

Rozpoznawanie człowieka metodami biometrii

Projekt 2. — Rozpoznawanie na podstawie głosu

Raport

Bartłomiej Dach

11 kwietnia 2019

Poniższy dokument stanowi sprawozdanie z implementacji aplikacji dokonującej rozpoznawania człowieka na podstawie zarejestrowanych próbek głosu. W dokumencie opisano zastosowaną metodę opartą na współczynnikach mel-cepstralnych oraz zawarto wyniki działania dla próbek zarejestrowanych przez studentów uczęszczających na zajęcia.

1 Wstęp

Ludzki głos jest cechą biometryczną z pogranicza cech biologicznych i behawioralnych. Podczas gdy barwa głosu jest determinowana przez wrodzone czynniki anatomiczne, ton głosu i akcent stosowany podczas wymawiania określonych fraz stanowi cechę nabytą podczas nauki mowy.

Rozpoznawanie człowieka na podstawie mowy to prężnie rozwijające się zagadnienie. Ze względu na rosnącą popularność urządzeń interpretujących frazy wymawiane przez użytkowników takich, jak Google Home czy Amazon Echo, identyfikacja na podstawie głosu stanowi ważny kierunek rozwoju.

W ramach projektu zaimplementowano prostą metodę klasyfikacji próbek na podstawie danego wcześniej zbioru treningowego, używającą współczynników mel-cepstralnych. Szczegółowy opis metody znajduje się w sekcji 3.

Nr	Nazwa	Opis	Licencja	
1	matplotlib 3.0.3	Tworzenie wykresów i wizualizacji	PSF	[1]
2	numpy 1.16.2	Wielowymiarowe tablicowe struktury danych	BSD	[4]
3	pandas 0.24.2	Struktury do manipulacji i analizy danych	BSD	[3]
4	seaborn 0.9.0	Rozszerzone wizualizacje danych	BSD	[5]
5	scipy 1.2.1	Algorytmy pomocnicze (transformata Fouriera, manipulacja dźwiękiem)	BSD	[2]

Tablica 1: Lista bibliotek użytych w projekcie

2 Opis aplikacji

2.1 Zastosowane biblioteki

2.2 Instrukcja obsługi

3 Opis metody

3.1 Wyznaczanie współczynników mel-cepstralnych

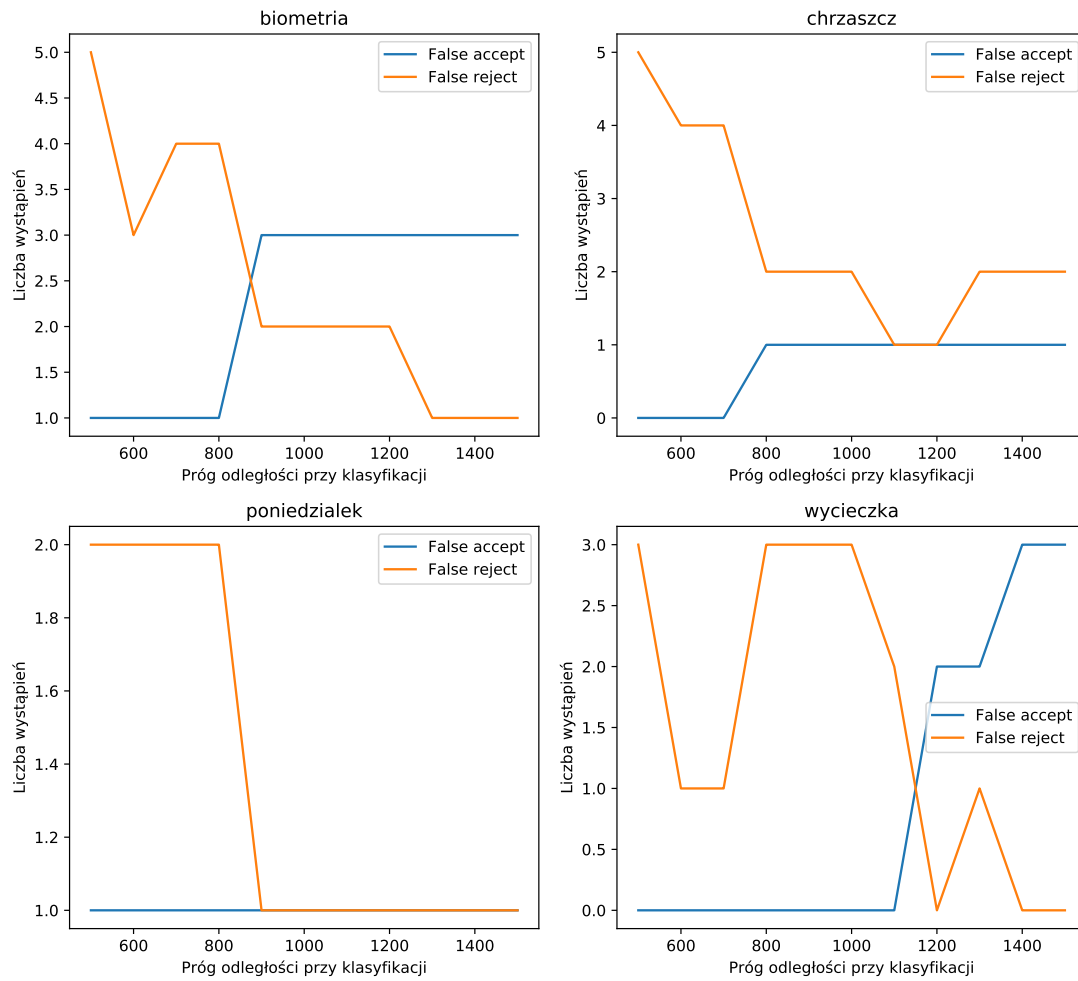
3.2 Klasyfikacja nowych próbek

4 Wyniki eksperymentalne

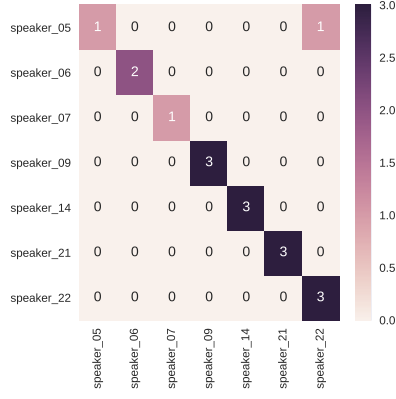
5 Podsumowanie

Literatura

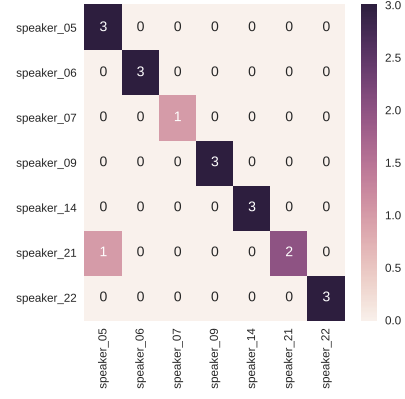
- [1] Hunter, J.D., „Matplotlib: A 2D graphics environment”, *Computing In Science & Engineering*, tom 9, nr 3, s. 90–95, 2007.
- [2] Jones, E., Oliphant T.E., Peterson P. i inni, „SciPy: Open source scientific tools for Python”. [Online]
Dostępne: <https://www.scipy.org/>. [Dostęp 7 kwietnia 2019]
- [3] McKinney, W., „Data Structures for Statistical Computing in Python”, *Proceedings of the 9th Python in Science Conference*, s. 51–56, 2010.
- [4] Oliphant, T.E., *A Guide to NumPy*, Trelgol Publishing, Stany Zjednoczone, 2006.
- [5] Waskom, M. i inni, „seaborn: statistical data visualization”. [Online]
Dostępne: <https://seaborn.pydata.org/>. [Dostęp 7 kwietnia 2019]



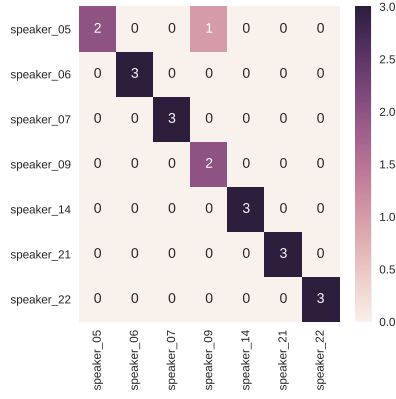
Rysunek 1: Liczba fałszywych pozytywów (ang. *false accept*) i fałszywych odrzuceń (ang. *false reject*) próbek głosów z testowanego zbioru w zależności od przyjętego progu odległości między próbkami podczas klasyfikacji. Oddzielono wyniki dla każdej z czterech rejestrowanych fraz.



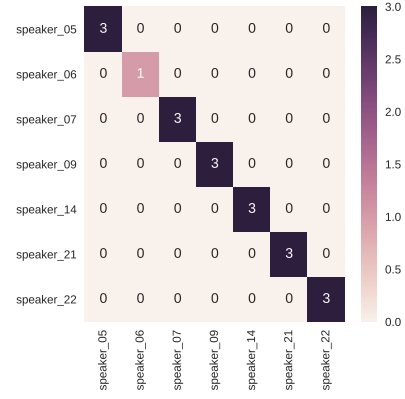
(a) Macierz pomyłek dla słowa *biometria* przy progu $t = 800$.



(b) Macierz pomyłek dla słowa *chrząszcz* przy progu $t = 800$.



(c) Macierz pomyłek dla słowa *poniedziałek* przy progu $t = 900$.



(d) Macierz pomyłek dla zdania *Jutro pojedę na wycieczkę, albo zostanę w domu* przy progu $t = 1100$.

Rysunek 2: Macierze pomyłek dla wartości progu minimalizujących liczbę fałszywych pozytywów i negatywów dla poszczególnych zarejestrowanych fraz.