**Коновалов Василий Павлович Методы переноса знаний для нейросетевых моделей обработки естественного языка**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Коновалов Василий Павлович

Введение

Глава 1. Применение переноса знаний к проблемам обработки

естественного языка

1.1 Парадигма переноса знаний

1.2 Таксономия переноса знаний

1.2.1 Последовательный перенос знаний

1.2.2 Многозадачное обучение

1.2.3 Межъязыковой перенос знаний

1.2.4 Адаптация домена

1.3 Архитектура Трансформер

1.3.1 Модуль самовнимания

1.3.2 Позиционный полносвязный слой

1.3.3 Варианты применения архитектуры Трансформер

1.3.4 Трансформер в задаче машинного перевода

1.4 Двунаправленная языковая модель BERT

1.4.1 Задачи предобучения BERT

1.4.2 Формат входных данных BERT

1.5 Последовательный перенос знаний на основе BERT

1.5.1 Набор данных GLUE

1.5.2 Результат переноса знаний на основе BERT

1.6 Многозадачное обучение на основе BERT

1.6.1 Архитектура модели MT-DNN

1.6.2 Результаты модели MT-DNN на задачах GLUE

1.7 Многоязычный BERT и межъязыковой перенос

1.7.1 Межъязыковой перенос на задачах разметки последовательностей

1.7.2 Результаты межъязыкового переноса

1.7.3 Адаптация многоязычного BERT для русского языка

1.7.4 Языковая фильтрация многоязычного BERT

1.8 Адаптация домена на основе BERT

1.8.1 Описание наборов данных

1.8.2 Результаты экспериментов

Глава 2. Обучение векторного представления слов для

малоресурсных языков

2.1 Методы построения векторных представлений слов

2.2 Корпуса для обучения

2.2.1 Бурятский язык

2.2.2 Эрзянский язык

2.2.3 Коми язык

2.3 Внутренняя оценка качества векторных представлений слов

2.4 Результаты сравнения методов

2.5 Внешняя оценка качества векторных представлений слов

2.6 Проект универсальных зависимостей

2.6.1 Бурятский язык

2.6.2 Эрзянский язык

2.6.3 Коми язык

2.7 Описание модели

2.8 Результаты сравнения методов

Глава 3. Межъязыковой перенос для модели

вопросно-ответных систем

3.1 Вопросно-ответные системы

3.2 Вопросно-ответные датасеты

3.3 Метрики качества вопросно-ответных систем

3.4 Базовые модели вопросно-ответных систем

3.4.1 Базовая модель R-Net

3.4.2 Базовая модель BiDAF

3.5 Вопросно-ответные модели на основе BERT

3.6 Цели и описание экспериментов

3.7 Результаты эксперимента

3.7.1 Кривые обучения

Глава 4. Отслеживание состояния диалога путем переноса

модели

4.1 Диалоговые системы

4.2 Отслеживание состояния диалога

4.3 Обзор решений

4.4 Схемоориентированный набор данных БСБ

4.5 Архитектура модели ССЬСЫЕ

4.5.1 Входные значения

4.5.2 Классификаторы ССЬСЫЕ

4.6 Метрики качества

4.7 Эксперименты и результаты

4.8 Исследование вклада отдельных компонентов модели

Заключение

Список сокращений и условных обозначений

Словарь терминов

Список литературы

Список рисунков

Список таблиц

Приложение А. Характеристики обученных вопросно-ответных

моделей

А.1 Сравнение количества обучающих итераций для

вопросно-ответных моделей

А.2 Характеристики используемых моделей для дообучения на

вопросно-ответную задачу