

SUMMARIES: STRUCTURE, TOPIC, MAIN POINTS

Jabborova Dildora Zokirjanovna

Department of English language theoretical aspects senior teacher PhD
Uzbekistan State World Languages University
dilyacity89@gmail.com

Abdulazizova Zarinabonu

Uzbekistan State World Languages university First English faculty
Foreign language and literature
zarinaabdulazizova834@gmail.com

Annotatsiya. Ushbu tezisdagi ekstraktiv referat tizimlari va ularning operatsion tamoyillari o'rganiladi. Mavzu ushbu tizimlarning bir yoki bir nechta hujjatlardan muhim jummalarni aniqlab, ularni birlashtirib ixcham va ravon xulosa yaratishiga qaratilgan. Tezis tuzilishi uch qismdan iborat: asosiy vazifa va mazmun tanlash, ekstraktiv usullarni abstraktiv usullardan afzal ko'rish asoslari hamda uchta mustaqil operatsion bosqich. Asosiy fikrlar shundan iboratki, ekstraktiv yondashuvlar foydalanuvchilarning axborot ehtiyojlariga moslasha oladi, bir va bir nechta hujjatlar bilan samarali ishlaydi hamda uchta vazifani o'z ichiga oladi: oraliq vakillik yaratish, shu vakillik asosida jummalarni baholash va bir necha jumladan iborat xulosani tanlash. Kirismuhim dizayn qarorlarini ta'kidlaydi va operatsion bosqichlarni tahlil qilish ayrim usullarning boshqalarga nisbatan afzalliklarini qanday ochib berishini tushuntiradi.

Kalit so'zlar: ekstraktiv referatlashtirish, jummalarni baholash, oraliq vakillik, bir hujjatli referatlashtirish, ko'p hujjatli referatlashtirish, matndan matnga generatsiya, abstraktiv referatlashtirish, axborot ehtiyoji

Аннотация. В данном тезисе рассматриваются экстрактивные системы реферирования и их операционные принципы. Тема посвящена тому, как эти системы создают краткие и связные резюме путём выделения и объединения значимых предложений из одного или нескольких документов. Работа структурирована в три части: основная функция и отбор содержания, обоснование выбора экстрактивных методов вместо абстрактных, а также три независимых операционных этапа. Основные положения включают следующее: экстрактивные подходы адаптируются к информационным потребностям пользователей, эффективно работают с одним или несколькими документами и включают три задачи — создание промежуточного представления, оценку предложений на основе этого представления и отбор резюме, состоящего из нескольких предложений. Введение подчёркивает ключевые проектные решения и объясняет, как анализ операционных этапов выявляет преимущества одних методов над другими.

Ключевые слова: экстрактивное реферирование, оценка предложений, промежуточное представление, реферирование одного документа, реферирование нескольких документов, генерация текста в текст, абстрактное реферирование, информационная потребность.

Annotation. This thesis examines extractive summarization systems and their operational principles. The topic focuses on how these systems generate concise, fluent summaries by identifying and concatenating salient sentences from single or multiple documents. The work is structured into three parts: core function and content selection, justification of extractive over abstractive methods, and three independent operational tasks.

The main points include that extractive approaches are adaptable to users' information needs, function effectively for both single- and multi-document inputs, and involve three tasks: creating an intermediate representation, scoring sentences based on that representation, and selecting a summary consisting of several sentences. The introduction highlights critical design choices and explains how analyzing operational stages reveals advantages of certain techniques over others.

Key words: extractive summarization, sentence scoring, intermediate representation, single-document summarization, multi-document summarization, text-to-text generation, abstractive summarization, information need.

Introduction

Extractive summarization systems constitute a fundamental approach within automatic text summarization, as their primary objective is to generate concise and fluent summaries that preserve the most salient information from the original input. The central topic of this chapter is the examination of extractive summarization systems designed specifically for producing short, paragraph-length summaries, and the elucidation of their internal mechanisms. In terms of structure, the discussion proceeds through three interconnected stages: first, the core function and content selection principles of extractive summarizers are delineated; second, the methodological justification for prioritizing extractive over abstractive techniques is critically assessed; and third, three independent tasks that underpin the operational pipeline of virtually all such systems are systematically distinguished.

With regard to the main points, the chapter advances several key arguments. First, extractive summarizers operate by identifying the most semantically significant sentences within a given input—whether the input comprises a single document or a cluster of topically related documents—and subsequently concatenating these sentences to produce a coherent summary. Importantly, the determination of content importance is driven primarily by the intrinsic features of the input text itself, rather than by external knowledge bases or user-specific queries. Second, although the deliberate focus on extractive techniques necessarily excludes the extensive body of text-to-text generation research developed for abstractive summarization, this methodological constraint affords several advantages. Specifically, extractive methods represent some of the most dominant and broadly adaptable approaches in the field; they can be readily modified to accommodate users' information needs and function effectively across both single-document and multi-document input scenarios. Third, a systematic examination of the sequential operational stages within extractive summarizers enables researchers to identify both commonalities and divergences among various summarization approaches. These differences often pertain to critical system components—such as sentence scoring mechanisms or representation schemes—and may explain the relative advantages of certain techniques over others. Finally, to facilitate a deeper understanding of system operation and to highlight the essential design choices that developers must make, the chapter distinguishes three relatively independent tasks performed by virtually all extractive summarizers: constructing an intermediate representation of the input that captures only its key aspects; computing a salience score for each sentence based on that representation; and selecting a subset of top-scoring sentences to form the final summary.

Conclusion

In conclusion, extractive summarization systems offer a principled and adaptable approach to generating concise, fluent summaries by selecting and concatenating salient

sentences from input texts. As discussed, the decision to focus on extractive techniques, while excluding abstractive text-to-text generation methods, is methodologically justified by the flexibility of these systems to accommodate users' information needs and their applicability to both single- and multi-document inputs. Furthermore, the examination of three independent operational tasks—namely, creating an intermediate representation that captures key aspects of the text, scoring sentences based on that representation, and selecting a summary consisting of several sentences—provides a coherent framework for understanding system design and performance. By distinguishing these tasks, researchers and developers can identify critical design choices, compare commonalities and differences across approaches, and explain the relative advantages of certain techniques over others. Thus, the analysis of extractive summarization not only clarifies how these systems operate but also lays the groundwork for future improvements in automatic text summarization.

References

- Siddiqui, M. K., Ahmad, A., Pal, O., & Ahmad, T. (2021). CoRank: A clustering cum graph ranking approach for extractive summarization.
- da Silva, V. C. (2022). Extractive text summarization using generalized additive models with interactions for sentence selection.
- Yoon, S., Chan, H. P., & Han, J. (2023). PDSum: Prototype-driven continuous summarization of evolving multi-document sets stream. In *ACM Web Conference 2023 (WWW '23)* (pp. 1650–1661). Association for Computing Machinery.
- Zhang, X., Wei, Q., Song, Q., & Zhang, P. (2024). TOMDS (Topic-Oriented Multi-Document Summarization): Enabling personalized customization of multi-document summaries. *Applied Sciences*, 14(5), 1880.
- Jin, H., et al. (2024). A comprehensive survey on process-oriented automatic text summarization with exploration of LLM-based methods.
- IEEE Access. (2025). Current trends and advances in extractive text summarization: A comprehensive review. *IEEE Access*, 13, 28150–28166.