

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

**«Автоматическое преобразование белорусской речи в текст методами
глубокого обучения»**

Трофимов Александр Сергеевич

Научный руководитель – кандидат технических наук, заведующий лаборатории
распознавания и синтеза речи ОИПИ НАНА РБ Гецевич Ю.С.

Минск, 2022

Реферат

Магистерская диссертация, 52 страницы, 13 рисунков, 5 таблиц, 3 формулы, 27 источников.

Ключевые слова: РАСПОЗНАВАНИЕ РЕЧИ, ПРЕОБРАЗОВАНИЕ РЕЧИ В ТЕКСТ, НЕЙРОННЫЕ СЕТИ, ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ, ГЛУБОКОЕ ОБУЧЕНИЕ, WAV2VEC2, БЕЛОРУССКИЙ ЯЗЫК.

Объектом исследования является автоматическое распознавание речи.

Предметом исследования является автоматическое распознавание общей белорусской речи методами глубокого обучения.

Целью работы являлось построение качественной модели распознавания общей белорусской речи, а также сбор необходимых для этого данных.

В ходе работы были проанализированы основные подходы к распознаванию речи а также предыдущие эксперименты по построению систем распознавания белорусской речи; собран большой корпус начитанных белорусских текстов; успешно обучена система распознавания речи на основе архитектуры Wav2Vec2 с использованием n-грамной лингвистической модели; полученная система опубликована на платформе Hugging Face; создан и опубликован на платформе Hugging Face демонстрационный веб-интерфейс, который позволяет распознавать белорусскую речь прямо через браузер.

Полученный результат можно использовать для создания системы голосового ввода текста, построения голосовых помощников, создания систем автоматизированного обучения белорусскому языку.

Abstract

Master thesis, 52 pages, 13 figures, 5 tables, 3 formulas, 27 resources.

Keywords: SPEECH RECOGNITION, SPEECH-TO-TEXT, SPEECH TO TEXT, NEURAL NETWORKS, DEEP NEURAL NETWORKS, DEEP LEARNING, WAV2VEC2, BELARUSIAN LANGUAGE.

The object of research is automated speech recognition.

The subject of study is automated speech recognition for general Belarusian language using deep learning methods.

The aim of this work is to build a good speech recognition model for general Belarusian language and to gather a required dataset for this task.

During the study main speech recognition approaches were analyzed as well as previous experiments to build speech recognition models for Belarusian language; a large dataset of voiced texts was gathered; a Wav2Vec2-based speech recognition system was successfully built using additional n-gram Language model; the system was published on Hugging Face platform; a demo web-interface was built and published on Hugging Face platform that allows to recognize Belarusian speech directly from a web-browser.

The result can be applied to develop voice input systems, build voice assistants, create a system to teach Belarusian language in an automated way.