

## О МОДЕЛИРОВАНИИ СМЫСЛА ТЕКСТОВ ПРИ РЕЦЕНЗИРОВАНИИ НАУЧНЫХ СТАТЕЙ

Бондарев Ю.А. (Университет ИТМО)

Научный руководитель – доктор технических наук, доцент Духанов А.В.  
(Университет ИТМО)

**Введение.** Анализ научных текстов, в частности их рецензирование – сложный процесс, требующий от рецензента глубокого понимания анализируемого текста. Автоматизация данного процесса в настоящее время проводится, однако направлена она преимущественно на синтаксис, грамматику и иные внешние составляющие текста. В некоторых случаях анализируется также связность между элементами и общее соответствие тематике исследования. С данными задачами хорошо справляются большие языковые модели [1]. Вместе с тем, необходимо также анализировать основной смысл текста, то есть главное содержание, либо же основную информацию, изложенную автором в виде конкретного текста. В связи с этим предполагается провести аналитический обзор основных методов извлечения смысла текста, существующих на сегодняшний день с целью выделения их особенностей работы, областей применения и основных недостатков.

**Основная часть.** Для анализа смысла, под которым обычно понимается основное информационное содержание, воплощенное в коммуникативной структуре текста, могут быть использованы различные методы. Чаще всего под анализом смысла текста понимается анализ семантических связей между элементами текста [2]. Подобный анализ проводится с помощью различных методов, в частности, методов векторизации текста, метода опорных векторов (SVM), а также с помощью различных языковых моделей. Однако, данные методы неприменимы для анализа смысла текста, так как основаны на работе с контекстом, что делает их схожими с механизмами внимания в больших языковых моделях. Данные механизмы могут быть полезны при работе со смыслами, однако по своей сути не предназначены для работы со смыслами.

Наиболее часто в последнее время для анализа смысла текста также используются графовые методы. Построение графов текстов, а также использование графовых нейронных сетей позволяет более тщательно анализировать смысл текста. Именно такие методы позволяют более тщательно проанализировать связи между понятиями и терминами и определить, насколько те или иные части текста влияют на общий смысл [3], и в результате дают относительно полное смысловое содержание. Данные методы применяются, например, при автоматическом реферировании текстов. Недостатком графовых методов является неспособность (на данный момент времени) анализировать сложные тексты, а также общая точность анализа.

**Выводы.** В ходе работы были проанализированы основные используемые в настоящее время методы и математические модели представления смысла текста, особенно применительно к научному тексту. Среди выделенных недостатков отмечается точность работы и неспособность обрабатывать сложные тексты, например, научные статьи или даже рецензии. На основании полученных результатов предполагается разработка методов автоматизации рецензирования, а также валидации рецензий, в том числе написанных с помощью искусственного интеллекта

### Список использованных источников:

1. Zhou R., Chen L., Yu K. Is LLM a Reliable Reviewer? A Comprehensive Evaluation of LLM on Automatic Paper Reviewing Tasks // Proceedings of the 2024 Joint International Conference

on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), Torino, Italia. ELRA and ICCL. 2024. P. 9340–9351.

2. Christianson N.H., Sizemore Blevins A.S., Bassett D.S. Architecture and evolution of semantic networks in mathematics texts. // Proceedings: Mathematical, Physical and Engineering Sciences. 2020, vol. 476, №2239.

3. Kouris, P., Alexandridis, G., Stafylopatis, A. Text summarization based on semantic graphs: an abstract meaning representation graph-to-text deep learning approach. // Journal of Big Data, 2024, vol. 11, №95.