

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

**«Анализ методов обработки естественного языка в задаче анализа
сложных вопросов»**

Копейко Владимир Сергеевич

Научный руководитель – кандидат физико-математических наук,
доцент Пилипчук Л.А

Минск, 2020

Реферат

Магистерская диссертация, 57 страниц, 36 рисунков, 14 таблиц, 18 источников.

ВЕКТОРНЫЕ ПРЕДСТАВЛЕНИЯ СЛОВ, МОДЕЛЬ BERT, МОДЕЛЬ BART, МОДЕЛЬ XLNET, АРХИТЕКТУРА ТРАНСФОРМЕР, МЕХАНИЗМ ВНИМАНИЯ, СЕТЬ ЭНКОДЕР-ДЕКОДЕР, TD-IDF МОДЕЛЬ.

Объект исследования – набор данных, состоящий из пар (вопрос, ответ) и описательных характеристик для каждой такой пары.

Цель работы – разработка моделей глубокого обучения для решения задачи классификации вопросов и ответов; сравнительный анализ работы реализованных моделей.

Методы исследования – методы машинного обучения, использующие векторные представления слов.

Результат – реализованные модели глубокого обучения, а также оценки эффективности каждой из реализованных моделей.

Область применения – автоматические системы нахождения ответов на вопросы.

Abstract

Master's thesis, 57 pages, 36 figures, 14 tables, 18 sources.

VECTOR REPRESENTATIONS OF WORDS, BERT MODEL, BART MODEL, XLNET MODEL, TRANSFORMER ARCHITECTURE, ATTENTION MECHANISM, ENCODER-DECODER NETWORK, TD-IDF MODEL

Object of research – data set consisting of pairs (question, answer) and descriptive characteristics for each such pair.

Research goal – the development of deep learning models to solve the problem of classification of questions and answers; comparative analysis of the implemented models.

Research methods – deep learning methods using vector representations of words.

Result – implemented deep learning models, as well as evaluating the effectiveness of each of the implemented models.

Application field – automatic systems for finding answers to questions.