

**Белорусский государственный университет  
Механико-математический факультет  
Кафедра дифференциальных уравнений и системного анализа**

**Аннотация к магистерской диссертации  
“Автоматическое распознавание речевых команд”**

**Бондарик Ирина Игоревна**

**руководитель Голубева Лариса Леонидовна**

**2021**

Магистерская диссертация содержит: 38 страниц, 24 иллюстрации, 2 таблицы, 4 приложения, 22 использованных источников литературы.

Ключевые слова: АУДИОСИГНАЛ, СПЕКТРОГРАММА, МЕЛ-ЧАСТОТНЫЕ КЕПСТРАЛЬНЫЕ КОЭФФИЦИЕНТЫ, СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ, ОСТАТОЧНАЯ НЕЙРОННАЯ СЕТЬ, РЕКУРРЕНТНАЯ НЕЙРОННАЯ СЕТЬ.

В магистерской диссертации изучается классификация аудиозаписей по 10 классам с помощью нейронных сетей.

Целью работы является исследование эффективности различных архитектур искусственных нейронных сетей для выполнения классификации аудиозаписей на категории.

Актуальность работы обусловлена тем, что очень много приложений реального мира начинают использовать голосовой интерфейс, так как это позволяет быстрее и комфортнее взаимодействовать с устройствами. Сложность заключается в том, что на данный момент нейронные сети не могут со 100-процентной вероятностью распознавать человеческую речь, поэтому возникает необходимость проведения исследований в данной области.

В магистерской диссертации были получены следующие результаты:

- описаны существующие методы и алгоритмы обработки сигналов и распознавания речи;
- обработан и проанализирован исходный набор аудиозаписей для обучения нейронных сетей;
- классифицированы аудиозаписи на 10 классов с использованием нейронных сетей;
- проведен сравнительный анализ результатов классификации.

Магистерская диссертация носит практический характер. Ее результаты могут быть использованы для улучшения результатов классификации аудиозаписей. Новизна результатов состоит в исследовании эффективности различных архитектур искусственных нейронных сетей. Обоснованность и достоверность полученных результатов обусловлена согласованностью с известными алгоритмами. Практическая часть была выполнена на языке программирования Python.

The master's thesis contains: 38 pages, 24 illustrations, 2 tables, 4 appendices, 22 used literature sources.

Key words: AUDIO SIGNAL, SPECTROGRAM, MEL-FREQUENCY CEPSTRAL COEFFICIENTS, CONVOLUTIONAL NEURAL NETWORK, RECURRENT NEURAL NETWORK, RESIDUAL NEURAL NETWORK.

The master's thesis examines the classification of audio recordings into 10 classes using neural networks.

The aim of this work is to study the effectiveness of various convolutional neural network architectures for classifying audio recordings into categories.

The relevance of the work is due to the fact that a lot of real-world applications are beginning to use the voice interface, as this allows you to interact with devices faster and more comfortably. The difficulty lies in the fact that at the moment neural networks cannot recognize human speech with 100% probability, so there is a need for research in this area.

The following results were obtained in the master's thesis:

- the existing methods and algorithms for signal processing and speech recognition are described;
- processed and analyzed the original set of audio recordings for training neural networks;
- classified audio recordings into 10 classes using neural networks;
- a comparative analysis of the classification results was carried out.

The master's thesis is of a practical one. Its results can be used to improve the results of the classification of audio recordings. The novelty of the results lies in the study of the effectiveness of various artificial neural network architectures. The validity and reliability of the results obtained is due to the consistency with known algorithms. The practical part was done in the Python programming language.