Section IV
Секція COMPUTER DIAGNOSTICS КОМП'ЮТЕРНА ДІАГНОСТИКА

UDC 519.7

Oleksandr BESARAB, PhD, Assoc. Professor,
Oleksandr GALKIN, DSc, Professor
Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine, e-mail: besarab@lll.kpi.ua, a.galkin@lll.kpi.ua

APPLICATION OF FUZZY LOGIC APPARATUS IN DISEASE DIAGNOSIS

Abstract. Systems based on artificial intelligence are penetrating deeper into all areas of human range, including medicine. The use of intelligent disease diagnosis systems allows to accumulate experience and knowledge of experienced doctors. This work is devoted to the development of an intelligent diagnostic system.

Keywords: fuzzy logic, expert systems, diagnosis, classifier, knowledge-based systems.

The rapid development of information technology has led to the use of artificial intelligence systems in virtually all areas of human activity, including the health care system. Knowledge-based systems or expert systems are one area of artificial intelligence. Fuzzy inference systems are knowledge-based systems. Fuzzy logic allows you to include inaccurate and uncertain information as part of the system. This ability to work with uncertainties and overlaps between class boundaries makes the model more consistent and mathematically easier to work with. The process of diagnosing diseases according to the patient's condition can be carried out by classifying, classifying according to the state of the disease. In [1] the author proposes to use the model of fuzzy inference as a universal classifier (Fig. 1).

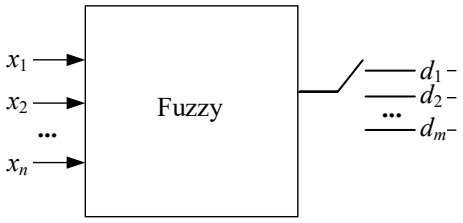

Fig. 1. Classifier based on fuzzy logic

Classification is carried out on the basis of fuzzy inference according to the production base of rules such as: If "variable_1 = value_A" and "variable_2 = value_B" and… and "variable_n = value_X" Then "variable_y = value_D".

Each such rule affects the final result, taking into account the implication of activated subconditions and aggregation. In this case, the author proposes not to carry out defuzzification. And as a result of the choice to accept the class for which there will be the highest degree of belonging.

A similar model was used for our research. We solved a problem that required a classifier that would classify data into two classes - according to patient data, it indicated the presence of 1 or the absence of disease 0.

Obviously, in addition to the discrete values of 0 and 1, it is advisable to have information, the distance from the cluster boundary, so that it is possible to draw a conclusion about the degree of cluster affiliation. As an indicator of the degree of belonging we will use the difference between the result of fuzzy inference from the limit values.

To develop and configure the model, we will use the Fuzzy Logic Toolbox package of the Matlab computing system. It has a wide range of tools for designing and researching fuzzy inference systems. The development of our model will be performed for the diagnosis of diabetes mellitus according to the data set presented in [2]. The data set used contains information on 768 female patients (268 diabetics and 500 non-diabetics). We will use rows only with cases where data is available for all columns. By selecting the rows with omissions, we obtained a table with 392 rows, including 130 cases with confirmed diabetes.

For the input linguistic variables, we will use 3 triangular linguistic variables, as they have a simple description and allow easy further adjustment of the fuzzy inference model.

The parameters of triangular functions are defined as follows. For the extreme membership functions, the vertex coordinate is assumed to be equal to the corresponding boundary of the universe of the linguistic variable. For other membership functions, the coordinates of the vertices are determined from the application of clustering of statistical data of the universe.

The universe of the linguistic variable is assumed to be equal to the minimum interval that covers the value of the column of the same data set

The next stage in the development of our model was the formation of the initial base of fuzzy rules. In our work, we applied the approach to the formation of the initial base of rules based on the fact that each example from the training sample is assigned a separate rule.

For each set of input and output variables were performed:

– the degree of belonging of the set values of variables to the terms of the corresponding linguistic variables is determined;

– the terms of linguistic variables were chosen, the degree of belonging to which turned out to be the greatest.

A database of rules has been formed for the selected terms. It is obvious that it turned out to be redundant and was reduced on a competitive basis.

The obtained classifier was further adjusted by automatically adjusting the weights and coefficients of the membership functions of the input and output variables.

A series of computational experiments was performed to further verify the developed model. The available sample of 392 data was randomly divided into two samples using a random number generator: training (300 cases) and test (92 cases). Such samples were formed several times. As a result of numerical experiments, it was found that the diagnosis error does not exceed 25%.

The next step in the study was to assess cases of misdiagnosis. The analysis of the obtained data of computational experiments showed that the erroneous results correspond to the range of results (0.4...0.6). Thus, the interpretation of the result of the developed model should be as follows: as a result (0…0.4) the disease will be pushed back; as a result [0.4…0.6] requires additional research; as a result (0.6…1.0) there is a disease.

A system of diagnosing using a classification based on fuzzy logic has been developed. It was established that the erroneous results of diagnostics according to the developed system belong to the border zone and thus the zone of reliability of the system operation results was revealed. Developed on the basis of fuzzy logic diagnostic system can be used for rapid diagnosis of diseases.

References

1. Shtovba S. D. Design of fuzzy classifiers in the system MATLAB / S. D. Shtovba, O. D. Pankevich // Works of the Scientific Conference «Design of Scientific and Engineering Applications in the Environment of MATLAB», 25-26 May 2004. P. 1320□1336.

2. Wu H, Yang S, Huang Z, He J, Wang X. Type 2 diabetes mellitus prediction model based on data mining. Informatics in Medicine Unlocked [Internet]. Elsevier BV; 2018;10:100–7. Available from: http://dx.doi.org/10.1016/j.imu.2017.12.006