

## РЕФЕРАТ

**Актуальність проблеми.** З початком ери автоматизації кількість нових розробок та досліджень зростає з експоненціальною швидкістю. Завдяки Інтернету більшість новітніх досліджень і відкриттів стають доступними для широких мас населення й активно використовуються. Це призвело до необхідності постійного самонавчання як викладачів, так і фахівців інших галузей.

Самонавчання пов'язане з проблемою контролю засвоєння вивченої інформації, тому постає питання в створенні автоматизованих систем, котрі допомогли б у цьому, наприклад систем автоматизованої генерації запитань. На сьогодні не існує систем, які здатні повноцінно розв'язувати таку задачу, адже тут необхідно не тільки виділити найважливішу інформацію з тексту, а й згенерувати запитання з виділеної інформації. Методики, що розроблювались для вирішення кожної з цих двох задач, досить ефективно працюють окремо, але в поєднанні не показують хороших результатів. Наприклад, існують окремо системи автоматизованого реферування, які виділяють важливу інформацію з тексту, системи генерації запитань до конкретних речень, проте їх поєднання не дає бажаного результату. Крім того, ці системи створені для мов із чітким порядком слів у реченні (англійська, німецька мови тощо), для текстів українською мовою згадані вище системи не розроблялися. Тому створення методик та програмних модулів для генерації відкритих запитань до текстів українською мовою є актуальною і важливою задачею з наукової, так і з практичної точки зору.

**Об'єкт дослідження** – процеси, моделі та методи автоматизованого аналізу текстової інформації.

**Предмет дослідження** – методи та методики автоматизованого реферування та пошуку ключової інформації в тексті, методи генерації запитань до текстів українською мовою.

**Мета роботи** – розробка та удосконалення методики пошуку ключової інформації в науково-технічних текстах та методів генерації відкритих запитань до науково-технічних текстів українською мовою.

**Методи дослідження:**

- аналіз, синтез, індукція, дедукція, абстрагування, конкретизація, класифікація, систематизація, схематизація;
- методи аналізу текстової інформації, методи кластеризації даних та методи автоматизованого реферування.

**Наукова новизна** роботи полягає в тому, що:

1. Розроблено методику пошуку ключових речень у науково-технічних текстах, яка відрізняється від наявних тим, що дозволяє враховувати тематичне спрямування і смислове навантаження речень аналізованого тексту.

2. Вперше запропоновано використовувати метод нейронних мереж для генерації запитань до речень, що дало змогу класифікувати речення текстів українською мовою (де немає чіткого синтаксичного порядку слів у реченні) і генерації запитань до них.

3. Розроблено принципово нові класи для класифікації речень, що дозволило трансформувати вхідні речення та отримувати в результаті запитання.

**Практична цінність** отриманих в роботі результатів полягає в тому, що запропоновані варіанти розв'язання проблем дають хороші результати, а методи та програмні модулі можуть використовуватись під час створення системи генерації запитань, яка спрощує перевірку рівня засвоєння інформації.

**Апробація роботи.** Основні положення і результати роботи були представлені на VII науковій конференції магістрантів та аспірантів «Прикладна математика та комп'ютеринг» ПМК-2015 (Київ, 15–16 квітня 2015 р.) та опубліковані у збірнику міжнародної науково-практичної конференції «Наука України: проблеми сьогодення та перспективи розвитку».

**Структура та обсяг роботи.** Магістерська дисертація складається зі вступу, чотирьох розділів, висновків та додатків.

У вступі викладень загальну характеристику роботи, подано оцінку

сучасного стану досліджуваної проблеми, обґрунтовано актуальність дослідження, сформульовано мету і завдання дослідження, наукову новизну отриманих результатів і практичну цінність роботи, наведено відомості про апробацію результатів і їх впровадження.

У першому розділі розглянуто основні методи автоматизованого реферування та методи кластеризації, подано результати їхнього порівняльного аналізу; на основі отриманих результатів обрано методи і методики, які будуть корисними для проведення дисертаційного дослідження.

У другому розділі проаналізовано принцип моделювання інформації і виявлення знань в текстах; удосконалено методику пошуку найбільш інформативних речень у тексті з урахуванням недоліків, які були виявлені в процесі аналізу методів.

У третьому розділі запропоновано власну класифікацію речень, котра буде застосовуватись для генерації запитань; висвітлено теоретичні аспекти методу, що використовується.

У четвертому розділі подано інформацію про розроблені програмні модулі; наведено приклади та результати тестування розроблених методик пошуку ключової інформації та методу генерації запитань.

У висновках викладено найбільш значущі наукові та практичні результати проведеного наукового дослідження та програмної реалізації, обґрунтовано достовірність отриманих результатів.

Загальний обсяг роботи становить 95 сторінок, основний зміст викладено на 59 сторінках. Робота містить 2 додатки, список використаних літературних джерел з 38 найменувань, 8 рисунків та 2 таблиці.

**Ключові слова:** метод автоматичного реферування, методи пошуку ключової інформації, метод класифікації, генерація запитань.