

УДК 004.8

**Яшкін Р.І.**, аспірант спеціальності 122 Комп'ютерні науки  
Науковий керівник: **Бердник М.Г.**, д.т.н., доцент, професор кафедри програмного забезпечення комп'ютерних систем  
(Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна)

## ВИКОРИСТАННЯ ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ МОВИ ДЛЯ ДІАГНОСТИКИ ЗАХВОРЮВАНЬ

Під час діагностики неврологічних та психічних захворювань, що виникають внаслідок нейродегенеративних змін у головному мозку, лікарі нерідко визначають ступінь порушення мовлення пацієнта. Для того, щоб визначити такі характеристики мовлення, як моногучність, монотонність, слабкість голосу, тощо, потрібно багато часу. Інтелектуальні технології розпізнавання мови надають унікальні можливості для автоматизації цього процесу та підвищення його точності.

За допомогою обробки мови та аналізу контексту, системи в основі яких лежить штучний інтелект (ШІ), можуть виявляти не лише фізичні, а й психологічні порушення у пацієнтів. Дослідження, що були сфокусовані на аналізі мовлення та його застосуванні для виявлення депресії [1], продемонстрували успішне застосування ШІ для виявлення ознак депресії у пацієнтів.

Мовлення людини можна розглядати як послідовність звукових сигналів, які можна математично описати за допомогою сигнальної обробки. Одним із розповсюджених методів цифрової обробки сигналів є дискретне перетворення Фур'є (ДПФ). ДПФ дозволяє розкласти аудіо сигнал на його складові частоти. Це означає, що ДПФ може представити складний сигнал, в нашому випадку мовлення людини, у вигляді спектру частот, де кожна частота представлена амплітудою [2]. Математично це може бути виражено наступною формулою (1):

$$X[k] = \sum_{n=0}^{N-1} (x[n] e^{-i\frac{2\pi}{N}nk}), \quad (1)$$

де  $X(k)$  - комплексний амплітудний спектр сигналу;  $x(n)$  - вхідний сигнал;  $N$  - кількість відліків у вхідному сигналі;  $i$  - уявна одиниця ( $i^2 = -1$ );  $k$  - частотний індекс, представляє частоту в дискретному домені;  $n$  - часовий індекс, що представляє часові проміжки в сигналі.

Таким чином ДПФ може бути використане для багатьох завдань, наприклад:

- 1) Розпізнавання мови. Може використовуватися для виділення фону мови;
- 2) Видалення шуму. Може використовуватися для видалення шуму з аудіо сигналу, наприклад, фонового шуму або шуму вітру;
- 3) Стиснення аудіо. Може використовуватися для стиснення аудіо сигналу, видаляючи частоти, які не чутні для людського вуха.

Окрім цього, для обробки мови можна використовувати такі методи, як токенізація та лематизація, для розбиття тексту на окремі слова (що можуть бути представлені у різних відмінках) та приведення їх до єдиної форми чи відмінку. Наприклад, застосування алгоритмів лематизації може перетворити слова "ходінням", "ходінню" у "ходіння", що, в свою чергу, дозволяє зменшити розмірність даних та полегшити їх подальший аналіз.

Отже, математичне моделювання сигналів мовлення за допомогою сигнальної обробки надає розширені можливості для об'єктивного аналізу характеристик мовлення та розробки ефективних методів діагностики захворювань на його основі.

Дослідження, проведені А. Ідрисоглу та його колегами, в області використання машинного навчання для аналізу мовлення у діагностиці захворювань [3] показали, що за допомогою алгоритмів машинного навчання (МН) можна ефективно аналізувати параметри мовлення, такі як тон, висота, резонанс і тривалість тощо, і використовувати цю інформацію для виявлення відхилень, пов'язаних з захворюваннями, що впливають на артикуляцію. Дослідження також підтвердили, що моделі МН можуть забезпечити достатньо високу точність в розпізнаванні мови, для таких захворювань як хвороба Паркінсона або рак гортані, адже такі моделі тренують на великих, структурованих даних. Це робить моделі МН ефективним, допоміжним інструментом для діагностики захворювань.

Математичне моделювання акустичних характеристик мовлення, таких як спектральні властивості звукових сигналів, частотні характеристики та інтенсивність звукових хвиль є одним із фізичних аспектів цього дослідження. Використання математичних моделей дозволяє аналізувати звукові сигнали з точки зору фізичних параметрів [4], і дозволяє більш об'єктивно оцінювати стан мовлення пацієнтів.

Отже, переваги використання ШІ та МН для розпізнавання мови можуть значно покращити процес діагностики захворювань, забезпечити швидший та більш точний аналіз тих захворювань, що впливають на мовлення.

#### Список використаних джерел:

1. Zhang, L., Li, S., & Wang, C. (2020). Speech Analysis and Its Application for Depression Detection: Perspectives and Challenges. *Journal of Psychiatry and Neurology*. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8514878/pdf/fnhum-15-713823.pdf>
2. Medium. Frequency analysis of audio signals with Fourier transform. URL: <https://medium.com/@ongzhixuan/frequency-analysis-of-audio-signals-with-fourier-transform-f89ac113a2b4> (дата звернення: 10.02.2023).
3. Idrisoglu, A., Moraes, A. L., Anderberg, P., & Berglund, J. (2023). Applied Machine Learning Techniques to Diagnose Voice-Affecting Conditions and Disorders: Systematic Literature Review. *Journal of Medical Internet Research*, 25, e46105. DOI: 10.2196/46105
4. Rabiner, L. R., & Juang, B. H. (1993). *Fundamentals of speech recognition*. PTR Prentice Hall.