

АНОТАЦІЯ

Пояснювальну записку виконано на 73 аркушах, вона містить 3 додатки та перелік посилань на використані джерела з 15 найменувань. У роботі наведено 22 рисунки та 7 таблиць.

Основною метою створення системи є автоматизація роботи користувача ресурсами персонального управління знаннями, забезпечення його необхідним математичним та програмним забезпеченням для використання спеціалізованого функціоналу, необхідного для оптимізації управління знаннями. Система створюється для використання в області персонального управління знаннями із забезпеченням базових функцій системи управління знаннями та допоміжними функціями оптимізації роботи користувача.

Під час реалізації проекту було проведено значний порівняльний аналіз методів класифікації, серед них методи C4.5, метод опорних векторів та метод класифікації байеса на основі наївної байєсовської моделі класифікації, методи та алгоритми кластерного аналізу даних, серед яких EM-алгоритм та алгоритм k-середніх кластеризації, алгоритми рубрикації та попереднього виокремлення термінів документу. Вибраними для використання були метод класифікації байеса на основі наївної байєсовської моделі класифікації та алгоритм k-середніх кластерного аналізу даних.

В результаті дипломного проектування спроектовано систему персонального управління знаннями, яка представляє програмний продукт на мові програмування Java з використанням системи керування базою даних Oracle 11gR2. Було проведено тестування розробленої системи, використаних алгоритмів класифікації документів та кластеризації даних з використанням документів заданих тем, результати досліджень відповідають правильній роботі алгоритмів.

Система розроблялася для персонального використання користувачами систем управління власними знаннями, які використовують додатковий спеціалізований функціонал для автоматизації роботи з власними даними,

виконання швидкої класифікації документів по визначеним темам, перевірка відповідності теми до обраної.

Ключові слова: система управління знаннями, класифікація документів, кластеризація, документ, тема, лексема, співпадіння, ймовірність.

