**Тихомиров Михаил Михайлович Методы автоматизированного пополнения графов знаний на основе векторных представлений**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Тихомиров Михаил Михайлович

Введение

Глава 1. Векторные представления в задачах автоматической

обработки текстов

1.1 Векторные представления слов

1.1.1 Матрица совместной встречаемости слов

1.1.2 Векторная модель Word2Vec

1.1.3 Векторная модель FastText

1.1.4 Векторная модель GloVe

1.2 Векторные представления графов

1.2.1 Графовая векторная модель DeepWalk

1.2.2 Графовая векторная модельNode2Vec

1.2.3 Графовая векторная модель TADW

1.2.4 Графовая векторная модель TransE

1.2.5 Графовая векторная модель Poincare Embeddings

1.2.6 Графовая сверточная сеть GCN

1.3 Мета-векторные представления

1.4 Контекстуализированные векторные представления

1.4.1 Контекстуализированная векторная модель ELMO

1.4.2 Контекстуализированная векторная модель ULMFiT

1.4.3 Контекстуализированная векторная модель BERT

1.5 Векторные представления и нейронные сети в задаче

пополнения таксономии

1.5.1 Подходы к предсказанию гиперонимов

1.6 Векторные представления и нейронные сети в задаче

извлечения именованных сущностей

1.6.1 Задача извлечения именованных сущностей

1.6.2 Методы в задаче извлечения именованных сущностей

1.6.3 Извлечение именованных сущностей в области информационной безопасности

Глава 2. Пополнение таксономии графов знаний новыми

понятиями

2.1 Постановка задачи

2.2 Подходы

2.2.1 Комбинированный подход на основе шаблонов и векторных представлений слов

2.2.2 Комбинированный подход на основе мета-векторных представлений слов

2.3 Описание данных и меры оценки

2.3.1 Набор данных RUSSE'2020

2.3.2 Набор данных Diachronie wordnets

2.3.3 Набор данных для адаптации таксономии на предметную область информационной безопасности

2.4 Эксперименты

2.4.1 Меры оценки

2.4.2 Эксперименты на наборе данных RUSSE'2020

2.4.3 Эксперименты на наборе данных Diachronie wordnets

2.4.4 Эксперименты на наборе данных OENTCyber

2.5 Выводы

Глава 3. Методы пополнения графов знаний именованными

сущностями в конкретной предметной области

3.1 Задача извлечения именованных сущностей

3.2 Постановка задачи

3.3 Используемый подход и модели

3.3.1 Методы дополнения данных

3.3.2 Подход на основе контекстуализированной векторной модели BERT

3.4 Описание данных для задачи извлечения именованных

сущностей в области информационной безопасности

3.4.1 Дополнение Sec\_col тренировочными данными

(порождение псевдоразметки)

3.5 Эксперименты

3.5.1 Оценка производительности

3.6 Выводы

Глава 4. Программный комплекс автоматизированного

пополнения графов знаний

4.1 Схема программного комплекса

4.2 Сервис предсказаний

4.3 Сервис разметки

4.4 Модуль обучения

Заключение

Список литературы

Список рисунков

Список таблиц