

РЕФЕРАТ

Бакалаврська дипломна робота: 63 с., 17 рис., 8 табл., 2 додатки та 15 джерел.

Об'єкт дослідження – сайти та методи просування в пошукових системах.

Предмет дослідження – методи класифікації текстових документів для фільтрації пошукових запитів на основі ключових слів за типом.

Мета роботи – розробка модуля для генерації пошукових запитів на основі ключових слів, а також розробка користувацького інтерфейсу до нього.

Методи дослідження – логістична регресія, пасивно-агресивний класифікатор, алгоритм адабуст.

Актуальність – доступні інструменти для генерації пошукових запитів не передбачають їх фільтрації за інформаційним та комерційним типом. Розрізнення цих типів запитів є важливим при формуванні семантичного ядра.

Результати роботи – запропоновано алгоритм класифікації пошукових запитів; реалізовано запропоновану конфігурацію класифікатора; створено додаток із зручним для користувача графічним інтерфейсом.

Шляхи подальшого розвитку предмету дослідження – детекція мови, тегування географічних назв, перевірка граматики, розпізнавання категорій товарів.

КЛАСИФІКАЦІЯ, МАШИННЕ НАВЧАННЯ, ОБРОБКА ПРИРОДНОЇ МОВИ, СТЕМЕР ПОРТЕРА, TF-IDF, ЛОГІСТИЧНА РЕГРЕСІЯ, АДАБУСТ, ПАСИВНО-АГРЕСИВНИЙ КЛАСИФІКАТОР.

