



**Politechnika
Śląska**

**Wydział Automatyki, Elektroniki
i Informatyki**

Systemy Interaktywne i Multimedialne
Projekt
Detekcja emocji w głosie

Natalia Stręk, Jakub Kula, Paweł Wójtowicz

Gliwice 2023

Spis treści

1	Wizja projektu	2
2	Cele projektu	2
3	Zakres projektu	2
4	Etapy w projekcie	2
5	Charakterystykę narzędzi	3

1 Wizja projektu

DO TO NAPISAĆ TO!

- prexyxyjna detekcja $acc_i=60\%$
- efektywność w czasie rzeczywistym

2 Cele projektu

- Opracowanie modelu uczenia maszynowego, który będzie potrafił rozpoznawać emocje w nagraniach głosowych.
- Stworzenie interaktywnego narzędzia umożliwiającego przesłanie nagrania głosowego i otrzymanie informacji zwrotnej na temat emocji wykrytych w głosie.

3 Zakres projektu

DO TO NAPISAĆ TO! Zakres projektu będzie obejmować wybór odpowiednich danych. Dane muszą zawierać odpowiednią ilość nagrań oraz obejmować odpowiednią ilość emocji. Następnym krokiem będzie wybór odpowiednich cech które posłużą do nauczania modelu. Po wyborze odpowiednich cech zostanie sporządzony skrypt przetwarzające wybrane dane. Kolejnym etapem będą badania nad achiterkórą i zestawem hisperparamterów które zapewnią najwyższy współcznnik dokładności. Następnie zostanie wybrana o

Zakres definiuje granice projektu przedstawiając co zostanie wykonane w ramach projektu, a co nie. To co zostanie wykonane powinno zostać ujęte za pomocą konkretnych, mierzalnych rezultatów. (ok 200 słów)

4 Etapy w projekcie

- Znalezienie nagrań głosowych przedstawiające różne emocje.
- Reserach na temat sygnałów dźwiękowych i ich cech
- Wybór cech które posłużą do nauki modelu uczenia maszynowego
- Przetworzenie nagrań głosowych, wyciąganie z nich cech charakterystycznych
- wybór metody krosvalidacji
- Strojenie i wybór architektury sieci oraz jej hisperparamterów przy użyciu wyników z krosvalidacji
- Stworzenie UI pozwalającego użytkownikowi dodanie pliku dźwiękowego oraz nagrania własnej wiadomości przy pomocy wbudowanego mikrofonu
- Testowanie projektu

5 Charakterystykę narzędzi

- Python - główny język programowania. W nim zostaną napisane skrypty przetwarzające dane oraz zostanie stworzony i nauczony model sieci neuronowe.
- Tensorflow i keras - Tensorflow będąca najpopularniejszą biblioteką uczenia maszynowego stanowi główne narzędzie do tworzenia sieci neuronowych. Keras udostępnia API do tworzenia modeli uczenia maszynowego.
- Numpy i Pandas - fundamentalne biblioteki do obliczeń naukowych w pythonie, umożliwiające efektywne przetwarzanie dużych ilości danych oraz ich wnikliwą analizę
- Librosa - biblioteka pozwalająca w łatwy sposób otworzyć i przetworzyć dane audio.
- Scikit-learn, Eli5 - biblioteki zawierające między innymi bogaty zbiór algorytmów uczenia maszynowego. Pozwalające na uproszczenie etapu tworzenia sieci.
- Conda - Conda jest środowiskiem wirtualnym i systemem do zarządzania pakietami. Pozwala ona na tworzenie odseparowanych środowisk dla różnych projektów.
- Tkinter - biblioteka pozwalająca nam uprościć proces tworzenia działającego interfejsu graficznego.