

АНОТАЦІЯ

Дипломну роботу виконано на 50 аркушах, вона містить 2 додатки та перелік посилань на використані джерела з 14 найменувань. У роботі наведено 5 рисунків та 3 таблиці.

Метою даної дипломної роботи є створення математичного та програмного забезпечення для визначення недоліків програмного забезпечення за допомогою кластеризації повідомлень технічної підтримки.

У роботі проведено аналіз існуючих рішень поставленої задачі – методів кластеризації – DBSCAN, k-середніх, ієрархічна кластеризація, expectation-maximization (EM). Виконано їх порівняння з погляду на складність алгоритмів, їх точність, ефективність та адекватність методів при застосуванні до текстових даних. Для розв'язання задачі в роботі обрано метод кластеризації - DBSCAN.

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

Ключові слова: аналіз текстів, кластеризація, визначення недоліків.