

Bachelor's thesis: 87 p., 19 fig., 7 tabl., 2 appendixes, 19 references.

Object of study – machine learning algorithms for analysis textual information.

Subject of study – using of methods for analysis textual information in the problem of text classification.

Purpose of study – to compare some machine learning methods for the analysis of textual information in the example of the problem of spam classification.

Relevance of study – in our time Data Mining is one of the most promising research and applied areas, which rapidly growing over the past decades and it has practical applications in many important spheres of human's social life. Text Mining is crucial part of this field and it's an important aspect of modern business.

Results – during the process of comparing the different methods for analysis textual information, were allocated advantages and disadvantages of each of them, and was analyzed the feasibility of their using in solving specific problems of machine learning.

Further improvements – expanding the range of tasks, which can be appropriate for using these methods. Expanding the list of methods for analysis textual information and analysis of their effectiveness.

DATA MINING, TEXT MINING, MACHINE LEARNING, TEXT CLASSIFICATION, CLASSIFICATION PROBLEM, SPAM CLASSIFICATION