

FourthBrain

Spotify - Wake Word Detection

Swetha Bezawada, Tiam Moradi, Mani Narasimhan



Agenda

- Problem
- Solution
- Data + Model
- MLE Stack
- EDA



Problem

- Advancement in technology towards automatic speech recognition has increased the demand for virtual voice assistants and similar devices.
- This market growth requires that the devices detect the wake word accurately
- In addition, the challenge is to find out if a specific wake word is required or can it be customized?
- Our project aims to solve these questions by developing and deploying a machine learning model that can classify a custom wake word.



Solution

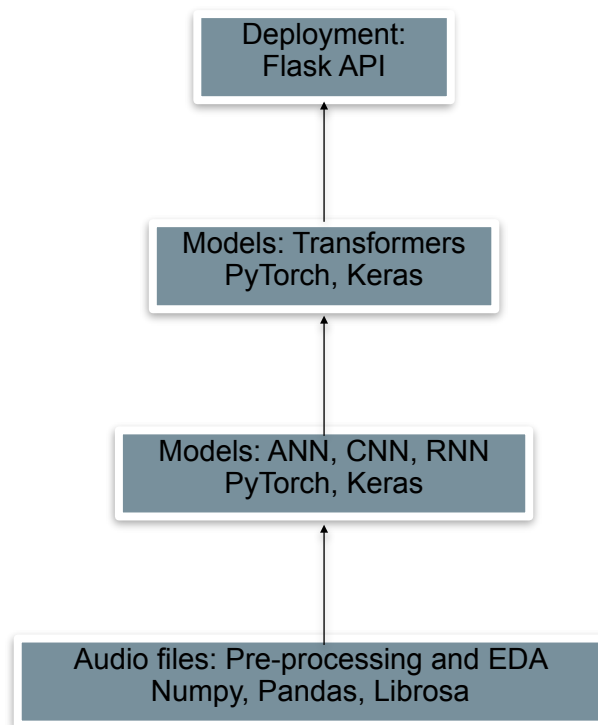
- Design a custom wake word - “Hey! Fourthbrain”, gather input data under diverse scenarios that include and exclude the wake word.
- Develop and train models using ANN, CNN and RNN architectures to effectively detect the wake word.
- Evaluate the models on different criteria such as accuracy, precision, recall, F1-score and AUC score.
- For a chosen sample of successful classifications generate search results for the corresponding text.
- Deploy the most effective model as a web application.



Data + Model

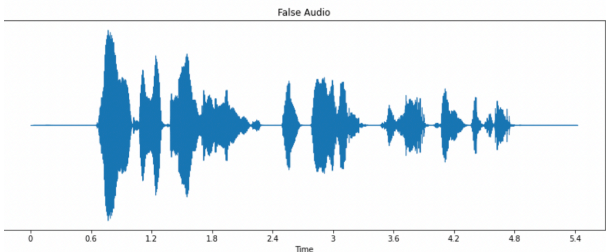
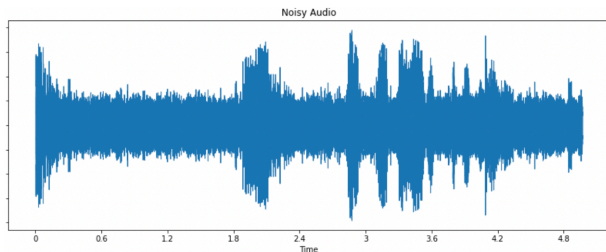
- Input data: We will leverage data in the public domain such as Mozilla common voice, Google speech commands and also generate custom data both including and excluding the wake word.
- Pre-process and do EDA using audio libraries such as Librosa.
- Models: ANN, CNN and RNN.
- Text to search: Generate search results using pre-trained Transformer models.
- Challenges:
 1. Input data to capture a diverse user base (age, gender, accent).
 2. Deploy the models with minimal latency to the end user

MLE Stack

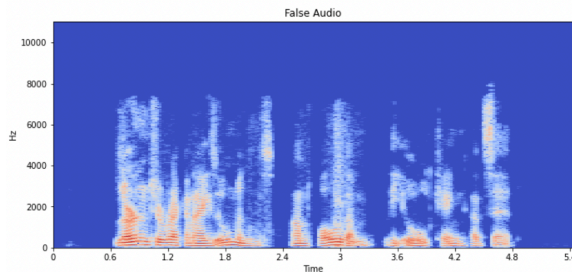
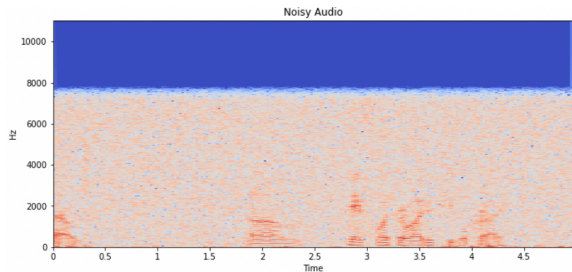




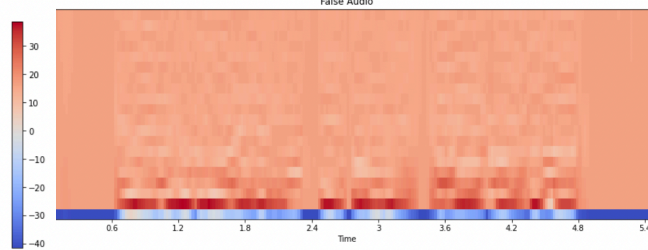
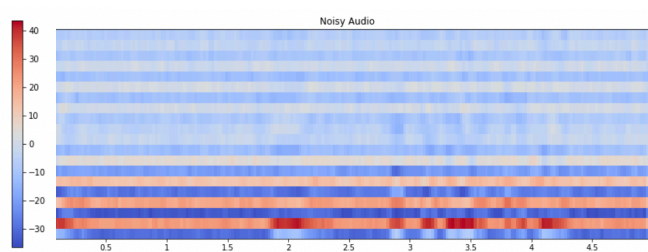
EDA



Audio Waves



Spectrogram



Mel Frequency Cepstral
Coefficients



Thank You! Questions?