232457 Разработка информационной системы анализа текста

Введение 2
1 Теоретическая часть 3
1.1 Классификация текстовых документов 3
1.2 Обзор и выбор метода классификации 16
1.2.1 Байесовский классификатор 16
1.2.2 Метод ближайших соседей 19
2 Аналитическая часть 24
2.1 Алгоритмы реализации метода ближайших соседей 24
2.1.1 Алгоритм взвешенных ближайших соседей 24
2.1.2 Алгоритм отыскания оптимальных параметров 24
2.1.3 Алгоритм отбора признаков 25
2.1.4 Алгоритм KNN 26
2.2 Постановка задачи 28
2.3 Выбор средств разработки 30
3 Проектная часть 46
3.1 Реализация интерфейса 46
3.2 Реализация серверной части автоматической классификации 54
3.3 Диаграммы UML 62
3.4 Анализ документов и выводы 70
Заключение 73
Список литературы 75