



24@KSMM'2023pb Wykrywanie zaburzeń psychicznych na podstawie mowy

Anna Bączyk, Marina Galanina, Anna Rekiel, Samuel Szurman
Opiekun: prof. dr hab. inż. Bożena Kostek

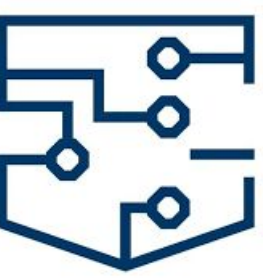


- **Anna Bączyk** - kierownik zespołu, organizacja pracy, ekstrakcja cech sygnału
- **Marina Galanina, Anna Rekiel** - preprocessing sygnału mowy, ekstrakcja cech sygnału
- **Samuel Szurman** - implementacja algorytmów klasyfikacji



Plan prezentacji

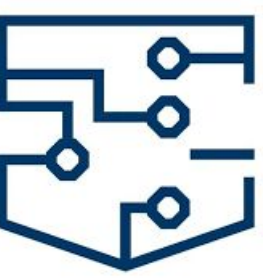
1. Cel projektu
2. Plan projektu
3. Wykorzystane bazy danych
4. Ekstrakcja cech
5. Algorytmy
6. Napotkane problemy
7. Plany na przyszłość



Celem projektu badawczego jest analiza powiązań pomiędzy próbką mowy a związkiem z chorobami psychicznymi (depresja).

Wśród przewidzianych zadań jest analiza sygnału mowy (oraz mowy) osób ze zdiagnozowanymi zaburzeniami psychicznymi przy użyciu przetwarzania sygnału (oraz analizy mowy).

W dalszej części przewidziana jest ekstrakcja cech związanych z sygnałem mowy i wykorzystanie algorytmów uczących się.



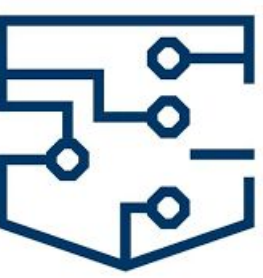
- 1. Przegląd baz danych**
- 2. Przeprowadzenie badań literaturowych**
- 3. Analiza innych projektów związanych z badaniem**
- 4. Preprocessing danych i ekstrakcja cech**
- 5. Zaplanowanie eksperymentów**
- 6. Stworzenie architektury modelu**
- 7. Trening modeli i dopasowywanie parametrów**
- 8. Wykonanie eksperymentów i ulepszanie algorytmu**
- 9. Analiza osiągniętych rezultatów**



DAIC-WOZ:


- Stworzony w 2014 roku,
- Język: angielski,
- Wywiady zebrane za pomocą wirtualnego asystenta Ellie,
- 189 folderów: transkrypcja+cechy twarzy+audio,
- Średnia długość nagrań: 16 minut,
- Używa PHQ8,

```
Pack\  
  
    300_P  
    301_P  
    ...  
    492_P  
    util  
    documents  
    train_split.csv  
    dev_split.csv  
    test_split.csv
```



EATD:

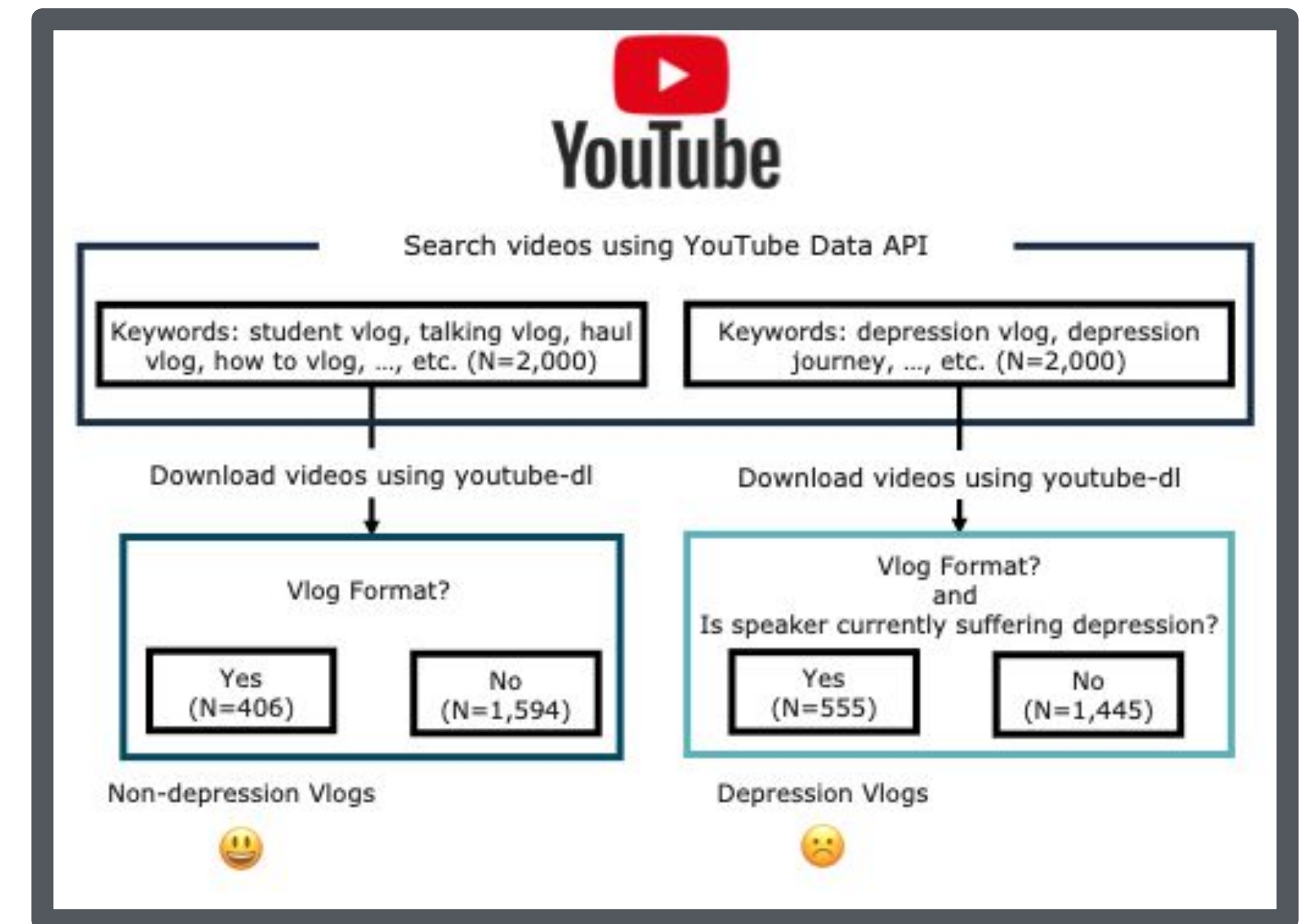
- Stworzony w 2022 roku,
- Język: chiński
- 162 foldery: transkrypcja+audio,
- Całkowita długość nagrań: 2.5 godziny,
- Row SDS*1.25 .

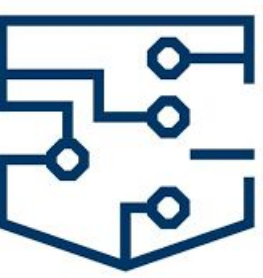
 t_1	Folder plików
 t_2	Folder plików
 t_3	Folder plików
 t_5	Folder plików
 t_6	Folder plików
 t_8	Folder plików
 t_9	Folder plików



D-Vlog:

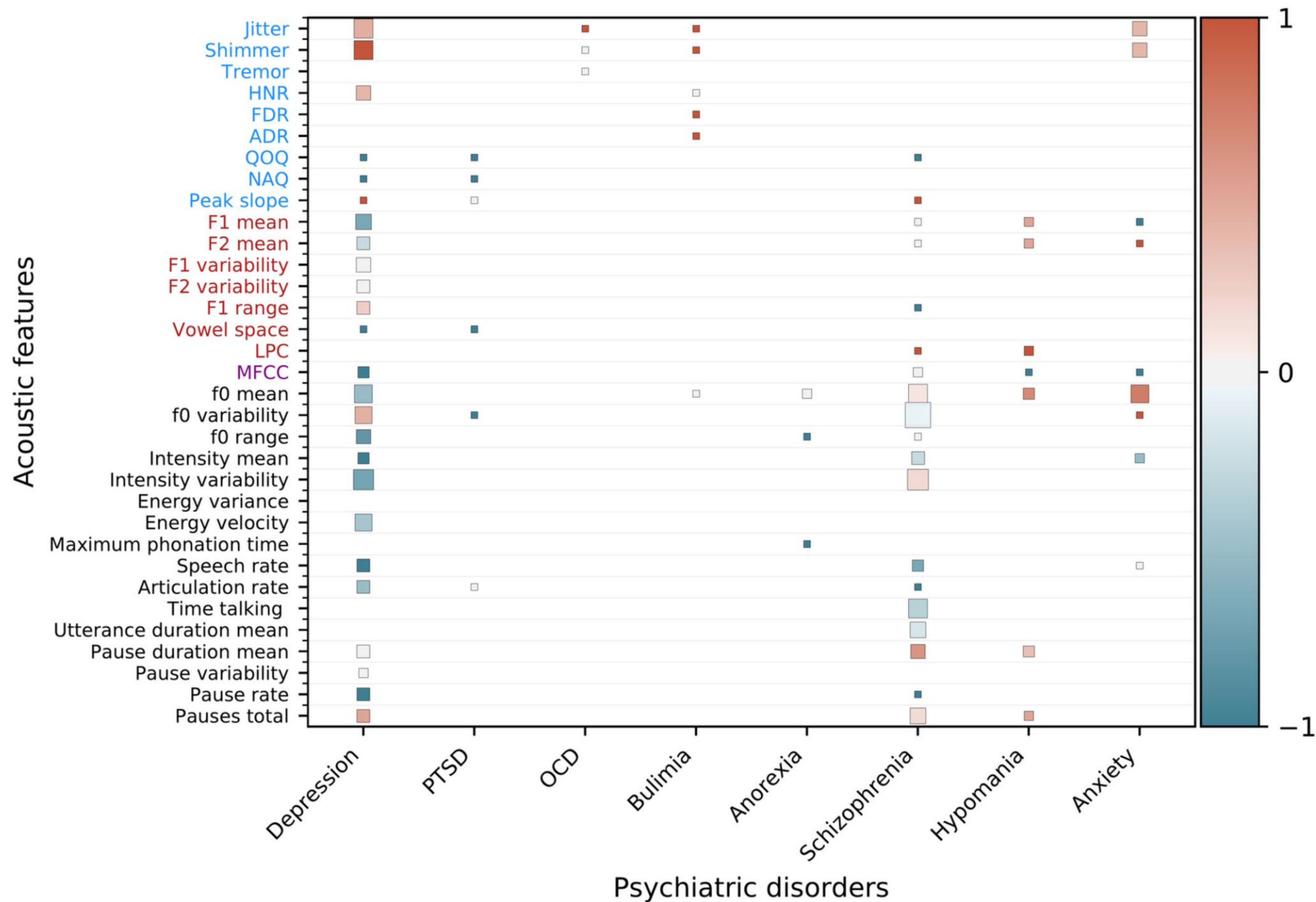
- Stworzony w 2022 roku,
- Język: angielski
- Wykrywanie depresji na podstawie vlogów,
- 961 vlog,
- Używa “słowa-klucze”,
- Zawiera już wyekstraktowane parametry.



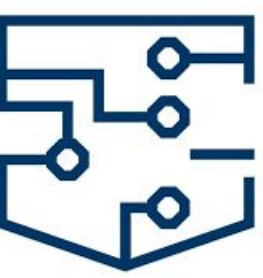


EMU:

- Stworzony w 2021 roku,
- Język: angielski
- Nagrania od 70 osób,
- Używa PH9,
- Zawiera już wyekstraktowane parametry.



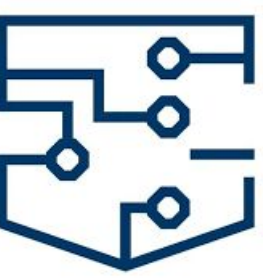
Rys. 1. Korelacja między zaburzeniami psychicznymi a cechami dźwięku [1]



Zastosowaliśmy następujące parametry:

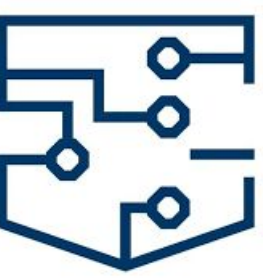
- 40 współczynników mel-cepstralnych
- Wartość skuteczną widma
- Częstotliwość fundamentalną
- Częstotliwości formantów
- Prędkość mowy
- Środek ciężkości widma
- Płaskość widmową
- Parametr opadania widma
- Jitter
- Shimmer

W dalszych etapach planowana jest selekcja cech oraz redukcja wymiarów przy wykorzystaniu metody PCA.



Algorytmy do przetestowania (podane razem z dokładnością) z artykułów naukowych, wytrenowanych na bazie DAIC-WOZ :

- liniowy klasyfikator SVM - 94,3% [2]
- ANN z 2 warstwami ukrytymi po 100 neuronów każda - 98,3% [3]
- 1D CNN z augmentacją danych - 95% [4]
- ResNet-50 - 77% [5]



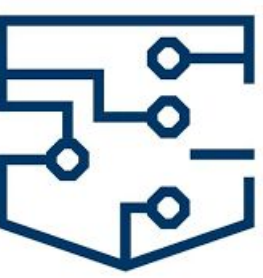
Napotkane problemy:

- problem z dostępnością korpusów
- brak korpusów w języku polskim



Plany na przyszłość:

- Rozwój aplikacji diagnostycznej
- Analiza mowy w innych językach
- Analiza języka naturalnego
- Współpraca z lekarzem psychiatrą



[1] Low, D. M., Bentley, K. H., Ghosh, S. S. Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*. 2020; 5: 96–116.

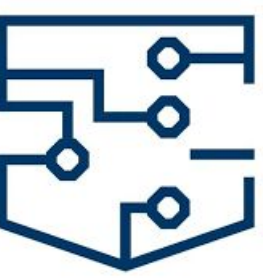
<https://doi.org/10.1002/lio2.354>

[2] Z. Huang, J. Epps and D. Joachim, "Investigation of Speech Landmark Patterns for Depression Detection," in *IEEE Transactions on Affective Computing*, vol. 13, no. 2, pp. 666-679, 1 April-June 2022,

<https://doi.org/10.1109/TAFFC.2019.2944380>

[3] Alghifari M. F., Gunawan T. S., Nordin M. A. W. , Kartiwi M., Borhan L., On the Optimum Speech Segment Length for Depression Detection, 2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), Kuala Lumpur, Malaysia, 2019, 1-5,

<https://doi.org/10.1109/ICSIMA47653.2019.9057319>



[4] Homsiang P. , Treebupachatsakul T. , Kiatrungrit K., Poomrittigul S. ,
Classification of Depression Audio Data by Deep Learning,2022 14th
Biomedical Engineering International Conference (BMEiCON), Songkhla,
Thailand, 2022, 1-4,

<https://doi.org/10.1109/BMEiCON56653.2022.10012102>

[5] Chlasta K., Wołk K., Krejtz I., Automated speech-based screening of
depression using deep convolutional neural networks, Procedia Computer
Science, 2019, 164, 618-628, <https://doi.org/10.1016/j.procs.2019.12.228>



Dziękujemy za uwagę!