**Разработка алгоритма семантической декомпозиции текста на ключевые элементы для решения задачи категоризации**

Содержание

Введение

1. Обзор предметной области

1.1. Категоризация

1.2. Обработка естественного языка

1.3. Семантический анализ

1.4. Семантическая декомпозиция

1.5. Обзор существующих алгоритмов (подходов)

1.6. Сравнение существующих алгоритмов (подходов)

1.7. Выбор алгоритма (подхода)

1.8. Выводы

2. Описание процесса разработки

2.1. Алгоритм семантической декомпозиции на n-граммы

2.1.1. Структура и описание работы алгоритма

2.1.2. Технические подробности реализации  
2.1.3. Сценарий использования (?)

2.1.4. Апробация алгоритма

2.2. Семантическая декомпозиция на простые предложения

2.2.1. Структура и описание работы алгоритма

2.2.2. Технические подробности реализации  
2.2.3. Сценарий использования (?)

2.2.4. Апробация алгоритма

2.3. Выводы

3. Исследование

3.1. Исследование времени работы

3.1.1. Сценарий эксперимента

3.1.2. Методика эксперимента

3.1.3. Используемые программные средства

3.1.4. Результат эксперимента

3.1.5. Сравнение и оценка результатов

3.2. Исследование точности результатов

3.2.1. Сценарий эксперимента

3.2.2. Входные данные

3.2.3. Результат работы алгоритмов

3.2.4. Экспертная оценка

3.2.5. Сравнение с экспертной оценкой

3.3. Выводы

4. Доп раздел

Заключение

Список использованных источников

План

Введение

Актуальность, цель, задачи, объект, предмет, практическая значимость.

1. Обзор предметной области

1.1. Категоризация

Что такое категоризация, для чего применяется, (если будем связывать с СППР, то надо будет сделать пункт про нее), что для нее нужно, в частности этапы. Говорим, что для нее надо подготовленный и размеченный текст, а это обработка естественного языка, и переходим к следующему пункту.

1.2. Обработка естественного языка

Что такое, для чего применяется, какие задачи мы рассматриваем. Переходим к семантическому анализу.

1.3. Семантический анализ

Опять, что такое, для чего применяется, в каком качестве мы его рассматриваем, какие есть подзадачи. Переходим к декомпозиции.

1.4. Семантическая декомпозиция

Что такое, ценность, какая бывает.

1.5. Обзор существующих алгоритмов (подходов)

Обзор существующих алгоритмов (подходов), их описание.

1.6. Сравнение существующих алгоритмов (подходов)

Сравнение конкретных показателей в таблице, описание таблицы.

1.7. Выбор алгоритма (подхода)

Вывод по таблице = выбор.

1.8. Выводы

Выводы по главе.

2. Описание процесса разработки

Я еще не знаю точно, но скорее всего алгоритмы для разделения текста на n-граммы и на простые предложения будут отличаться. Также не знаю, успею ли я поработать над двумя, но пока пишу по двоим.

2.1. Алгоритм семантической декомпозиции на n-граммы

2.1.1. Структура и описание работы алгоритма

Логика работы, схема, ее расшифровка, описание входных/выходных данных.

2.1.2. Технические подробности реализации

На каком языке написан алгоритм, список использованных библиотек с их описанием.

2.1.3. Сценарий использования

«Нужно ввести текст в файл Текст.txt, зайти в командную строку ввести такую-то команду и т.д.».

2.1.4. Апробация алгоритма

Здесь о том, что разработанный алгоритм до настоящего времени не внедрялся в программные продукты организации, но в будущем может быть использован в СППР для службы технической поддержки ООО ГП.

2.2. Семантическая декомпозиция на простые предложения

2.2.1. Структура и описание работы алгоритма

2.2.2. Технические подробности реализации  
2.2.3. Сценарий использования

2.2.4. Апробация алгоритма

2.3. Выводы

Выводы по главе.

3. Исследование

3.1. Исследование времени работы

3.1.1. Сценарий эксперимента

Как проводилось исследование времени работы.

3.1.2. Методика эксперимента

Характеристики ВМ (ОС, системные характеристики), какие входные данные.

3.1.3. Используемые программные средства

Язык программирования, библиотеки.

3.1.4. Результат эксперимента

Результат исследования времени работы алгоритма (+ при внедрении в СППР – время работы СППР с этим алгоритмом).

3.1.5. Сравнение и оценка результатов

Сравнение и оценка.

3.2. Исследование точности результатов

3.2.1. Сценарий эксперимента

Как проводилось исследование точности работы.

3.2.2. Входные данные

Какие входные данные.

3.2.3. Результат работы алгоритмов

Какой результат дали алгоритмы.

3.2.4. Экспертная оценка

Какой результат должен был получиться.

3.2.5. Сравнение с экспертной оценкой

Сравнение

3.3. Выводы

Выводы по главе.

4. Доп раздел

Заключение

Подведение итогов по работе

Список использованных источников