ABSTRACT

The explanatory note is 83 pages long and contains 6 illustrations, 18 tables, 4 appendices, and 25 references.

The relevance of the topic is due to the rapid development of e-commerce and the constant growth of product information presented on numerous Internet resources in an unstructured and diverse format. Manual collection, updating, and analysis of data on product characteristics, prices, availability, and reviews is labor-intensive and inefficient, and does not ensure the timely receipt of up-to-date information. Traditional data processing methods do not meet the requirements of speed, accuracy, and scalability in the context of large amounts of information. In this regard, there is a need to create intelligent tools capable of automatically collecting, structuring, and analyzing product information, taking into account the semantic relationships between data. The use of artificial intelligence methods can increase the efficiency of these processes, making the development of automated systems for collecting product information a relevant scientific and practical task.

Research objective. The objective of this work is to increase the speed and accuracy of data collection and analysis from Internet resources, as well as to expand the possibilities of searching for the necessary information by improving search algorithms, including the use of semantic approaches.

Research object: software for automated collection, processing, and analysis of information from Internet resources.

Research subject: methods, models, and algorithms for collecting, searching, and classifying data, as well as ways to integrate artificial intelligence tools into web applications for automated information processing.

Tasks solved in the work:
• analysis of existing tools for collecting and processing information;
• research of modern search methods, including semantic search;
• development of a methodology for integrating artificial intelligence algorithms into the process of data collection and analysis;

- creation of a web application architecture focused on efficient information processing and support for intelligent search;

- evaluation of the effectiveness of the proposed solution through comparative analysis with traditional methods;

The scientific novelty of the master's thesis results lies in the development of a software solution that combines automated collection of information from Internet resources and a system of personalized recommendations. The result was achieved by using intelligent product search based on TF-IDF vector representation of texts instead of simple search string matching and subsequent structured grouping of selected results using a modified K-means model.

Practical significance:

The developed web application can be used in the fields of marketing, cybersecurity, journalism, scientific research, and public administration for automated monitoring of information resources. Improved search algorithms provide higher quality relevant results, reduce the time spent on analysis, and increase the effectiveness of decision support.