

## **232460 Создание высокопроизводительной организации данных и искусственного интеллекта на Azure Databricks**

### **Исследование актуальности проблемы высокопроизводительных структур данных 1**

#### **Введение 2**

#### **1.1. Актуальность темы 2**

#### **1.2. Цели и задачи исследования 3**

#### **1.3. Обзор литературы 9**

### **2. Теоретические основы высокопроизводительных структур данных 14**

#### **2.1. Понятие высокопроизводительности в контексте структур данных 14**

#### **2.2. Основные типы высокопроизводительных структур данных 16**

#### **2.3. Преимущества и недостатки каждого типа структур данных 21**

### **3. Анализ существующих высокопроизводительных структур данных 27**

#### **3.1. Описание и классификация существующих структур данных 27**

#### **3.2. Анализ производительности существующих структур данных 35**

#### **3.3. Выбор наиболее подходящих структур данных для дальнейшего исследования 39**

4. Разработка новой высокопроизводительной структуры данных	55
4.1. Описание принципов разработки новой структуры данных	55
Известные методы применения кластерного анализа	67
4.2. Разработка алгоритмов и структуры новой структуры данных	73
4.3. Оценка производительности новой структуры данных	78
5. Реализация и тестирование новой структуры данных	79
5.1. Описание процедуры реализации новой структуры данных	79
Сравнительный анализ рассмотренных методов	80
5.2. Тестирование производительности новой структуры данных	88
Функциональное тестирование	89
Разработка эксперимента при заданных условиях первого значения	89
Разработка эксперимента при заданных условиях второго значения	91
Разработка эксперимента при заданных условиях третьего значения	94
6. Заключение	98
6.1. Обобщение результатов исследования	98
Сравнение алгоритмов кластеризации информационного графа программы в соответствии с разработанной методикой	109



Для сравнения алгоритмов кластеризации разработано программное обеспечение, описание которого приведено ниже. Для запуска программы необходимо запустить файл Project.exe, после чего появится главная форма программы. 109

10.3. Разработка и тестирование платформы данных 117

12.2. Практическая значимость работы 121

Azure Databricks122

12.3. Рекомендации для дальнейших исследований 134

Список использованной литературы 137

