

ABSTRACT

The thesis is presented in 57 pages. It contains 2 appendixes and bibliography of 21 references. 3 figures and 9 tables are given in the thesis.

The goal of the thesis is to develop mathematical and software tools for solving the problem of Preprocessing of Numerical Data for Use in Logico-Linguistic.

In the thesis, existing solutions are analyzed, such as methods for pre-processing data for data mining, in particular – discretization of continuous variables. They are compared in terms of the possibility of using existing solutions in a stand-alone automatic and user's possibility of change parameters of discretization system. Some modification of basics discretization algorithms is chosen. to solve the task.

Discretization algorithm is implemented according to the general requirements and task. The automated system implementing the chosen method is developed. The developed system is tested.

Keywords: data mining, machine learning, discretization, data analysis, data preprocessing.

