ABSTRACT

Master's thesis: 4 p., 0 fig., 0 tabl., 1 appendixes, 24 sources.

The topic of the research "Composition methods of classification and regression"

The paper provides overview of basic machine learning theory and its applications in text mining.

In this paper the author considered following methods of text analysis: semantic analysis, probabilistic latent semantic analysis, latent Dirichlet allocation. Main features, advantages and disadvantages are researched.

Described composition methods are used to problems which based on real data. Analysis of results and comparison of used methods is conducted.

These results show the validity of these methods and provide the opportunity for their future use for solving text mining problems.

MACHINE LEARNING, TEXT MINING, SEMANTIC ANALYSIS, TOPIC MODEL, ADDITIVE REGULARIZATION.