**Малых Валентин Андреевич Методы сравнения и построения устойчивых к шуму программных систем в задачах обработки текстов**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Малых Валентин Андреевич

Введение

Глава 1. Современные методы оценки устойчивости и общие

подходы к построению систем обработки текстов

1.1 Моделирование шума и оценка устойчивости к шуму в задачах обработки естественного языка

1.2 Векторные представления текстов и слов

1.2.1 TF-IDF

1.2.2 Система Word2Vec

1.2.3 Модель fastText

1.2.4 Система GloVe

1.2.5 Другие системы векторного представления слов

1.3 Общие подходы к построению нейронных сетей

1.3.1 Полносвязные нейронные сети

1.3.2 Рекуррентные нейронные сети

1.3.3 Сверточные нейронные сети

1.4 Задача классификации текстов

1.5 Задача извлечения именованных сущностей

1.6 Задача извлечения аспектов

1.7 Выводы к главе

Глава 2. Устойчивые к шуму вектора слов

2.1 Постановка задачи

2.1.1 Описание задачи

2.2 Векторные представления слов

2.2.1 Векторные представления слов на уровне символов

2.2.2 Описание используемых подходов к обработке текстов

2.2.3 Ячейка нейронной сети Simple Recurrent Unit

2.3 Разработанная система: RoVe

2.3.1 Представление слов BME

2.3.2 Архитектура системы

2.4 Исследуемые варианты системы RoVe

2.4.1 Системы с рекуррентными кодировщиками

2.4.2 Системы, основанные на сверточном кодировщике

2.5 Наборы данных

2.6 Эксперименты на прикладных задачах обработки текстов

2.6.1 Метод сравнения систем векторных представлений слов

2.6.2 Используемые для сравнения задачи

2.6.3 Постановка экспериментов

2.6.4 Метрики оценки качества для тестируемых задач

2.6.5 Наивный Байесовский классификатор

2.6.6 Обучение систем

2.7 Результаты экспериментов для векторных представлений слов

2.7.1 Эксперименты для английского языка

2.7.2 Эксперименты для русского языка

2.7.3 Эксперименты для турецкого языка

2.8 Изучение влияния шума

2.9 Анализ результатов сравнения систем на прикладных задачах обработки текстов

2.10 Выводы к главе

Глава 3. Устойчивая к шуму классификация текстов

3.1 Системы классификации текстов

3.1.1 CharCNN

3.1.2 FastText

3.1.3 CharCNN-WordRNN

3.1.4 RoVe

3.2 Эксперименты по сравнению систем классификации текстов

3.2.1 Наборы данных для задачи классификации текстов

3.2.2 Метод сравнения систем классификации текстов

3.3 Результаты экспериментов для задачи классификации текстов

3.3.1 Набор данных SentiRuEval-2015

3.3.2 Набор данных Airline Twitter Sentiment

3.4 Выводы к главе

Глава 4. Распознавание именованных сущностей в шумных

текстах

4.1 Базовая архитектура системы распознавания именованных сущностей

4.2 Наборы данных для задачи распознавания именованных сущностей

4.2.1 Набор данных ^N^'03

4.2.2 Корпус Peгsons-1000

4.2.3 Корпус CAp'2017

4.3 Эксперименты с вариантами системы biLSTM-CRF

4.3.1 Метод сравнения систем распознавания именованных сущностей

4.3.2 Варианты системы распознавания именованных сущностей

4.4 Результаты для задачи распознавания именованных сущностей

4.5 Выводы к главе

Глава 5. Извлечение аспектов в шумных данных

5.1 Система извлечения аспектов на основе внимания (ABAE)

5.2 Модификации системы ABAE

5.3 Эксперименты по сравнению систем извлечения аспектов

5.3.1 Метод сравнения систем извлечения аспектов

5.3.2 Постановка экспериментов для задачи извлечения аспектов

5.4 Результаты экспериментов для задачи извлечения аспектов

5.5 Выводы к главе

Глава 6. Комплекс программ для оценки устойчивости к шуму

систем для задач обработки текстов

6.1 Пакет программ для построения и оценки устойчивости систем векторных представлений слов

6.1.1 Система RoVe

6.1.2 Система Woгd2Vec

6.1.3 Система fasttext

6.2 Пакет программ для построения и оценки устойчивости систем классификации текстов

6.2.1 Система CharCNN

6.2.2 Система Fasttext-GRU

6.2.3 Система CharCNN-WordRNN

6.2.4 Система RoVe

6.3 Пакет программ для построения и оценки устойчивости систем

распознавания именованных сущностей

6.3.1 Система LSTM-CRF

6.4 Пакет программ для построения и оценки устойчивости систем извлечения аспектов

6.5 Выводы к главе

Заключение

Список сокращений и условных обозначений

Словарь терминов

Список литературы

Список рисунков

Список таблиц