

## АНОТАЦІЯ

Дипломну роботу виконано на 57 аркушах, вона містить 1 додаток та перелік посилань на використані джерела з 16 найменувань. У роботі наведено 7 рисунків та 3 таблиці.

Метою даної дипломної роботи є створення математичного та програмного забезпечення для системи аналізу унікальності текстових документів та цифрових зображень та автореферування текстів.

У роботі проведено аналіз існуючих рішень указаних задач — системи «Антиплагіат», програми Advego Plagiatus, сервісу Unplag для виявлення плагиату в текстах; для автореферування текстів було розглянуто методи Extractive та Abstractive груп, такі як LexRank, LSA, Gensim's TextRank summarization module, та було обрано Gensim's TextRank. Виконано їх порівняння з погляду точності отримуваних розв'язків, ефективності алгоритмів та пристосованості методів до використання даних. Для аналізу унікальності текстів був обраний алгоритм шинглів, а для зображень - спосіб секторного порівняння гістограм.

Розроблено автоматизовану систему, що реалізує обрані методи. Виконано тестування розробленої системи.

Робота виконується на замовлення ТОВ "DataRoot".

Ключові слова: аналіз унікальності, плагіат, автореферування, обробка текстів, обробка зображень.