

## ABSTRACT

Thesis: 85 p., 19 fig., 18 tabl., 2 append., 9 sources.

MATHEMATICAL MODEL, QUALITY CRITERIA, CLUSTERING METHODS, CATEGORY DATA, CLOPE

Object of research: modern methods of machine learning.

Purpose: use of various clustering methods for categorical data, comparing them, identifying strengths and weaknesses.

Subject of research: application of methods of machine learning for solving the problem of data clustering

The diploma project is devoted to the implementation of methods of clustering categorical features, including clustering of transactions. This work is especially relevant because large database clustering can not only structure data, but also find abnormalities in it, and build simplified predictions. The work analyzes the work of existing algorithms and methods of clusterization, various metrics and criteria of the quality of the work adequacy of the models, criteria of the quality of forecasts estimates, which are used in forecasting. A decision support system has been created.