

## ABSTRACT

The thesis is presented in 50 pages. It contains 2 appendices and bibliography of 20 references. Eight images and 1 table are given in the thesis.

Diploma work is dedicated to mathematical and programming software development for system of data from Internet network collection and ranging.

In the thesis existing solutions for data collection from Internet network and methods of their clustering were analyzed. Their advantages and disadvantages were distinguished.

Based on Scrapy framework system for information from news web-sites search and storage was created. Collected information was analyzed with the help of Python programming language realization of Fuzzy clustering method which allowed to filter and group news according to their general characteristics.

Key words: data collection system, fuzzy clustering, C-means clustering, Scrapy, Python.

