

ALGORITHM FOR SIGNS SET OPTIMIZATION TO SOLVE CLASSIFICATION PROBLEMS

Akinin P.V., Karp V.P.

Moscow State Institute of Radio-Engineering, Electronics and Automation,
Faculty of Cybernetics, Information Systems Dept.,
78 Vernadskii prospect, Moscow, 119454, Russia
+7(909)923-22-33, E-mail: pavak@mail.ru

One of the most important problems solving by intelligent information systems is the classification problem. Generally, a researcher considers obviously superfluous set of signs because of the absence of sufficient a priori information. This set of signs bears noisy information about classes and generates unreliable dependence. In such cases, there is a need to optimize the signs set [1].

The idea of the algorithm is based on two iterative procedures: signs collection signs exception. Each procedure consists of several stages. Signs collection procedure begins with sorting out all signs through the estimation of their diagnostics quality in decreasing order (Q). The sign, which has a maximum value of Q , becomes the first sign in the set. Then for the sign given another one, which demonstrates together with the first one the best Q estimation, is added. Further to these two signs the third one is added to achieve the best Q of the signs set. The process continues until among the not included signs there will not remain the single one, whose inclusion in the set would improve the evaluation of Q . After this the reverse procedure starts: the algorithm excludes from the given signs set the signs that do not worsen the Q estimation of the formed set. Upon completion, another attempt to find out the sign that would improve Q is taken place. After the completion of the procedure mentioned signs elimination process begins. Thus, an iterative system is being formed, which via an odd iteration strives to include signs that improve the assessment of the quality of diagnosis, while an even iteration – strives to exclude signs, that do not worsen it. This results in a set of signs, which gives the best diagnostic results for the chosen method of classification. It is important to note that the engine optimization described is universal as it does not imply the use of any specific classification algorithm. Testing of the algorithm is planned, primarily on the formation of a diagnostic signs set by sorting conjunctions.

References

1. *Karp V.P.* Metodi i sredstva kontrolya i diagnostiki slojnoorganizovannih sistem. – M.: MIREA, 2008. 196 p.