very precise in uncertainty estimation. Confidence intervals show neighboring alternative classes (if they exist). The distance from the case considered to decision borders may be analyzed in this way. Analysis of complex cases, which often lie on the decision border, is much more reliable using probabilistic confidence intervals than logical rules. It is very easy to find which features are important and which may be omitted.

Properties of probabilistic intervals of confidence make them a very useful diagnostic tools. Artificial neural networks may be interpreted using such tools, breaking the myth that neural networks are *black boxes*. And logical rules may be used as complementary diagnosis tools.

References

[1] J. N. Buther, W. G. Dahlstrom, J. R. Graham, A. Tellegen, and B. Kaem-mer. *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. University of Minnesota Press, Minneapolis, 1989.

[2] J Gomuła and T. Kucharski. Wprowadzenie do kwestionariusza MMPI-WISKAD. Toruń, 1998.

[3] S.R. Hathaway and J.C. Mc Kinley. *The Minnesota Multiphasic Personality Inventory Manual Revised*. The Psychological Corporation, New York, 1951.

[4] S.R. Hathaway and J.C. Mc Kinley. Wielowymiarowy inwentarz osobowości. autoryzowany przekład mieczysława choynowskiego. Technical Report T-7 (zeszyt testowy), Pracownia Psychometryczna PAN, Warszawa, 1964.

[5] N. Jankowski. Approximation and classification in medicine with IncNet neural networks. In *Machine Learning and Applications. Workshop on Machine Learning in Medical Applications*, pages 53–58, Chania, Greece, July 1999.(PDF).

[6] N. Jankowski. *Ontogenic neural networks and their applications to classification of medical data*. PhD thesis, Department of Computer Methods, Nicholas Copernicus University, Toruń, Poland, 1999. (PDF).

[7] N. Jankowski and V. Kadirkamanathan. Statistical control of growing and pruning in RBF-like neural networks. In *Third Conference on Neural Networks and Their Applications*, pages 663–670, Kule, Poland, October 1997.

[8] Z. Płużek. Polska adaptacja testu MMPI. In *I Konferencji Psychologii Klinicznej w Łomnicy Tatrzańskiej*, 1964.

[9] Z. Płużek. *Wartość testu WISKAD-MMPI*. KUL, Lublin, 1976.

[10] J. Quinlan. Programs for machine learning, 1993.

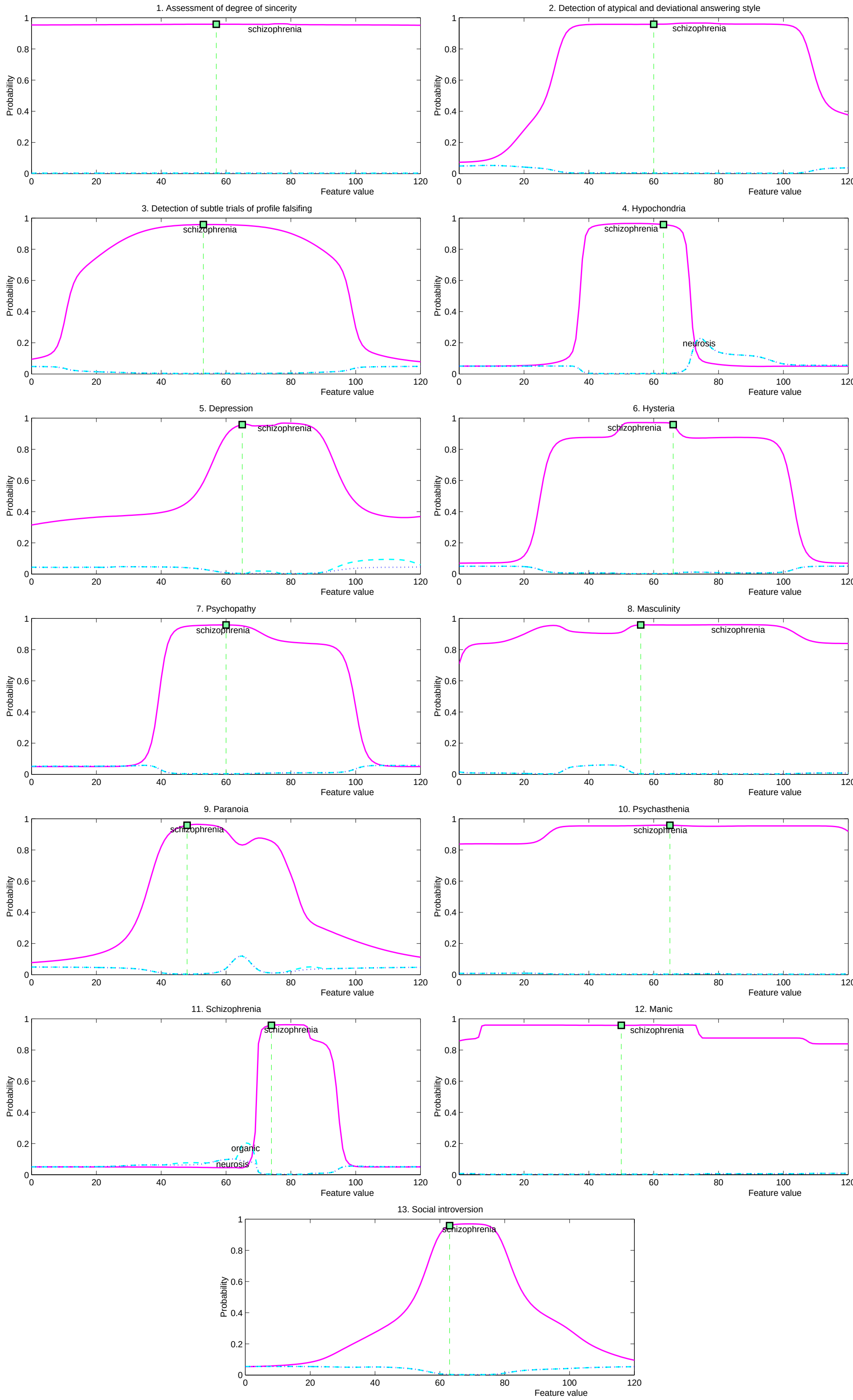

Figure 2: First examination – J.K._1: Class: Schizophrenia (prob. 0.95). Axis OY displays the class probability. Probability equal to 1 mean 100% of given disease, and probability 0 mean 0%.

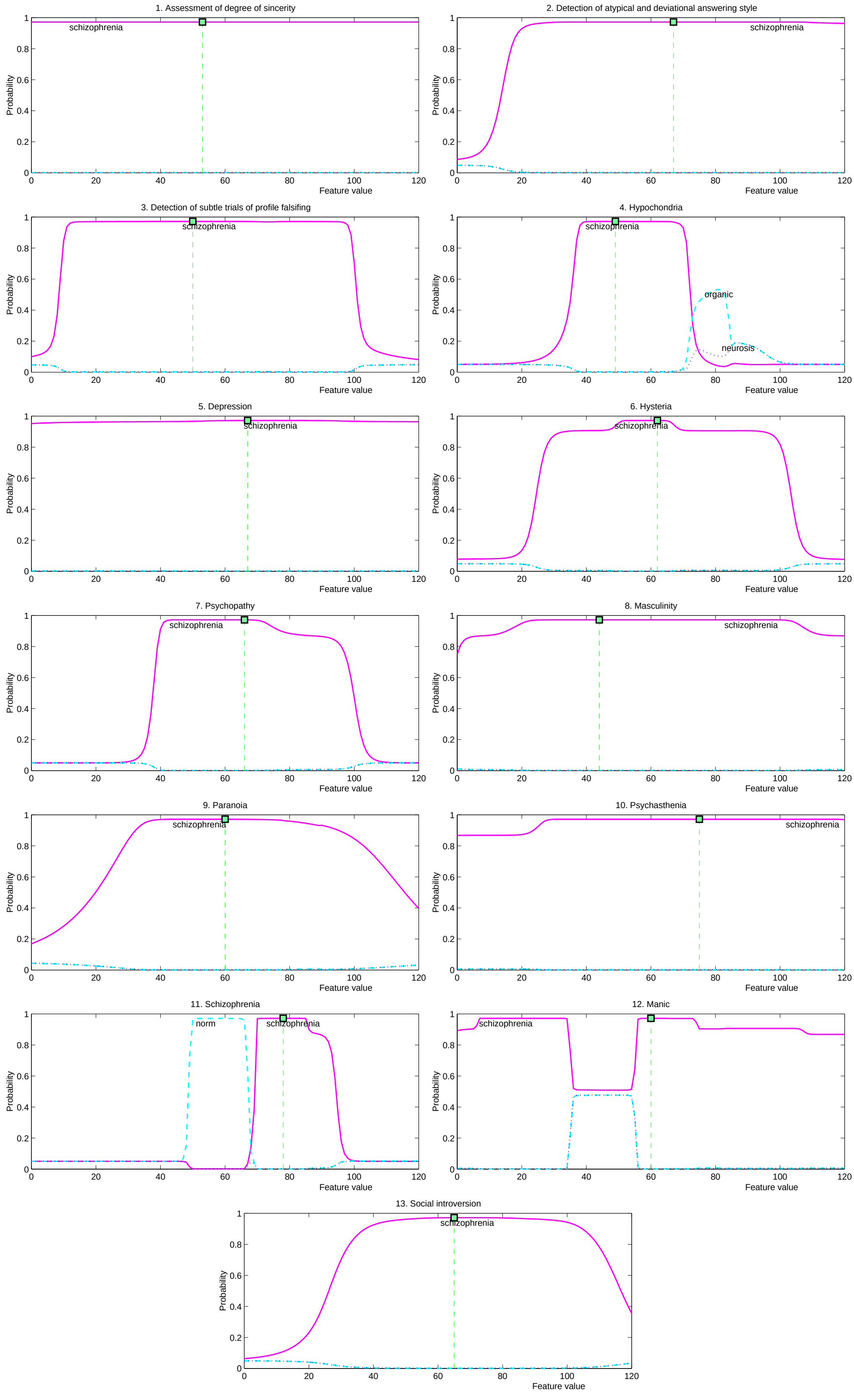

Figure 3: Second examination – J.K._2: Class: Schizophrenia (prob. 0.97). Axis OY displays the class probability. Probability equal to 1 mean 100% of given disease, and probability 0 mean 0%.