

DOI: 10.61726/6666.2024.86.71.001

УДК 004.738;615.47

СТРУКТУРА СЕТИ ИНТЕРНЕТА ВЕЩЕЙ И АЛГОРИТМЫ ДЛЯ ДИАГНОСТИКИ НЕВРОЛОГИЧЕСКОГО ЗАБОЛЕВАНИЯ

В.А. ВИШНЯКОВ¹, С. ИВЕЙ²

¹Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»,

ул. П. Бровки, 6, Минск, 220013, Беларусь
ORCID: <https://orcid.org/0000-0003-2929-8958>

²Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»,

ул. П. Бровки, 6, Минск, 220013, Беларусь
ORCID: <https://orcid.org/0000-0002-0269-4320>

Поступила в редакцию 03 апреля 2024

Цель статьи – исследовать инновационный подход, использующий технологии машинного обучения, нейронных сетей, обработки сигналов, сети Интернета вещей (IoT) для раннего выявления болезни Паркинсона. Представлена модель и алгоритмы обработки звуковых сигналов от пациентов, исследуемых на вероятность заболевания болезнью Паркинсона. Приведена структура и функции сети интернета вещей, которая использует Thingspeak в качестве платформы, позволяя собирать голосовые данные с помощью мобильных телефонов, извлекать из них соответствующие функции и обрабатывать данные. В рамках этого процесса для расчета вероятности болезни Паркинсона используется модель полностью подключенной нейронной сети (FCNN), которая предоставляет медицинским работникам и пациентам удобный, точный и ранний диагностический ИТ-инструмент. В исследовании рассматриваются структура, алгоритмы и роль FCNN в сети Интернета вещей, подчеркивается ее влияние на сектор ИТ медицины.

Ключевые слова: технология Интернет вещей, удаленная ИТ-диагностика, анализ голосовых данных, шумоподавление, нейронная сеть.

Введение. По мере дальнейшего развития общества и прогресса науки и техники все большее внимание уделяется раннему выявлению и диагностике проблем со здоровьем. Болезнь Паркинсона как хроническое неврологическое заболевание оказывает глубокое влияние на качество жизни заболевших людей. Ранняя диагностика имеет решающее значение для обеспечения более эффективного лечения и ухода [1, 2]. В работе [3] авторы представили подход к распознаванию болезни Паркинсона на основе голосового анализа. Основываясь на нем, авторы предложили использовать сеть Интернета вещей, которая собирает голосовые данные с помощью мобильных телефонов, извлекает соответствующие функции, облегчает передачу и обработку данных и, в конечном счете, выводит вероятность болезни Паркинсона. Внедрение этого метода обещает предоставить более удобный, точный и ранний диагностический инструмент как для медицинских работников, так и для пациентов. В этой статье авторы рассматривают структуру и алгоритмы работы этой сети Интернет.

Методы исследования. Первоначально голосовые данные собираются и предварительно обрабатываются с помощью мобильного телефона. Это включает в себя сбор речевых данных от пациентов с болезнью Паркинсона (БП) в течение 5 секунд с частотой дискретизации 44,1 кГц. Для повышения качества сигнала используется алгоритм спектрального вычитания [4] для устранения окружающего шума.

Впоследствии характеристики извлекаются из предварительно обработанных речевых данных после подавления шума. Затем они передаются на голосовой канал платформы Интернета вещей Thingspeak. В платформе Thingspeak данные передаются в функцию анализа MATLAB.

Модуль анализа MATLAB играет ключевую роль в интерпретации данных путем загрузки 3-уровневой модели FCNN, развернутой в облаке. Он обрабатывает данные и генерирует значение вероятности, указывающее на вероятность возможного диагноза болезни Паркинсона.

Наконец, результаты передаются с платформы Thingspeak IoT на мобильный телефон через функцию ThingHTTP и отображаются на телефоне для дальнейшего изучения и оценки.

Модель и алгоритмы сбора и обработки данных. I. Модель и алгоритм вычитания спектра для голосовых данных. Процесс включает следующие шаги:

1. Разделение исходного звукового сигнала на кадры, длина каждого кадра составляет 256 сэмплов, используя 50 % перекрытие, чтобы разделить кадры и получить серию кадров сигнала. Пусть исходный звуковой сигнал обозначается как $S[n]$, где n – индекс выборки. Пусть длина кадра равна $L = 256$ выборкам, а процент перекрытия равен 50 %. Определим индекс кадра k таким образом, чтобы начальный индекс k -го кадра задавался формулой:

$$n_k = (k - 1) \times \frac{L}{2}, k = 1, 2, 3, \dots \quad (1)$$

затем k -й кадр сигнала задается следующим образом:

$$s_k[n] = s[n_k + n_{k+1}], 0 \leq n < L. \quad (2)$$

2. Выполнение преобразования Фурье для каждого кадра сигнала, чтобы получить соответствующий спектр; применение оконной функции $w[n]$ к кадру, чтобы уменьшить спектральную утечку, такую как окно Хэмминга или окно Ханнинга. Вычисление преобразования Фурье оконного фрейма, чтобы получить комплексный спектр $X_k[f]$:

$$X_k[f] = \text{sum} \left(w[n] \times s_k[n] \times \exp \left(-j \times 2 \times \pi i \times f \times \frac{n}{L} \right) \right), 0 \leq f \leq \frac{L}{2} \quad (3)$$

3. Усреднение спектра всех кадров, чтобы получить средний спектр всего сигнала, и использование его в качестве основы для расчета спектра шума. Пусть $X_k[f]$ – комплексный спектр k -го кадра сигнала. Вычисление среднего спектра всех кадров следующим образом:

$$X_{avg}[f] = \left(\frac{1}{K} \right) \times \text{sum}(X_k[f]), \quad (4)$$

где K – число кадров.

Использование алгоритма спектрального вычитания $X_{avg}[f]$ в качестве основы для вычисления спектра шума.

4. Фильтрация спектра шума, чтобы получить более точную оценку шума. Пусть $X_{noise}[f]$ – спектр шума, оцененный по $X_{avg}[f]$. Применение фильтра Винера к $X_{noise}[f]$, чтобы получить более точную оценку:

$$X_{noise}^{filtered}[f] = \frac{X_{noise}[f]}{(X_{noise}[f] + \alpha * X_s[f])}, \quad (5)$$

где α – параметр сглаживания, а $X_s[f]$ – спектр оригинального сигнала.

5. Сравнение спектра каждого кадра со спектром шума, отобранного фильтром Винером, вычисление отношения сигнал/шум и рассмотрение частотных составляющих с отношением сигнал/шум менее 10 дБ в качестве шумовых составляющих [5]. Пусть $X_k[f]$ – комплексный спектр k -го кадра сигнала. Вычисление отношения сигнал/шум (SNR) для каждой частотной составляющей в виде:

$$SNR[f] = 10 \times \log_{10} \left(\frac{|X_k[f]|^2}{|X_{filtered}[f]|^2} \right). \quad (6)$$

Определение частотных составляющих с $SNR < 10$ дБ в качестве составляющих шума.

6. Вычитание частотных составляющих шума, установив коэффициент равным 0,5, чтобы получить спектр, удаленный от шума. Пусть $X_k^{noiseless}$ – бесшумный спектр k -го кадра сигнала. Для каждой частотной составляющей, если $SNR[f] < 10$ дБ, установить величину $X_k^{noiseless}[f]$ равной 6:

$$|X_k^{noiseless}[f]| = 0.5 \times |X_k[f] - X_{noise}^{filtered}[f]|. \quad (7)$$

7. Чтобы получить полный сигнал после удаления шума, нужно преобразовать удаленный шумом спектр обратно во временную область и наложить каждый кадр. Пусть $X_k^{noiseless}[f]$ – бесшумный сигнал в частотной области k -го кадра сигнала, а $X_k^{noiseless}$ – соответствующий сигнал во временной области. Аналогично пусть $Y_k[f]$ – удаленный от шума спектр k -го кадра, а $Y_k[n]$ – соответствующий сигнал во

временной области. Чтобы преобразовать удаленный шумом спектр обратно во временную область, мы можем применить обратное преобразование Фурье к $Y_k[f]$, что дает нам $Y_k[n]$:

$$Y_k[n] = IFFT(Y_k[f]). \quad (8)$$

Затем можно объединить удаленный от шума сигнал каждого кадра, чтобы получить полный сигнал без шума:

$$x_{noiseless}[n] = \text{sum}_k(x_k^{noiseless}[n] \times w_k[n]), \quad (9)$$

где $w_k[n]$ – оконная функция, применяемая к каждому кадру, и сумма берется по всем кадрам.

На рис. 1 показана блок-схема предварительной обработки данных.



Рис. 1. Блок-схема предварительной обработки данных

II. Голосовые данные после устранения шума отображаются в окне. Основное преимущество использования окна Хэмминга [6] для извлечения окна сигнала заключается в том, что оно может уменьшить эффект колебаний на краю окна, сохраняя основные компоненты сигнала внутри окна. Размер окна составляет 1024, а частота голосовых данных – 44,1 кГц. Частота голосовых данных составляет 44,1 кГц, а частота перекрытия окна составляет 50 %, поэтому время речи в окне составляет около 23 мс.

Пусть $x[n]$ – исходный сигнал во временной области с частотой дискретизации $f_s = 44,1$ кГц. Пусть $w[n]$ – окно Хэмминга размером $N = 1024$. Оконный сигнал $x_w[n]$ получается путем умножения $x[n]$ на значение окна $w[n]$ и сдвига окна на размер перехода $H = \frac{N}{2}$:

$$x_w[n] = x[n] \times w[n - n_0], \quad (10)$$

где $n_0 = k \times H$ целых k .

Окно Хэмминга $w[n]$ определяется как:

$$w[n] = 0.54 - 0.46 \times \cos\left(2 \times \pi i \times \frac{n}{N-1}\right), 0 \leq n \leq N-1 \quad (11)$$

Длительность каждого оконного сегмента равна $T = \frac{N}{f_s} = \frac{1024}{44100}$, $S = 0.023$ sec (23 ms), а перекрытие между соседними сегментами равно $H/T = 2$.

Извлечение признаков. Функция извлечения голосовых данных [7] выполняется в указанном окне, как показано в табл. 1 ниже.

Таблица 1

Извлечение всех объектов

Номер	Название объекта	Описание
1	MDVP:Fo(Hz)	Средняя основная частота вокала
2	MDVP:Fhi(Hz)	Максимальная основная частота вокала
3	MDVP:Flo(Hz)	Минимальная основная частота вокала
4	MDVP:Jitter(%)	Мера изменения основной частоты (в процентах)
5	MDVP:Jitter(Abs)	Мера изменения основной частоты (абсолютное значение)
6	MDVP:RAP	Мера изменения основной частоты (относительное амплитудное возмущение)
7	MDVP:PPQ	Мера изменения основной частоты (коэффициент возмущения периода основного тона)
8	Jitter:DDP	Мера изменения основной частоты (среднее значение абсолютных разностей разностей между соседними периодами)
9	MDVP:Shimmer	Мера изменения амплитуды (локальное изменение амплитуды)
10	MDVP:Shimmer(dB)	Мера изменения амплитуды (локальное изменение амплитуды в дБ)
11	Shimmer:APQ3	Мера изменения амплитуды (коэффициент амплитудного возмущения, 3-балльный метод)
12	Shimmer:APQ5	Мера изменения амплитуды (коэффициент амплитудного возмущения, 5-балльный метод)
13	MDVP:APQ	Мера изменения амплитуды (коэффициент возмущения средней амплитуды)
14	Shimmer:DDA	Мера изменения амплитуды (средняя абсолютная разница амплитуд между последовательными периодами)
15	NHR	Отношение шума к тональным составляющим в голосе
16	HNR	Отношение гармоник к шуму в голосе
17	RPDE	Мера нелинейной динамической сложности
18	D2	Мера нелинейной динамической сложности
19	DFA	Показатель фрактального масштабирования сигнала
20	spread1	Нелинейная мера изменения основной частоты
21	spread2	Нелинейная мера изменения основной частоты
22	PPE	Нелинейная мера изменения основной частоты

Передача и обработка данных. Чтобы загрузить 22 функции в Thingspeak [8] IoT для анализа выполняется алгоритм.

1. Создание канала. На рис. 2 ниже показана настройка голосового канала.

Channel Settings

Percentage complete 30%

Channel ID 2045599

Name

Description

Field 1

Field 2

Field 3

Field 4

Field 5

Field 6

Field 7

Field 8

Рис. 2. Настройки голосового канала

2. Получение API-ключа записи/чтения голосового канала и идентификатора канала. На рис. 3 ниже показан API-ключ записи/чтения голосового канала и идентификатор канала.

voice

Channel ID: 2045599
Author: mwa0000029248662
Access: Public

Private View Public View Channel Settings Sharing API Keys

Write API Key

Key: RT2VJK2XDGSU4G7U

Generate New Write API Key

Read API Keys

Key: Y7TRGS2U54UWG92P

Note:

Save Note Delete API Key

Add New Read API Key

Рис. 3. Ключ API для записи/чтения и идентификатор голосового канала

3. Использование API Thingspeak. Код библиотеки API использует протокол HTTP для загрузки 22 голосовых функций в голосовой канал. На рис. 4 показан алгоритм загрузки данных ThingSpeak API.

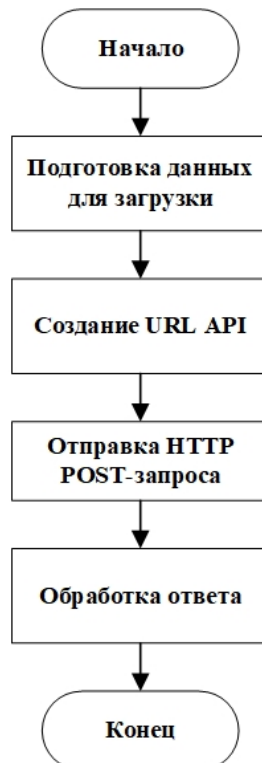


Рис. 4. Алгоритм загрузки данных ThingSpeak API

4. Загрузка предварительно обученной модели CNN [9] в модуль анализа MATLAB и ввод 22 данных в модель для анализа, чтобы получить результаты. Модель была обучена с использованием общедоступного набора данных по болезни Паркинсона [10]. На рис. 5 показаны схемы 3-слойного FCNN.

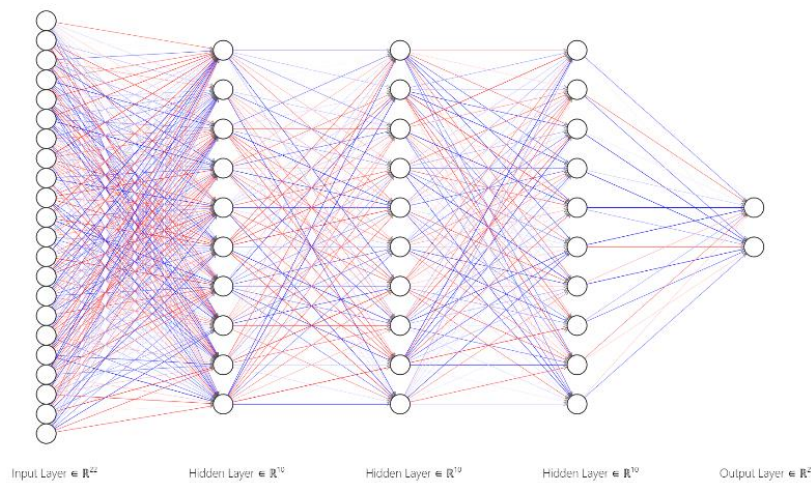


Рис. 5. Схема 3-слойной CNN

В табл. 2 приведены гиперпараметры 3-слойного CNN.

Таблица 2

Гиперпараметры 3-слойной CNN

Имя	Гиперпараметры
Размер первого слоя	10
Размер второго слоя	10
Размер третьего слоя	10
Функция активации	ReLU
Предел итерации	1000
Скорость обучения	0.01
Алгоритм обновления скорости обучения	SGD
Сила регуляризации (Лямбда)	0
Стандартизировать данные	Yes

На платформе Thingspeak сохраняются результирующее значение file1 в голосовом канале, а затем телефон считывает значение file1 из голосового канала.

Результаты исследований. На рис. 6 показан процесс ИТ-диагностики. Первоначально голосовые данные записываются на мобильный телефон с последующим извлечением голосовых функций. Затем мобильный телефон передает эти данные на платформу Интернета вещей, используя протокол HTTP. После обработки с помощью модели CNN платформа отправляет результаты идентификации обратно на мобильный телефон по протоколу HTTP для дальнейшего просмотра и анализа.

Процесс начинается с получения исходного сигнала от микрофона, далее сигнал проходит ряд этапов предварительной обработки, включая выделение высоких частот, подавление шума и сегментацию речевого кадра с использованием окна Хэмминга. Затем из этих обработанных данных извлекаются ключевые характеристики и компилируются в наборы данных. Эти наборы данных передаются в нейронную сеть для обучения и оптимизации модели, визуально представленную на рисунке в виде многослойной структуры. Завершением обучения и оптимизации является создание файла модели, способного классифицировать голосовые вводимые данные.

Платформа Интернета вещей Thingspeak предоставляет функциональные возможности для визуализации, хранения и управления данными. Оптимизированная модель развертывается и взаимодействует с внешними приложениями, как показано в запросах API и голосовом интерфейсе. Эта платформа объединяет моделирование нейронных сетей с облачным управлением данными для получения точных результатов распознавания.

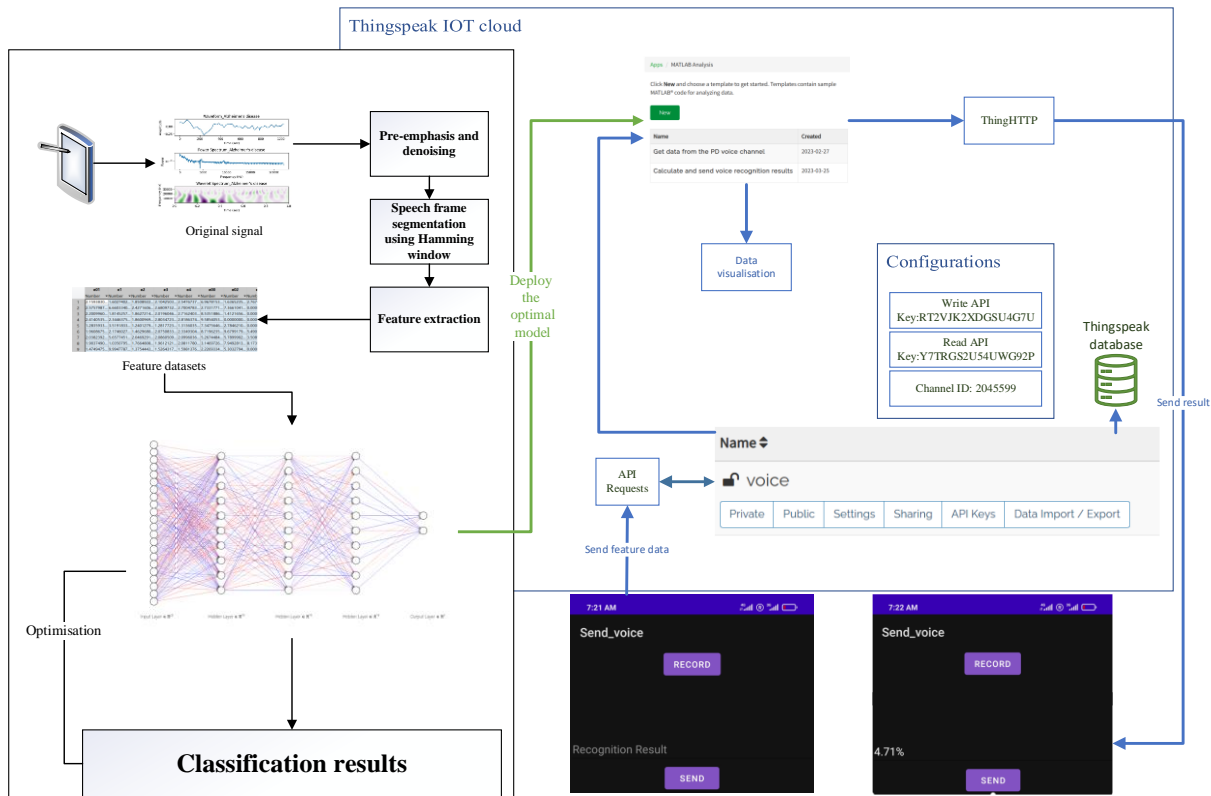


Рис. 6. Результат программирования сети Интернета вещей для ИТ-диагностики

Эксперимент был проведен на международном наборе данных [11]. Результаты тестовых экспериментов в сети ИВ для диагностики БП пациентов по изменению речи приведены в табл. 3.

Таблица 3

Данные тестовых экспериментов по распознаванию БП по изменению речи

Набор данных/ показатели	Средняя точность	Средняя чувствительность	Средняя F1 оценка	Точность тестирования
БП по речи	92,95 %	92,95 %	92,95 %	94,7 %

В сети ИВ достигнуто 94,7 % точности при диагностировании болезни Паркинсона на основе речевых данных и показателя F1 92,95 %. На том же наборе данных один из лучших показателей зарубежных исследований составляет 95,8 % [12], что свидетельствует как о хороших результатах распознавания и возможности внедрения сети ИВ для отечественной ИТ-диагностики БП.

Заключение. 1. Описан подход, использующий технологии машинного обучения, нейронных сетей, обработки сигналов, сети Интернета вещей (IoT) для раннего выявления болезни Паркинсона. Представлена модель и алгоритмы обработки звуковых сигналов от пациентов, исследуемых на вероятность заболевания болезнью Паркинсона. Сеть Интернета вещей собирает голосовые данные от пациентов с БП, обрабатывает их для устранения шума, извлекает важные характеристики и использует трехуровневую модель FCNN для получения результатов диагностики, основанных на вероятности, предлагая решение для своевременного выявления болезни Паркинсона.

2. Данная работа подчеркивает роль Интернета вещей в развитии ИТ-диагностики пациентов. Благодаря беспроводному подключению устройств сеть IoT не только обеспечивает удаленную диагностику, но и способствует расширению возможностей пациентов, врачей, персонала и повышению качества оказания медицинской помощи. В результате экспериментов на международном наборе данных в сети ИВ достигнуто 94,7 % точности при диагностировании болезни Паркинсона на основе речевых данных.

THE STRUCTURE OF THE INTERNET OF THINGS NETWORK AND ALGORITHMS FOR THE DIAGNOSIS OF NEUROLOGICAL DISEASE

U.A. VISHNIAKOU, XIA IWAY

Abstract

The purpose of this article is to explore an innovative approach using machine learning technologies, neural networks, signal processing, the Internet of Things (IoT) network for early detection of Parkinson's disease. The model and algorithms for processing audio signals from patients examined for the likelihood of Parkinson's disease are presented. The structure and functions of the Internet of Things network, which uses Thingspeak as a platform, allowing to collect voice data using mobile phones, extract relevant functions from them and process data, are given. As part of this process, a fully connected neural network (FCNN) model is used to calculate the probability of Parkinson's disease, which provides healthcare professionals and patients with a convenient, accurate and early diagnostic IT tool. The study examines the structure, algorithms and role of FCNN in the Internet of Things network, emphasizes its impact on the IT medicine sector.

Список литературы

1. Whitehouse, P. J. Ethical issues in early diagnosis and prevention of Alzheimer disease / P. J. Whitehouse // *Dialogues in Clinical Neuroscience*. – 2019. – № 21 (1). – P. 101–108.
2. Giannakopoulou, K. M. Internet of things technologies and machine learning methods for Parkinson's disease diagnosis, monitoring and management: a systematic review / K. M. Giannakopoulou, I. Roussaki, K. Demestichas // *Sensors*. – 2022. – № 22 (5). – 1799 p.
3. Vishniakou, U. A. IT Diagnostics of Parkinson's Disease Based on the Analysis of Voice Markers and Machine Learning / U. A. Vishniakou, X. YiWei // *Doklady BGUIR*. – 2023. – № 21 (3). – 102–110.
4. Upadhyay, N. Speech enhancement using spectral subtraction-type algorithms: A comparison and simulation study / N. Upadhyay, A. Karmakar // *Procedia Computer Science*. – 2015. – № 54. – P. 574–584.
5. Dendrinos, M. Speech enhancement from noise: A regenerative approach / M. Dendrinos, S. Bakamidis, G. Carayannis // *Speech Communication*. – 1991. – № 10 (1). – P. 45–57.
6. Al-Barhan, H. A. Modified speech separation deep learning network based on Hamming window / H. A. Al-Barhan, // In : *IOP Conference Series : Materials Science and Engineering*. – 2021. – Vol. 1076 (1).
7. Little, M. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection / M. Little [et al.] // *Nature Precedings*. – 2007. – P. 1–11.
8. Pasha, S. ThingSpeak based sensing and monitoring system for IoT with Matlab Analysis / S. Pasha // *International Journal of New Technology and Research (IJNTR)*. – 2016. – № 2 (6). – P. 19–23.
9. Sainath, T. N. Convolutional, long short-term memory, fully connected deep neural networks / T. N. Sainath [et al.] // In: *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE. – 2015. – P. 1–5.
10. Little, M. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease / M. Little // *Nature Precedings*. – 2008. – P. 1–27.
11. Parkinson Data Set [Electronic resource]. – Mode of access : <https://archive.ics.uci.edu/ml/datasets/parkinsons>. – Date of access : 12.9.2023.
12. Sakar, B. E. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings / B. E. Sakar [et al.] // *IEEE Journal of Biomedical and Health Informatics*. – 2013. – Vol. 17 (4). – P. 828–834.