

## РЕФЕРАТ

Магістерська дисертація: 93 ст., 11 рис., 32 табл., 40 джерел та 2 додатки.

Протягом останніх двох століть людство експотенційно накопичує об'єми текстових даних. Зростання об'ємів накопичених документів ставить нові вимоги до швидкості та якості пошуку.

З 1954-го року проводяться дослідження в області семантичного аналізу текстів. Новітні результати у даній області ведуться шляхом покращення моделі “Word2Vec”, анонсованої компанією Google у 2013-му році.

Метою даної роботи є розробка системи мультимодального текстового пошуку. У роботі проаналізовано методи вирішення задачі обробки текстів, побудови векторного простору для врахування семантики слів, також розглянуто задачі знаходження тем документу на основі багатьох модальностей.

Об'єктом дослідження є системи текстового пошуку, предметом дослідження - системи мультимодального пошуку та їх компоненти. Деякі результати даної роботи опубліковано на міжнародній конференції “ICTERI 2018”.

Результати роботи:

- реалізовано систему мультимодального текстового пошуку;
- реалізовано методи пошуку у багатовимірних розріджених просторах;
- вдосконалено існуючі алгоритми текстового пошуку.

Результати даної роботи рекомендовано до використання у пошуку по Інтернет-магазинах, каталогах продукції, інформаційних веб-сайтах, файлових системах, користувацьких повідомленнях. Проведено вимірювання, перевірено гіпотези, що підтверджують практичну значимість результатів та можливості їхнього застосування.

При подальшому розвитку системи доцільно додати нові методи оцінки релевантності документів, а також створити більш гнучку мову для керування даними.

МУЛЬТИМОДАЛЬНИЙ ТЕКСТОВИЙ ПОШУК, ІНФОРМАЦІЙНИЙ ПОШУК, МАШИННЕ НАВЧАННЯ, ПОВНОТЕКСТОВИЙ ПОШУК,

ОКАРІ-ВМ25, ЕМ-АЛГОРИТМ, ЙМОВІРНІСНИЙ ВИСНОВОК,  
СЕМАНТИЧНИЙ АНАЛІЗ ТЕКСТІВ