

Sprawozdanie: Rozpoznawanie płci mówcy na nagraniu głosowym

Marek Kamiński 136732

Grzegorz Bryk 136686

Zadanie polegało na stworzeniu w Pythonie programu rozpoznającego płeć mówcy na nagraniu dostarczonym w formacie WAVE. W tym celu przygotowaliśmy program, który:

- wczytuje zbiór plików WAVE
- filtruje wczytane pliki filtrem band-pass Butterwortha w zakresie 80Hz-3000Hz
- obkłada dane „Oknem Kaisera” z $\beta=15$
- obkłada dane dźwiękowe Transformatą Fouriera
- w zakresie od 65Hz do 300Hz szuka częstotliwości podstawowej mowy
- znalezionej częstotliwości podstawową, wraz z wartością trzeciego kwartyłu z zakresu 80-3000Hz wykorzystuje jako wejście do regresji logistycznej.

Training loss regresji logistycznej wynosi w naszej implementacji 0,0(3)%. Macierz błędów przedstawia się następująco:

Wartość prawdziwa	Wartość przewidziana	
	Mężczyzna	Kobieta
Mężczyzna	44	1
Kobieta	2	43

Poniżej wykres przestrzeni danych wejściowych regresji logistycznej wraz z jej rozwiązaniem:

