

## АНОТАЦІЯ

Дипломну роботу виконано на 45 аркушах, вона містить 2 додатки та перелік посилань на використані джерела з 26 найменувань. У роботі наведено 3 рисунки та 6 таблиць.

Метою даної дипломної роботи є створення математичного та програмного забезпечення для побудови специфічних словників для розпізнавання мовленнєвих сигналів за тематикою.

У роботі проведено аналіз існуючих систем розпізнавання злитого мовлення та словників, які ними використовуються. Виконано їх порівняння з погляду точності отримуваних розв'язків, вартості, призначення, а також об'єму словника.

Сформовані вимоги до нормалізації текстів. Для кожної кожної тематики побудовано відповідні словники. Розроблено автоматизовану систему, що будує тематичні словники. Виконано тестування розробленої системи.

Ключові слова: корпуси текстів, файли ключових слів, нормалізація, лематизація, лінгвістична модель, метрика TFIDF, кластеризація.