

Master's dissertation consists 64 pages, 17 images, 33 tables, 38 referring sources.

Topicality: Today, the importance of textual data processing is increasing. This is due to the large amount of textual information available through the Internet. Because millions of content characters are generated every day, people do not have the physical ability to process all information.

The aim of the study: Improvement of real-time analysis of Ukrainian-language streaming text data and anomaly detection

To achieve this goal, the following tasks were formulated:

- justify the choice of anomaly detection method;
- to create a mathematical model of the chosen method of anomaly detection;
- to perform software implementation of the method of anomaly detection;
- to investigate the effectiveness of the anomaly detection method.

Object of study: streams of Ukrainian-language text data.

Subject of research: anomalies detection in streaming text data.

Research methods: text mining methods, data mining methods.

Scientific novelty: The most significant scientific results of the master's thesis are:

- development of an adapted Isolation Forest method for detecting anomalies in Ukrainian-language text data streams.

The practical significance of the obtained results is determined by the fact that a modified Isolation Forest algorithm is proposed, which supports the detection of anomalies in Ukrainian-language data.

Relationship with working with scientific programs, plans, topics: work was performed at the Department of Automated Information Processing and Management Systems of the National Technical University of Ukraine «Kyiv Polytechnic Institute. Igor Sikorsky» within the topic «Methods and technologies of high-performance computing and processing of large data sets». State Registration Number 0117U000924.

Testing: The main points of the work were reported and discussed at the Third All-Ukrainian Scientific and Practical Conference of Young Scientists and Students "Information Systems and Management Technologies" (ISTU-2019) - Kyiv: NTUU "KPI them. Igor Sikorsky", November 20-22, 2019

Publications: Scientific provisions of the dissertation published in Afanasieva O.E. Detection of anomalies in text data streams / O.E. Afanasieva, Y.O. Oliynyk // Proceedings of the Third All-Ukrainian Scientific and Practical Conference of Young Scientists and Students "Information Systems and Management Technologies" (ISTU-2019) - Kyiv: NTUU "KPI them. Igor Sikorsky", November 20-22, 2019

Keywords: DATA FLOWS, ANOMALY DETECTION, ISOLATION FOREST METHOD, UKRAINOMATIC DATA, TEXT DATA, DATA MINING.