ABSTRACT

The size of the explanatory note is 99 pages, containing 6 illustrations, 16 tables, 3 appendices, and 30 references.

Topicality. The modern scientific and technical sphere is characterized by an exponential growth in information volumes, which leads to the problem of "information overload" for specialists and complicates the timely analysis of publications. Existing methods of automatic summarization have significant drawbacks: extractive algorithms (TextRank, LexRank) provide factual accuracy but low text coherence, while modern abstractive models (based on the Transformer architecture) have limitations on input text length and are prone to "hallucinations." Furthermore, most solutions ignore tabular and graphical data, which are critically important for scientific documents. This necessitates the development of hybrid methods adapted for full-text scientific articles, taking into account multimodal data.

The aim of the study is to improve the quality and accuracy of automatic summarization of scientific and technical texts by developing a software solution that implements a hybrid method combining extractive and abstractive summarization approaches, as well as the processing of tables, figures, and graphs.

Object of research: the process of automatic summarization of scientific and technical texts.

Subject of research: hybrid models, methods, and algorithms of automatic summarization that combine extractive and abstractive (Transformer-based) approaches for document processing.

To achieve the set goal, it is necessary to solve the **following tasks:**

— To analyze the current state of the subject area, investigate existing methods (extractive, abstractive, hybrid) and software tools for summarization, identifying their limitations when working with scientific texts.

— To theoretically substantiate and develop an improved hybrid summarization method that includes structural text segmentation, work with tables and graphical elements, extractive selection of key sentences, and abstractive generation of the final summary.
— To develop a software solution that implements the proposed hybrid method using modern natural language processing libraries.
— To conduct an experimental evaluation of the effectiveness of the developed method, and compare its quality with baseline approaches using standard metrics (ROUGE).

The scientific novelty of the work lies in the fact that a software solution has been developed for automatic summarization of scientific and technical texts, which combines extractive and abstractive approaches with preliminary text partitioning into sections, and also takes into account the structured data of tables and graphical elements of the document. The result was achieved by improving the hybrid summarization method through the combination of structural segmentation, LexRank graph ranking, BART abstractive generation, and Vision-Language models.

The practical value lies in the creation of a ready-made software solution that will allow researchers, engineers, and students to significantly reduce the time for literature review and rapid analysis of large arrays of scientific and technical documents.

Relationship with working with scientific programs, plans, topics. The work was performed at the Department of Computer Science and Software Engineering of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

Approbation. The scientific provisions of the thesis underwent approbation at the IX International Scientific and Practical Conference of Young Scientists and Students "Software Engineering and Advanced Information Technologies" (SoftTech-2025) – Kyiv.

Publications. The scientific provisions of the thesis are published in:

Bublikov K.I. Preliminary sectioning and its impact on reducing lead bias in hybrid extractive-abstractive summarization / Bublikov K.I., Polupan Yu.V. // Proceedings of the IX International Scientific and Practical Conference of Young Scientists and Students "Software Engineering and Advanced Information Technologies" (SoftTech-2025) – Kyiv: NTUU "Igor Sikorsky KPI", November 26-28, 2025.

Keywords: AUTOMATIC SUMMARIZATION; LEAD BIAS; TEXT SECTIONING; HYBRID MODELS; EXTRACTIVE SUMMARIZATION; ABSTRACTIVE SUMMARIZATION; COHERENCE; STRUCTURAL BALANCE.