

## РЕФЕРАТ

Розмір пояснювальної записки – 99 аркушів, містить 6 ілюстрацій, 16 таблиць, 3 додатки, 30 посилань на джерела.

**Актуальність теми.** Сучасна науково-технічна сфера характеризується експоненціальним зростанням обсягів інформації, що призводить до проблеми «інформаційного перевантаження» фахівців та ускладнює своєчасний аналіз публікацій. Існуючі методи автоматичного реферування мають суттєві недоліки: екстрактивні алгоритми (TextRank, LexRank) забезпечують фактологічну точність, але низьку зв'язність тексту, тоді як сучасні абстрактні моделі (на базі архітектури Transformer) мають обмеження на довжину вхідного тексту та схильні до «галюцинацій». Крім того, більшість рішень ігнорує табличні та графічні дані, які є критично важливими для наукових документів. Це зумовлює необхідність розробки гібридних методів, адаптованих для повнотекстових наукових статей з урахуванням мультимодальних даних.

**Мета дослідження** полягає у підвищенні якості та точності автоматичного реферування науково-технічних текстів шляхом розробки програмного рішення, що реалізує гібридний метод, який поєднує екстрактивні та абстрактивні підходи реферування, обробку таблиць, рисунків та графіків.

**Об'єкт дослідження:** процес автоматичного реферування науково-технічних текстів.

**Предмет дослідження:** гібридні моделі, методи та алгоритми автоматичного реферування, що поєднують екстрактивні та абстрактивні (Transformer-based) підходи для обробки документів.

Для досягнення поставленої мети необхідно вирішити **наступні задачі:**

— Проаналізувати сучасний стан предметної області, дослідити існуючі методи (екстрактивні, абстрактивні, гібридні) та програмні засоби реферування, виявивши їхні обмеження при роботі з науковими текстами.

— Теоретично обґрунтувати та розробити вдосконалений гібридний метод реферування, що включає структурну сегментацію тексту, роботу з таблицями і графічними елементами, екстрактивний відбір ключових речень та абстрактивну генерацію фінального реферату.

— Розробити програмне рішення, що реалізує запропонований гібридний метод, з використанням сучасних бібліотек обробки природної мови.

— Провести експериментальну оцінку ефективності розробленого методу, порівняти його якість з базовими підходами за стандартними метриками (ROUGE).

**Наукова новизна** роботи полягає в тому, що розроблено програмне рішення для автоматичного реферування науково-технічних текстів, яке поєднує екстрактивні й абстрактивні підходи з попереднім розділенням тексту на секції, а також враховує структуровані дані таблиць та графічні елементи документу. Результат досягнуто удосконаленням гібридного методу реферування шляхом поєднання структурної сегментації, графового ранжування LexRank, абстрактивної генерації BART та Vision-Language моделей.

**Практичне значення** полягає у створенні готового програмного рішення, що дозволить дослідникам, інженерам та студентам значно скоротити час на огляд літератури та швидкий аналіз великих масивів науково-технічних документів.

**Зв'язок з науковими програмами, планами, темами.** Робота виконувалась на кафедрі інформатики та програмної інженерії Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського".

**Апробація.** Наукові положення дисертації пройшли апробацію на IX Міжнародній науково-практичній конференції молодих вчених та студентів «Інженерія програмного забезпечення і передові інформаційні технології» (SoftTech-2025) – м. Київ.

**Публікації.** Наукові положення дисертації опубліковані в:

Бубліков К.І. Попереднє секціонування і його вплив на зменшення lead bias у гібридному екстрактивно-абстрактивному реферуванні / Бубліков К.І., Полупан Ю.В. // Матеріали 6 IX Міжнародної науково-практичної конференції молодих вчених та студентів «Інженерія програмного забезпечення і передові інформаційні технології» (SoftTech-2025) – м. Київ: НТУУ «КПІ ім. Ігоря Сікорського», 26-28 листопада 2025 р.

**Ключові слова:** АВТОМАТИЧНЕ РЕФЕРУВАННЯ; LEAD BIAS; СЕКЦІОНУВАННЯ ТЕКСТУ; ГІБРИДНІ МОДЕЛІ; ЕКСТРАКТИВНЕ РЕФЕРУВАННЯ; АБСТРАКТИВНЕ РЕФЕРУВАННЯ; КОГЕРЕНТНІСТЬ; СТРУКТУРНА ЗБАЛАНСОВАНІСТЬ.