

# PRETEEN AGE: THE ANALYSIS OF THE MULTILEVEL PSYCHO-DIAGNOSTIC DATA BASED ON NEURAL NETWORK MODELS

**Elena Slavutskaya**

Chuvash State Pedagogical University, Russian Federation

**Leonid Slavutskii**

Chuvash State University, Russian Federation

**Abstract.** *The use of the artificial neural network (ANN) models for vertical system analysis of psycho-diagnostic data is suggested. It is shown that the ANN training, as the problem of nonlinear multi-parameter optimization, allows to create effective algorithms for the psycho-diagnostic data processing when the results of psychological testing for the different level's characteristics have different numerical scales. On the example of processing the author's data of psycho-diagnostics (preadolescent schoolchildren), it is shown that neural network models can be used to estimate latent (hidden) connections between psychological characteristics. The proposed algorithms are based on a statistical assessment of the quality of such models, do not require a large sample of respondents. The quantitative statistical criteria for evaluating the quality of the models are estimated. The approach is sufficiently clear for practical use by psychologists who do not have a special mathematical preparation.*

**Keywords:** *preadolescent schoolchildren, psychodiagnostic data, vertical system analysis, artificial neural networks.*

## Introduction

For the psychological researches, it is important to assess the presence of indirect (latent, hidden) links between the diagnosed traits. Vertical system analysis in psychology implies, as a rule, finding connections between the psychological characteristics of different levels (Hebb, 1961). That is, at the same time, data processing can include solving the problem of classifying or the constructing of a hierarchical model. The traditional for psychology the correlation and factor analysis (Glass & Stanly, 1970), as the mathematical tools, allow to evaluate statistically the presence of only linear connections between psychodiagnostic data and do not give a result in the event that such connections are nonlinear. These tasks can be solved using artificial neural networks (ANN)(Haykin, 1999).

In this paper, the task is not to find the numerical meanings of psychological traits, but only to establish links between them (Baxt, 1994). This makes it possible to use the itself analyzed data for ANN training. In this case, the sample

can be quite limited (up to 50-100 subjects), the main thing is that the number of neurons in each ANN layer should be much smaller than the sample size (Grossberg, 1974). The numerical dimension of the initial psychological characteristics is not of fundamental importance, the signs can even have symbolic meanings (for example, a sex sign). Thus, the use of ANN makes it possible to establish latent and nonlinear links between psychodiagnostic data, even if they are obtained by tests with different numerical scales. To assess the relations between the characteristics of different levels, it is proposed, in particular, to evaluate the quality of the training of the corresponding ANN models. Training of the neural network requires professional skills, and the overwhelming majority of psychologists prefer to use the traditional statistical methods of analysis (Glass & Stanly, 1970). Most often, the ANN is used to automate the procedure itself of medical and psychological diagnosis (Reznichenko et al., 2013) or for creating the models in psychophysiology (Usher & Zakay, 1993).

Preadolescent age is quite complex and saturated from a psychological point of view. J. Piaget defines this period as the time of transition from the stage of concrete to the stage of formal operations of thinking (Piaget, 1972). Reflexive abilities of 10-12-year-old schoolchildren are manifested in the ability to deeper generalization, analysis, the ability to explain abstract concepts. It was revealed that by the age of ten, schoolchildren can analyze complex emotions, understand the ambivalence of feelings and the relativity of moral judgments (Harter, 1990). Many researchers note the peculiarities in the development of the personality of children 9-12 years old. The development of moral consciousness of children aged 10-13 is at the conventional level. By the end of this period, their judgments are made on the basis of established order, with respect for norms and rules, and subordination of authority (Kohlberg, 1984). From the point of view of the psychoanalytic theory of S. Freud, the age from 5 to 10 years is a latent stage in the child's psychosexual development. E. Erikson, focusing on the social stages of development, considered the age of 6-12 years as the period of transfer to the child of systematic knowledge and skills. In children at this age, the ability to master the environment surrounding it most intensively develops (or does not develop). With a positive outcome of this stage of development, the child develops an experience of his own skill, with an unsuccessful - a sense of inferiority and inability to be on a par with other people (Erikson, 1950). W. Collins argues that between 9 and 12 years there is such a serious crystallization of personality structures that many psychological and social characteristics of schoolchildren of this age give a basis for reliable predictions of his behavior for 4-6 years ahead. Data on a child up to grade 3 do not have predictive power (Collins, 1984). Thus, the study of the multilevel psychological

characteristics of children and the interrelationships of these characteristics in preadolescents are of considerable interest.

### **Sample and Psychodiagnostic Data**

The study used the standardized methods of psychodiagnostics, the choice of which was determined by the following criteria: the identification of a variety of different levels psychological indicators of schoolchildren, a sufficient number of signs for the intellectual analysis of data and for the subsequent system analysis of the characteristics of mental development in this age segment.

The investigation involved 90 schoolchildren of the 5th grade (age 11-12).

The psychodiagnostics of the following multilevel psychological characteristics of schoolchildren was carried out:

- the highest, personality level is represented by the indicators of relations: R1 - the relation to immediate family- the father, the mother; R2 - relationships with peers; R3 - attitude towards the school, teachers; R4 - attitude to yourself (test "Unfinished Sentences" by V. Mikhal for children aged 7-12) and motivational characteristics ("Need for Achievements"(NA) by Yu.M. Orlov);
- the properties of the second level were tested using the R. B. Cattells 12-factor questionnaire (twelve personal traits for children 8-12 years old (Cattell, 1990));
- the level of mental processes and states was determined through the free cultural intellectual test by R.B. Cattell ("fluid intelligence", expressed in the form of an IQ (Cattell, 1990));
- the type of the nervous system according to the psychomotor parameters (the tapping T- test of E.P. Ilyin) was considered as the basic lower level of innate (biological) characteristics and the physiological support of mental processes.

Creating the free cultural test of intelligence, R.B. Cattell suggested that it would identify a biologically predetermined intellect, formed on the basis of the architectonics of the cortex of the cerebral hemispheres. Therefore, in our study we combined the indicators of biological capabilities and the characteristics of providing mental processes and states to one level.

### **Method of Neural Network Analysis**

Based on the theoretically and empirically established rules for the use of ANN in analyzing the data, we will rely on the following general provisions:

1. An adequate ANN learning implies the presence, structure and certain stability of the connections between the input and output data of a neural network, even if these links are indirect and nonlinear.

2. The quality of the ANN model can be estimated from the scattering diagram, the statistical distribution of errors in the training and testing of the ANN.
3. It is better to use several ANN with one "output" than one ANN with several "outputs".

ANN is a network of single neurons connected to each other. The number of neurons and the scheme of their connections can be different. We use the most common type of ANN - direct propagation (feed forward), in which neurons are combined into layers (Figure 1). The ANN scheme consists of one input layer, one "hidden" layer and one output layer. Neurons in the layer are not connected to each other, but they are connected to the neurons of the previous and next layer according to the principle "every with each".

The analytical program "Deductor" was used, which contains tools for assessing the quality of data, cleaning them from duplicates and contradictions and "anomalous" values, preliminary analysis of data (partial processing, factor and correlation analysis), Data Mining, including ANN, various means of visualizing the results of analyzed models.

An example of the evaluation of the ANN training quality is shown in Figure 1. In the network input values are personality traits (Cattell, psychological characteristics of the second level). IQ in the output of ANN refers to the psychological characteristics of the baselevel (individual psychological, biologically determined characteristics).

Interpretation of the personality traits in Cattells 12-factor questionnaire (CPQ): A - gregariousness - isolation; B - abstract-concrete thinking; C - emotional stability -instability; D – excitement - balance; E-independence - obedience; F - carefree - concern; G - high-low discipline; H - boldness-shyness; I - softness-hardness; O - anxiety-calmness; Q3 - high-low self-control; Q4 - tension-relaxation.

During ANN training, various input data (in this case, Cattell's personality traits) are fed to the ANN input level, and the obtained IQ OUT values are compared with the known IQ IN target values from the psychodiagnostic data. The "Synaptic Weights", that contribute the most to the error, are modified (using the backpropagation error algorithm). This cycle is repeated many times until an acceptable accuracy of displaying the input information in the output (or - the error ceases to decrease) is achieved. The number of cycles can reach 10000.

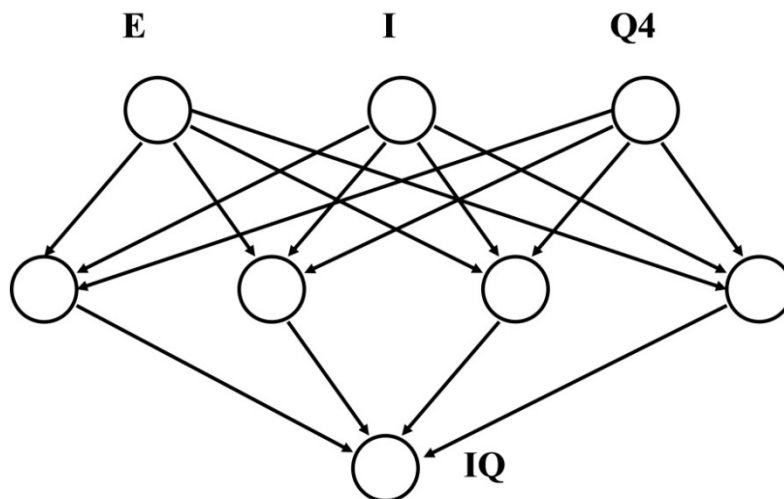


Figure 1. Example of ANN

Figure 1 shows a neural network model, created separately for boys. In this case, the input of the neural network is consistently given values of three personality traits (three neurons in the input layer). The number of neurons in the "hidden" layer (4) is chosen empirically. It corresponds to the best quality of training by the criteria, which are further described in this article. The one neuron in the output layer of the network, in this case corresponds to the intelligence coefficient.

The quality of the neural network model can be estimated from the scattering diagram, which presents the spread of the values IQ OUT during the network learning. In the "ideal" training of ANN, all points must "fit" the shown straight line (see Figure 2). The quantitative criteria can be the root mean-square deviation (error) and the percentage of recognizable links for a given error.

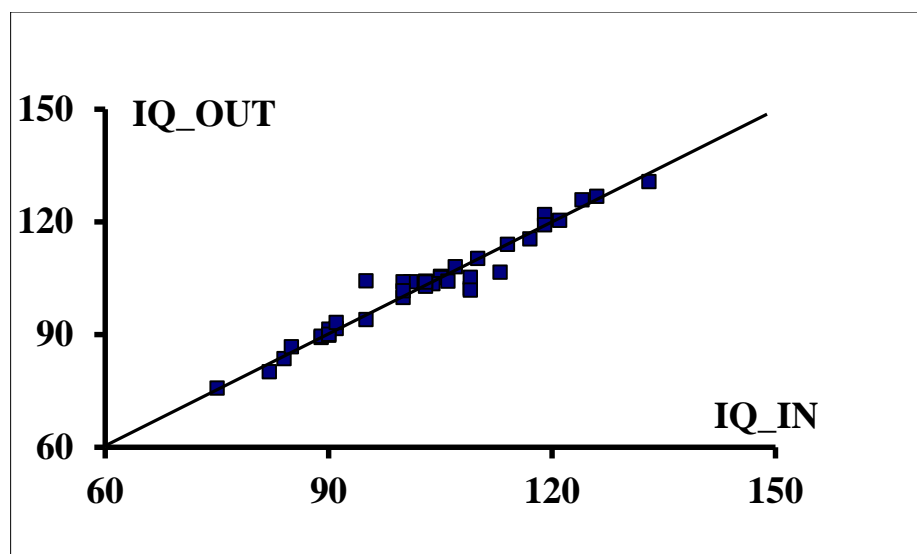


Figure 2. Example of the scattering diagram (boys)

An example of the root-mean-square error distribution over the intervals is shown in Figure 3. Comparing such histograms for different models, you can choose the best one, i.e., the model with the most preferred kind of distribution function (distribution density) of errors.

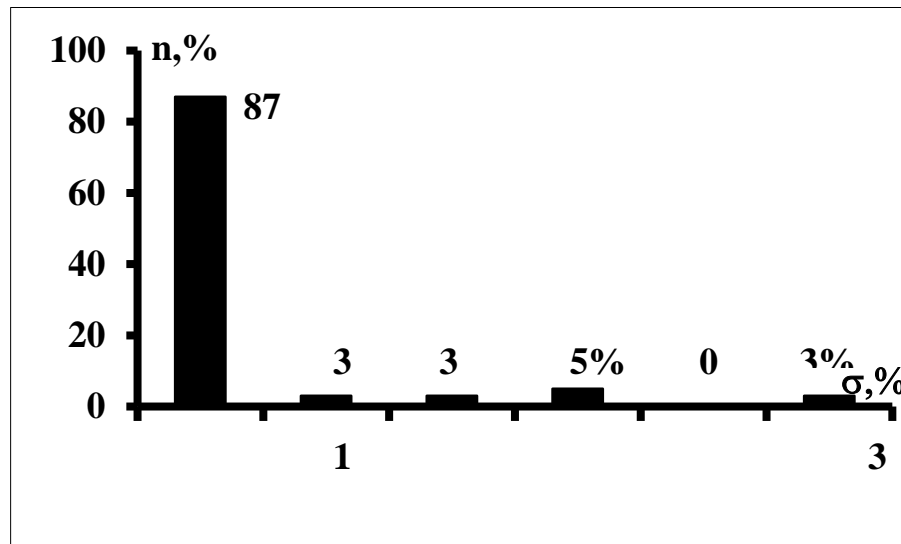


Figure 3. Example of the root-mean-square error distribution

Numerical criteria can be the maximum of the root-mean-square error  $\sigma_m$  (in this case  $\sigma_m = 2,6\%$ ), and the number of examples  $N$  (in absolute units or in percentages) that have an error of 5% (0.05) or 1% (0.01) (in this case  $N_5 = 100\%$  and  $N_1 = 90\%$ ). Thus, the quality of the ANN model and the percentage of connections recognized in the training of the network lends itself to a traditional statistical description with a confidence interval and a significance level. In order to assess the presence of latent links, one can change the number of input parameters of the network (traits) and, after evaluating the quality of the model, draw conclusions about those input parameters (psychological attributes) that are most important. The above example of analysis and numerical criteria for evaluating ANN models allow, by changing the structure of ANN, input characteristics and output function, to carry out a vertical system analysis of level psychological characteristics.

## Results and discussions

For the system analysis of gender differences, the ANN apparatus can be used after the separation of the respondents based on gender. For example, Figure 2 shows the results of ANN training for boys separately (corresponds to the model of figure 1). In the case when the values of all 12 Cattell's personality

traits (girls with boys) were submitted to the neural network input, the quality of the ANN model, according to the statistical criteria, was much worse ( $\sigma_m = 7,8\%$ ,  $N_5 = 97\%$  and  $N_1 = 80\%$ ). This allows us to draw a preliminary conclusion about more structured connections of personality traits with IQ for the boys than for the girls (Slavutskaya & Slavutskiy, 2012). It turns out that in order to obtain a qualitative ANN model for boys, it is sufficient to use at the input of a neural network the only 3 of the 12 Cattell's personality traits. For the girls the corresponding result can not be achieved. For them, with a decrease in the number of input attributes (in any combination), the quality of ANN training deteriorates. The percentage of recognizable connections is lower than in the general sample ( $\sigma_m = 13,6\%$ ). This confirms the earlier conclusion about more similar links of psychological signs for boys at this age.

The ANN model to evaluate the inderected (latent) connections of level psychological characteristics was constructed. To the input of ANN the values of biologically determined characteristics of the baseline level (sex, IQ, the results of the T test, characterizing the type of the nervous system) was given, and the output of the neural network consisted of personal characteristics of the highest level (the NA, the system of relations: in the family-R1, with peers - R2, in school - R3, to yourself-R4) (see Figure 4).

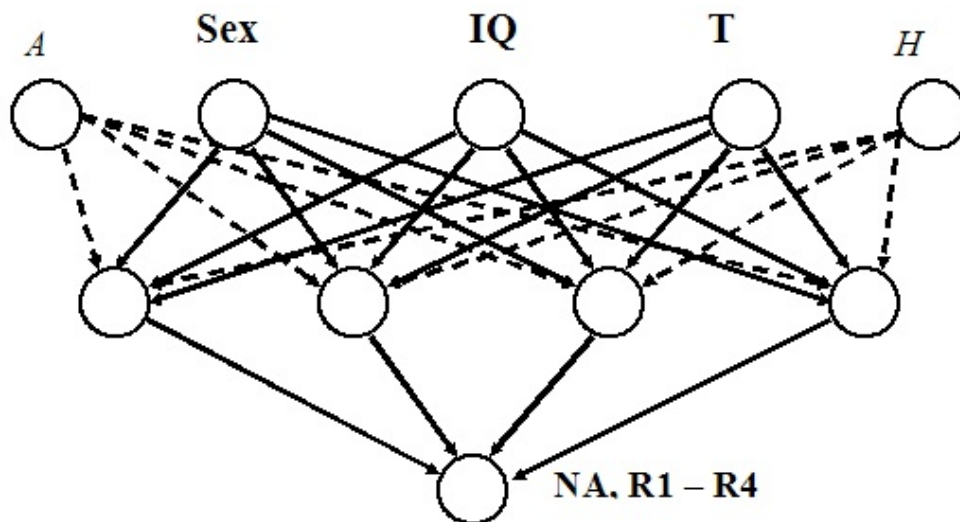


Figure 4. ANN for the latent connections analysis

Based on the results of a consistent analysis of the links between the psychological characteristics of the lower (base level) and upper personal level, it is established that the quality of the ANN model in Figure 4 is significantly dependent on the objective function - the highest level personal characteristic. For

R4 and R2 network training allows us to talk about the existence of a stable connection between the input and output characteristics ( $\sigma_m=0,77\%$ ,  $N_1=100\%$  and  $\sigma_m=6,2\%$ ,  $N_5=95\%$ ,  $N_1=80\%$ , respectively). For the NA, R1, R3, the quality of network training is very low. For NA  $\sigma_m=12\%$ ,  $N_5=86\%$ ,  $N_1=0\%$ ; for R1  $\sigma_m=16\%$ ,  $N_5=72\%$ ,  $N_1=0\%$ ; for R3  $\sigma_m=20\%$ ,  $N_5=76\%$ ,  $N_1=0\%$ .

Thus, it can be assumed that there is a significant relationship between the characteristics of the base lower level - the intelligence (IQ) and the type of the nervous system (T) with such higher-level indicators as attitudes toward oneself (R4) and peers (R2). Motivation, attitude to the family and to the school depend on more factors.

To assess these factors, it is necessary to take into account the psychological characteristics of the middle level. It was found that such characteristics can be communicative features in a 12-factor Cattell's questionnaire (CPQ): A - gregariousness - isolation; H - boldness-shyness in communication (see Figure 4). If the values of these attributes are fed to the input of the ANN model together with the characteristics of the lower level, then the quality of the training of the neural network is significantly improved for the objective functions, the quality of the ANN models with which turned out to be low for the structure of the neural network with three input neurons. For NA  $\sigma_m=4,8\%$ ,  $N_5=100\%$ ,  $N_1=86\%$ ; for R1  $\sigma_m=16\%$ ,  $N_5=87\%$ ,  $N_1=0\%$ ; for R3  $\sigma_m=5,7\%$ ,  $N_5=81\%$ ,  $N_1=62\%$ . In addition, to obtain a qualitative ANN model, one of the communicative features A or H can be replaced at the input of the network of Figure 4 by the volitional characteristic G (observance of norms and rules of behavior, strong-willed indicator).

Thus, it can be considered that the relationship of psychosocial characteristics of the upper level of NA, R3 and R1 with the basic characteristics of the lower level is mediated by communicative and volitional personality traits A, H, G.

If the quality of the ANN model is sufficiently high by statistical criteria, then the dependence of the output characteristic of the neural network on the value of each of the input attributes can be constructed for each respondent (Slavutskaya & Slavutskiy, 2014). For example, such dependencies can be NA(H), NA(G), NA(A).

In general, the statistical evaluation of the quality of neural network models allows to solve the following system analysis tasks during the psychodiagnostic data processing:

1. To assess the existensy (in principle) of the links between psychodiagnostic data, to what extent these are stable and structured.



2. Selectively to identify the links between the psychological traits for each respondent. To determine in what degree these dependences are nonlinear.
3. If the values of psychological signs are submitted to the ANN input, and gender is on the output, then according to statistics, feminine and masculine characteristics can be established.
4. System analysis in psychology involves mandatory classification of psychological characteristics. The finding of mediated links between them and the grouping of psychological attributes by certain properties is in fact the task of data classifying. The evaluation of gender characteristics (the previous paragraph) can also be considered as the classification task.

### **Conclusions**

On the example of the psychodiagnostic data processing for the children of preteen age (Collins, 1984) it is shown that the evaluation of the quality of ANN models training, as problem of non-linear multi-parameter optimization, allows solving some tasks of vertical system analysis in psychology. The detection and evaluation of the latent nonlinear relationships between the psychological characteristics is a problem that is difficult to solve by traditional statistical methods. The proposed approach allows psychologists to solve this problem with a fairly limited sample of respondents.

### **Acknowledgements**

The authors express their deep gratitude to professor V.S. Abrukov for fruitful discussions.

### **References**

- Baxt, W.G. (1994). Complexity, chaos and human physiology: the justification for non-linear neural computational analysis. *Cancer Lett*, 77(2-3), 85-93.
- Cattell, R.B. (1990). *Advanced in Cattellian Personality Theory. Handbook of Personality. Theory and Research*. New York: The Guilford Press.
- Collins, W.A. (1984). *Development during middle childhood: The years from six to twelve*. Washington, DC: Natl. Acad. Press.
- Erikson, E.H. (1950). *Childhood and Society*. New York: Norton.
- Glass, J. V. & Stanly, Y.C. (1970). *Statistical methods in education and psychology*. New Jersey.
- Grossberg, S. (1974). Classical and instrumental learning by neural networks. *Progress in theoretical biology, Vol. 3*. (pp. 51-141). New York: Academic Press.
- Harter, S. (1990). *Self and identity development. At the threshold: The developing adolescent*. Cambridge, MA: Harvard Univ. Press.

*Slavutskaya & Slavutskii, 2018. Preteen Age: the Analysis of the Multilevel Psycho-Diagnostic Data Based on Neural Network Models*

- Haykin, S. (1999). *Neural networks: A comprehensive Foundation*. New York: Prentice Hall.
- Hebb, D. (1961). *Organization of behavior*. New York: Science Edition.
- Kohlberg L. (1984). *The psychology of moral development: The nature and validity of moral stages*. B. Y.: Holt, Rinehart and Winston.
- Piaget J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development*, 15, 1-12.
- Reznichenko, N.S., Shilov, S.N. & Abdulkin, V.V. (2013). Neuron Network Approach to the Solution of the Medical-Psychological Problems and in Diagnosis Process of Persons with Disabilities (Literature Review). *Journal of Siberian Federal University. Humanities and Social Sciences*, 9 (6), 1256-1264.
- Slavutskaya, E.V. & Slavutskiy L. A. (2012) Using artificial neural networks for analysis of gender differences in younger teenagers. *Psikhologicheskie Issledovaniya*, 5(23), 4. <http://psystudy.ru> (in Russian, abstr. in English).
- Slavutskaya, E.V. & Slavutskiy L. A. (2014) Neural network analysis of the relationship between verbal and nonverbal intelligence in younger adolescents *Psychological Journal*, 35(5) 28-36.
- Usher M. & Zakay D. (1993). A neural network model for attribute-based decision processes *Cognitive Science*, 17, 349-396.