Pay Attention to the cough

Early Diagnosis of COVID-19 using Interpretable Symptoms Embeddings with Cough Sound Signal Processing

Ankit Pal (Aaditya Ura)
Research Engineer, Saama AI Research Lab
ankit.pal@saama.com

Malaikannan Sankarasubbu
VP, Saama AI Research Lab
malaikannan.sankarasubbu@saama.com

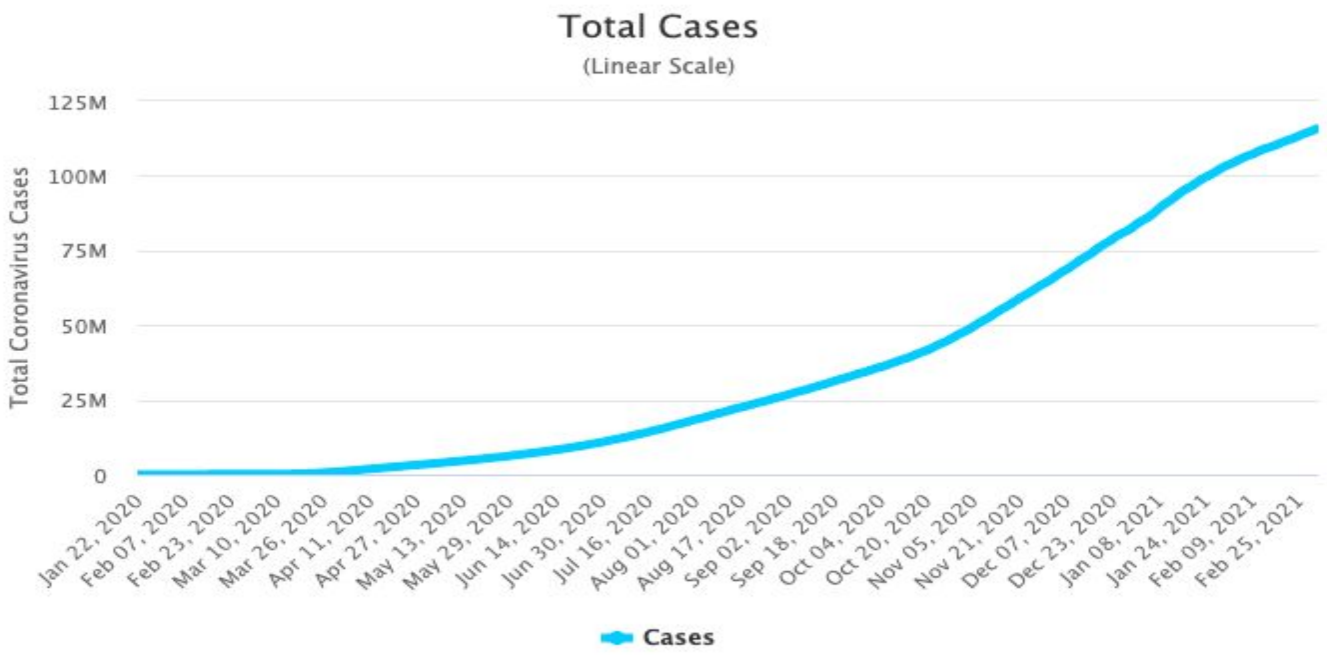
About Me

- Research Engineer @ Saama AI Research Lab
- Research interests involve
  Representation Learning on Graphs and Manifolds
- Interpretable Natural Language Processing,
  and their applications in Healthcare data
- Respiratory, Neurophysiological (EEG, ECG, EMG etc.) based Signal Processing

Email : ankit.pal@saama.com
Website : aadityaura.github.io
twitter : @aadityaura

COVID-19

The novel coronavirus (COVID-19) disease has affected over 113 million lives, claiming more than 2.5 million fatalities globally, representing an epoch-making global crisis in health care.



Outline

- **Introduction**
- **Problem Statement**
- **Motivation**

- **COVID-19 Data**
    - Data Collection, Statistics & Preprocessing

- **Feature Extraction**
    - Cough Features

- **Model Architecture**
    - Symptoms & Cough Embeddings, Final Model & loss

- **Experiments**
    - Results

- **Interpretability**
- **In depth Analysis**
- **Scope & Future work**

The Problem statement

The current diagnosis of COVID-19 is done by

- **Thermometer**
  - It does not give an accurate estimation of deep body temperature

- **Reverse-Transcription Polymer Chain Reaction (RT-PCR)**
  - Time-consuming, expensive, and not easily available in straitened regions

Motivation

- Lack of a fast and reliable testing method challenge
- A low-cost, rapid, easily accessible testing solution is required to increase diagnostic capability and devise a treatment plan
- It is an essential tool in the fight to slow and reduce the virus's spread and impact

Motivation

At the beginning of March 2020, we thought about a simple idea:

"Is it possible to do mass testing with current technology?"

COVID-19 Symptoms

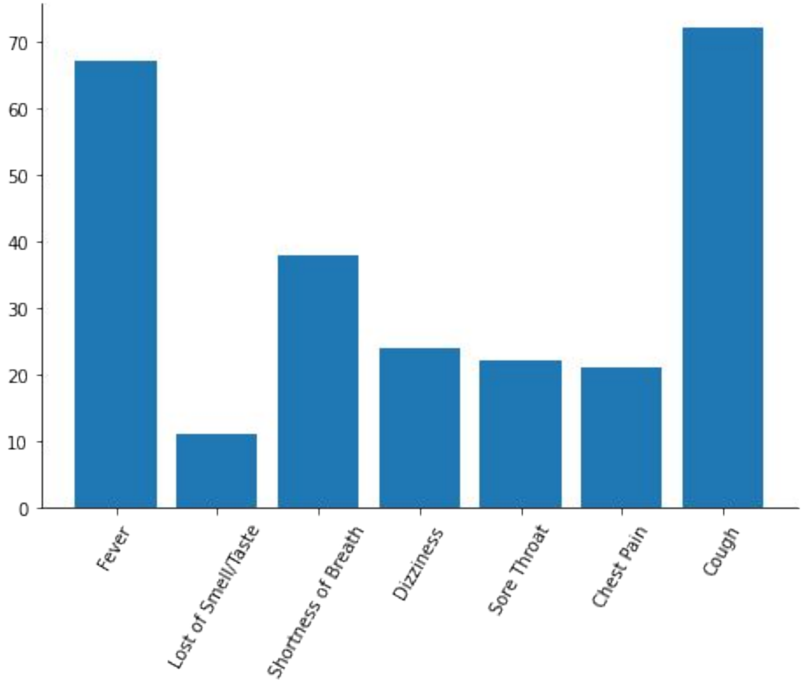
The main symptoms of COVID-19 given by WHO and CDC official report are

- **Fever**
- **Shortness of breathing**
- **COVID-19 Dry cough**

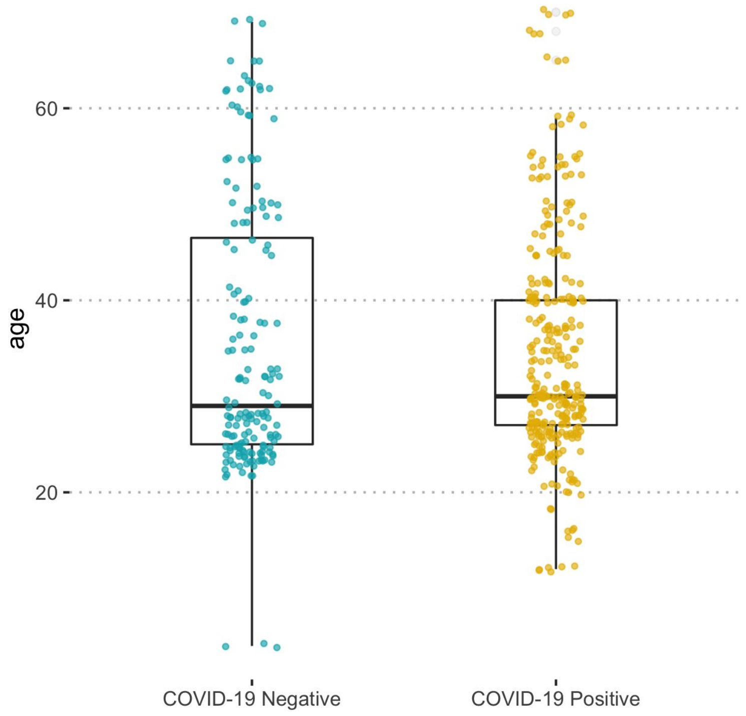
Data collection

- Data in this study were obtained from Dr. Ram Manohar Lohia Hospital, New Delhi, India
- Out of 100 were confirmed positive from COVID-19 RT-PCR results
- Bronchitis and Asthma cough samples were also collected from different online & offline sources
- Additional data
  - Breathing sounds
  - counting 1 to 10 (natural voice samples)
  - sustained phonation of ’a,’ ’e,’ ’o’ vowel

Data Statistics

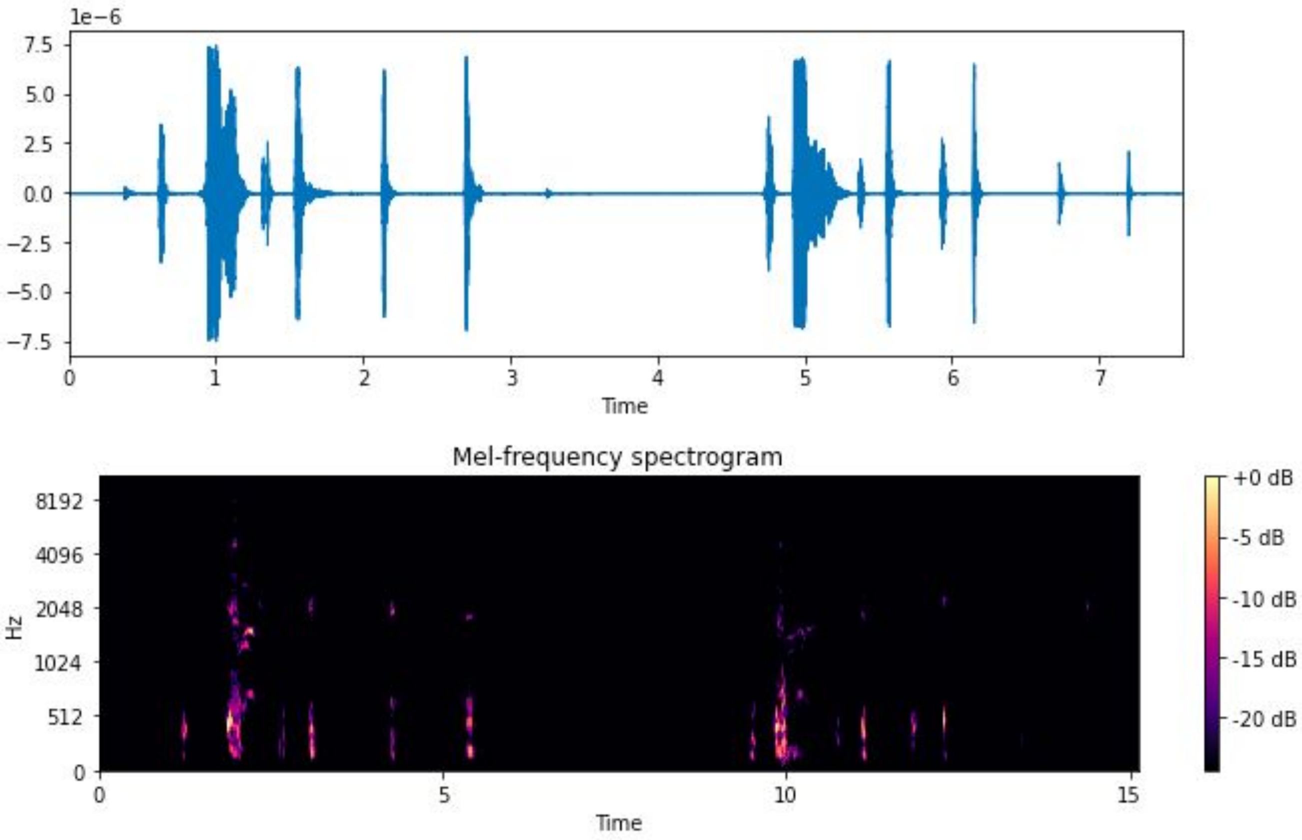


Symptoms count



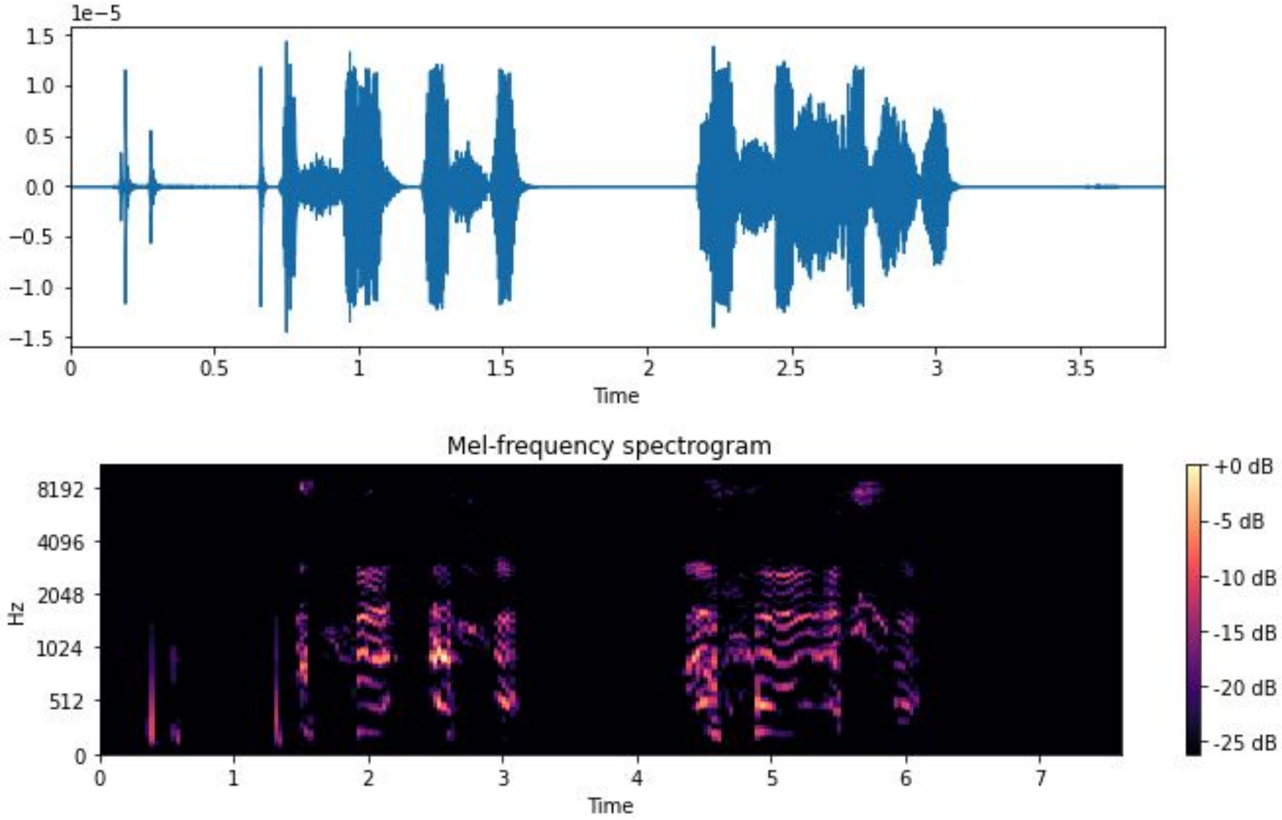
Age distribution

Healthy Cough



Random Sample of a Healthy Cough from collected dataset

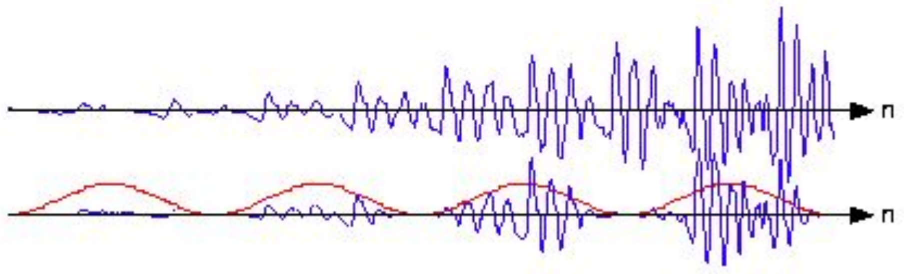
COVID-19 Cough



Random Sample of a COVID-19 Cough from collected dataset

Data Preprocessing

- Each cough recording was downsampled to 16 kHz
- Normalization was applied to the cough signal level with a target amplitude of -28.0 dBFS
- Normalized features were split based on the silence threshold
- A High Pass Filter(HPF) was applied to reduce the noise
- Divided into sub-segments of non-overlapping Hamming-windowed frames



An example of Non-Overlapping windows [1]

Feature Extraction

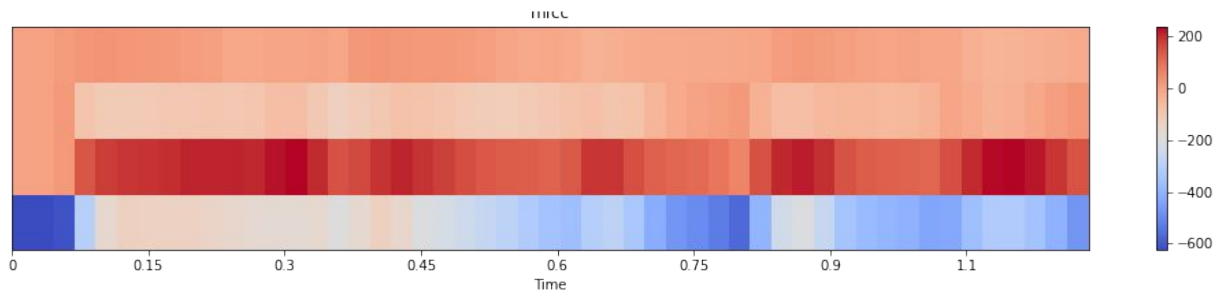
After Preprocessing, Two types of features were extracted :

- The cough Features
- Symptoms & Demographic features

Cough Features

For each Hamming-windowed frames, the following features were extracted

- **Mel Frequency Cepstral coefficients (MFCCs)**
    - The hearing mechanism of human beings inspires MFCC

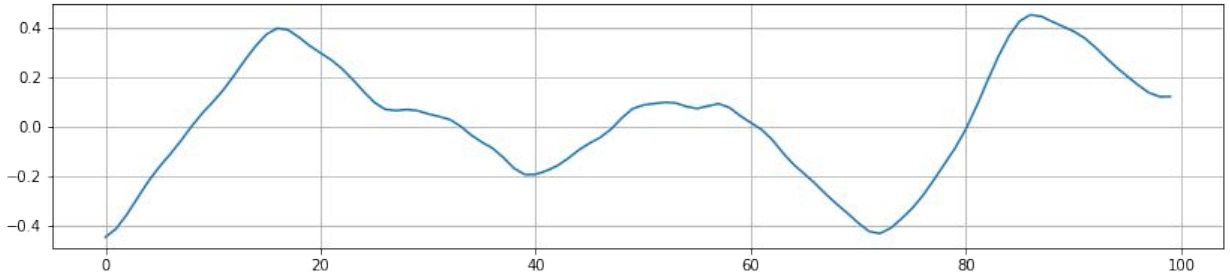


- **Skewness**
    - Measures the symmetry in a probability distribution

Cough Features

For each Hamming-windowed frames, the following features were extracted

- **Zero crossing rate(ZCR)**
    - ZCR is used to calculate the number of times a signal crosses the zero axis

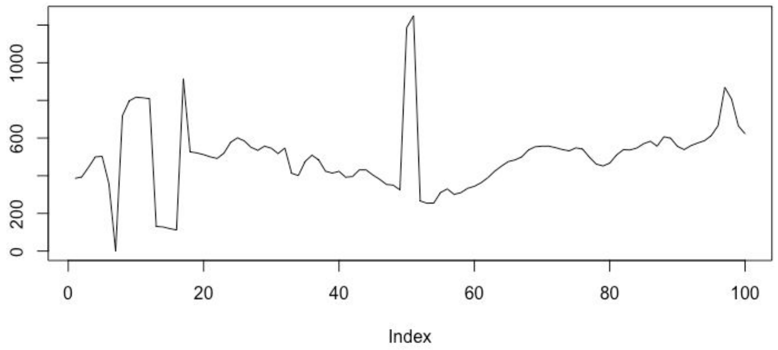


- **Entropy**
    - Capture the difference between signal energy distributions

Cough Features

For each Hamming-windowed frames, the following features were extracted

- **Formant frequencies**
    - Formant frequencies capture the vocal tract resonance characteristics

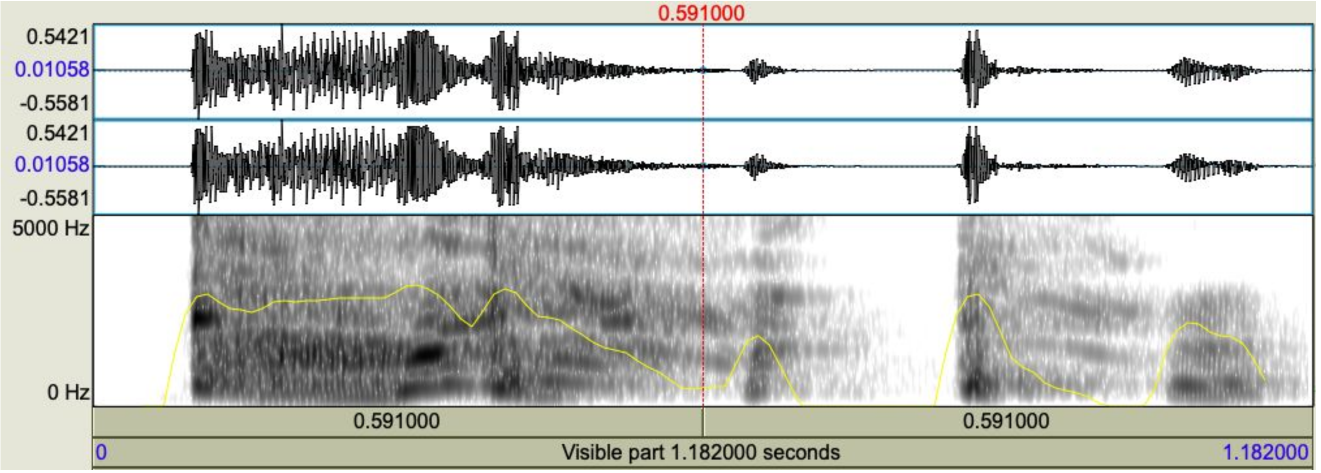


- **Kurtosis**
    - Measures the peakiness or heaviness associated with the cough sub-segment probability distribution

Cough Features

For each Hamming-windowed frames, the following features were extracted

- **Fundamental frequency(F0)**
  - F0 is the frequency at which vocal cords vibrate in voiced sounds



Model Architecture

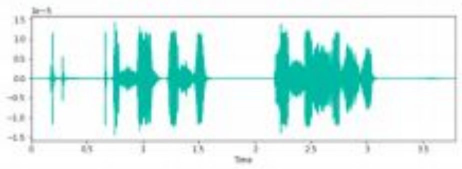
The model architecture consists of two subnetworks components

- Symptoms Embeddings

Fever	Chest pain	Dyspnea	Cough	Thyroid	...	Asthma
0	1	1	0	0	...	0
1	0	0	1	0	...	0
1	0	1	1	0	...	0
0	1	0	0	0	...	1

Symptoms and Demographic data

- Cough Embeddings



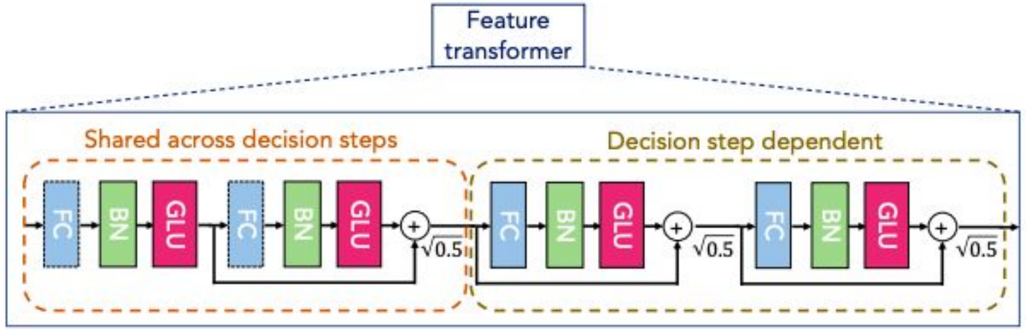
Patient Cough Audio

Symptoms Embeddings

- Symptoms Embeddings capture the hidden features of patient characteristics, diagnosis, symptoms
- We utilized the transformer-based tabular model called “TabNet,” which uses a sequential attention mechanism to generate the Symptoms embeddings
- It consists of four main components :
  - Gated Linear Unit
  - Feature Transformer
  - Attentive Transformer
  - Decision Step

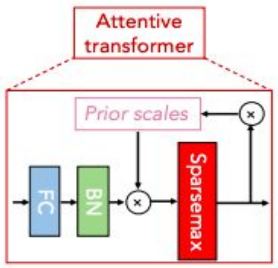
Symptoms Embeddings

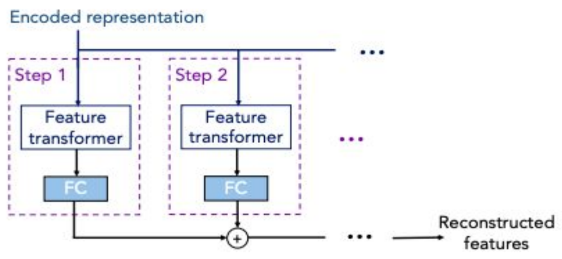
- **Gated Linear Unit (GLU)**
  - GLU functions like a filter; it regulates which parts of the signal should be allowed into the unit
- **Feature Transformer (FT)**
  - Feature Transformer process the filtered features by looking at all the symptoms features assessed and deciding which ones indicate which class



Symptoms Embeddings

- **Attentive Transformer (AT)**
  - Utilizes sparse intense wise features selection based on learned symptoms dataset and focusing on specific symptoms features only
- **Decision Step (DS)**
  - Decision steps are composed of a Feature Transformer(FT), Attentive Transformer(AT), and feature masking

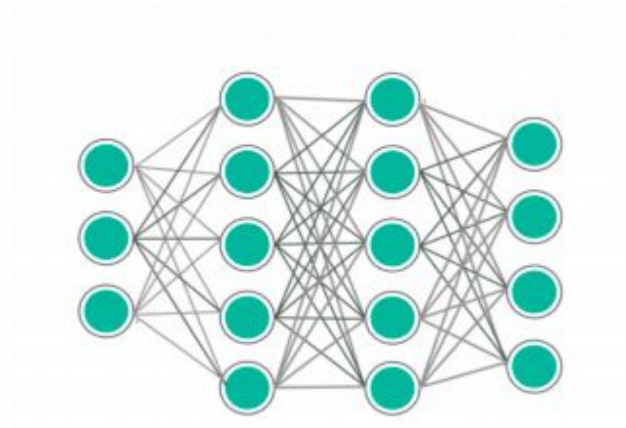




Cough Embeddings

- Cough Embeddings learn and capture more in-depth features in temporal acoustic characteristics of cough sounds

- We used Deep Neural Network (DNN) to learn the embeddings from cough features extracted in the previous feature engineering part

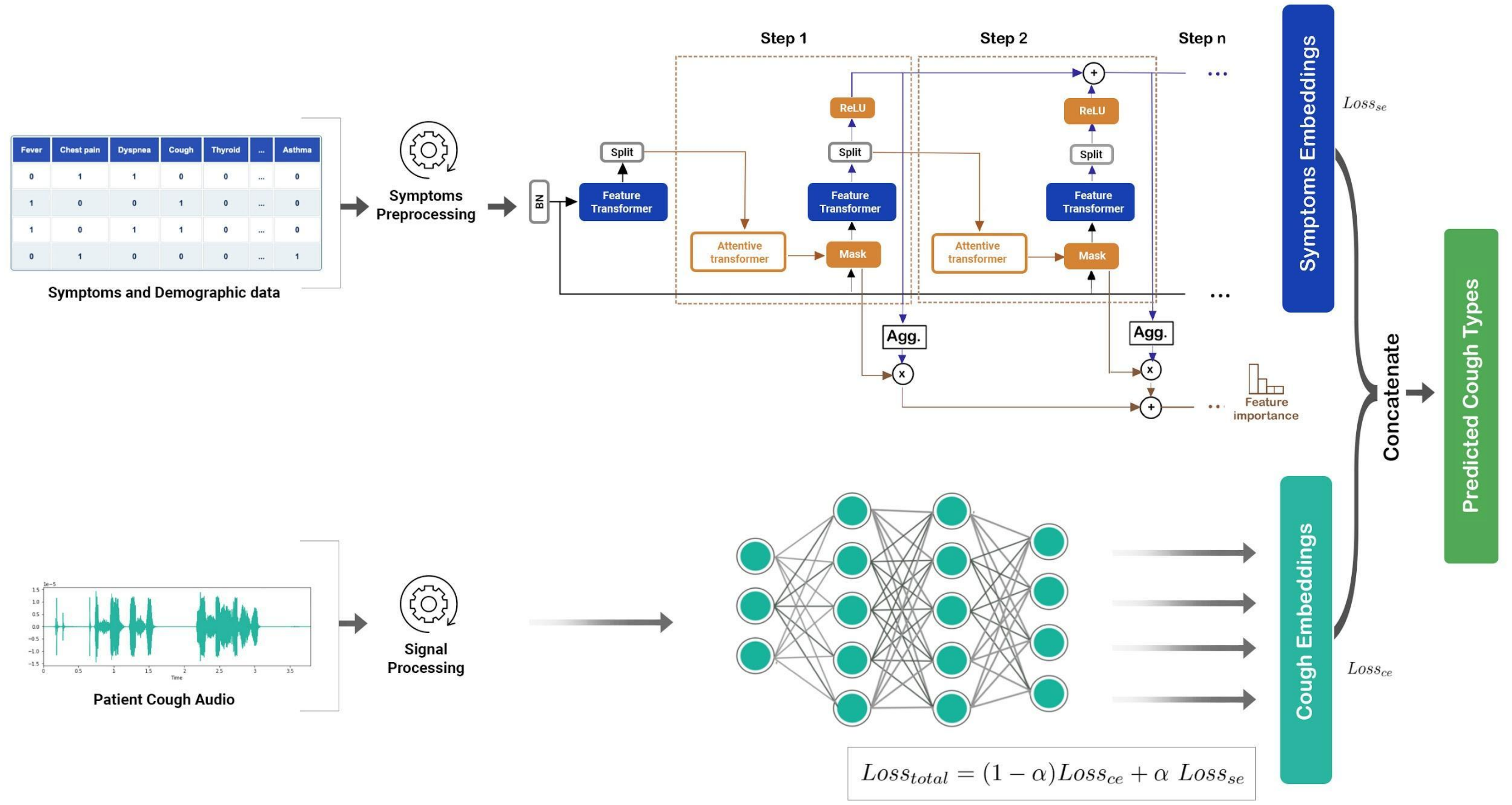


Final Model

Concatenating Symptoms Embeddings with Cough Embeddings followed by a Fully-Connected(FC) layer

$$\hat{y} = \underbrace{[S_e, C_e]}_{\text{Concatenate}} \cdot FC$$

Final Model



Fever
Chest pain
Dyspnea
Cough
Thyroid
...
Asthma
Symptoms and Demographic data
Symptoms Preprocessing
BN
Step 1
Step 2
Step n
Split
Feature Transformer
Attentive transformer
Mask
ReLU
Agg.
Feature importance
Symptoms Embeddings
$Loss_{se}$
Concatenate
Predicted Cough Types
Patient Cough Audio
Signal Processing
Cough Embeddings
$Loss_{ce}$
$Loss_{total} = (1-\alpha)Loss_{ce} + \alpha\ Loss_{se}$

Total Loss

After this, the Total loss was calculated as follows

$$Loss_{total} = \underbrace{(1-\alpha)\ Loss_{ce}}_{\text{Cough Embeddings Loss}} + \underbrace{\alpha\ Loss_{se}}_{\text{Symptoms Embeddings Loss}}$$

Where α is a small constant value to balance the contribution of the different losses.

Experiments

Based on the collected dataset, the model was trained on the following combination of features

- **Task 1**, Using cough data only
- **Task 2**, Using Demographic & Symptoms Data Only
- **Task 3**, Using Both
- **Task 4**, Using Both with different Cough Types

Results (Binary Classification : Task 1)

		F1-score	Precision	Sensitivity	Specificity	Accuracy
Cough data	Covid-19 Positive	$90.6 \pm 0.2\%$	$89.1 \pm 0.4\%$	$86.2 \pm 0.3\%$	$92.4 \pm 0.2\%$	$89.3 \pm 0.1\%$
	Covid-19 Negative	$90.6 \pm 0.1\%$	$91.7 \pm 0.1\%$	$94.3 \pm 0.3\%$	$89.3 \pm 0.1\%$	$92.4 \pm 0.1\%$
	Overall	$90.6 \pm 0.3\%$	$90.4 \pm 0.5\%$	$90.1 \pm 0.6\%$	$90.3 \pm 0.3\%$	$90.8 \pm 0.2\%$
Symptoms data	Covid-19 Positive	$91.5 \pm 0.2\%$	$86.9 \pm 0.5\%$	$87.8 \pm 0.2\%$	$86.0 \pm 0.3\%$	$94.1 \pm 0.6\%$
	Covid-19 Negative	$91.5 \pm 0.1\%$	$93.7 \pm 0.3\%$	$93.3 \pm 0.2\%$	$94.1 \pm 0.1\%$	$86.0 \pm 0.2\%$
	Overall	$91.5 \pm 0.3\%$	$90.3 \pm 0.8\%$	$90.5 \pm 0.4\%$	$90.8 \pm 0.3\%$	$91.1 \pm 0.8\%$
Both	Covid-19 Positive	$96.8 \pm 0.4\%$	$95.1 \pm 0.1\%$	$94.6 \pm 0.3\%$	$95.6 \pm 0.1\%$	$97.3 \pm 0.2\%$
	Covid-19 Negative	$96.8 \pm 0.1\%$	$97.6 \pm 0.3\%$	$97.8 \pm 0.4\%$	$97.3 \pm 0.2\%$	$95.6 \pm 0.3\%$
	Overall	$96.8 \pm 0.5\%$	$96.3 \pm 0.4\%$	$96.2 \pm 0.7\%$	$96.5 \pm 0.3\%$	$96.5 \pm 0.5\%$

Model performance metrics across different models on Task 1 Using cough data only

Results (Binary Classification : Task 2)

		F1-score	Precision	Sensitivity	Specificity	Accuracy
Cough data	Covid-19 Positive	$90.6 \pm 0.2\%$	$89.1 \pm 0.4\%$	$86.2 \pm 0.3\%$	$92.4 \pm 0.2\%$	$89.3 \pm 0.1\%$
	Covid-19 Negative	$90.6 \pm 0.1\%$	$91.7 \pm 0.1\%$	$94.3 \pm 0.3\%$	$89.3 \pm 0.1\%$	$92.4 \pm 0.1\%$
	Overall	$90.6 \pm 0.3\%$	$90.4 \pm 0.5\%$	$90.1 \pm 0.6\%$	$90.3 \pm 0.3\%$	$90.8 \pm 0.2\%$
Symptoms data	Covid-19 Positive	$91.5 \pm 0.2\%$	$86.9 \pm 0.5\%$	$87.8 \pm 0.2\%$	$86.0 \pm 0.3\%$	$94.1 \pm 0.6\%$
	Covid-19 Negative	$91.5 \pm 0.1\%$	$93.7 \pm 0.3\%$	$93.3 \pm 0.2\%$	$94.1 \pm 0.1\%$	$86.0 \pm 0.2\%$
	Overall	**$91.5 \pm 0.3\%$**	$90.3 \pm 0.8\%$	$90.5 \pm 0.4\%$	$90.8 \pm 0.3\%$	$91.1 \pm 0.8\%$
Both	Covid-19 Positive	$96.8 \pm 0.4\%$	$95.1 \pm 0.1\%$	$94.6 \pm 0.3\%$	$95.6 \pm 0.1\%$	$97.3 \pm 0.2\%$
	Covid-19 Negative	$96.8 \pm 0.1\%$	$97.6 \pm 0.3\%$	$97.8 \pm 0.4\%$	$97.3 \pm 0.2\%$	$95.6 \pm 0.3\%$
	Overall	$96.8 \pm 0.5\%$	$96.3 \pm 0.4\%$	$96.2 \pm 0.7\%$	$96.5 \pm 0.3\%$	$96.5 \pm 0.5\%$

Model performance metrics across different models on Task 2 Using Demographic & Symptoms Data Only

Results (Binary Classification : Task 3)

		F1-score	Precision	Sensitivity	Specificity	Accuracy
Cough data	Covid-19 Positive	$90.6 \pm 0.2\%$	$89.1 \pm 0.4\%$	$86.2 \pm 0.3\%$	$92.4 \pm 0.2\%$	$89.3 \pm 0.1\%$
	Covid-19 Negative	$90.6 \pm 0.1\%$	$91.7 \pm 0.1\%$	$94.3 \pm 0.3\%$	$89.3 \pm 0.1\%$	$92.4 \pm 0.1\%$
	Overall	$90.6 \pm 0.3\%$	$90.4 \pm 0.5\%$	$90.1 \pm 0.6\%$	$90.3 \pm 0.3\%$	$90.8 \pm 0.2\%$
Symptoms data	Covid-19 Positive	$91.5 \pm 0.2\%$	$86.9 \pm 0.5\%$	$87.8 \pm 0.2\%$	$86.0 \pm 0.3\%$	$94.1 \pm 0.6\%$
	Covid-19 Negative	$91.5 \pm 0.1\%$	$93.7 \pm 0.3\%$	$93.3 \pm 0.2\%$	$94.1 \pm 0.1\%$	$86.0 \pm 0.2\%$
	Overall	$91.5 \pm 0.3\%$	$90.3 \pm 0.8\%$	$90.5 \pm 0.4\%$	$90.8 \pm 0.3\%$	$91.1 \pm 0.8\%$
Both	Covid-19 Positive	$96.8 \pm 0.4\%$	$95.1 \pm 0.1\%$	$94.6 \pm 0.3\%$	$95.6 \pm 0.1\%$	$97.3 \pm 0.2\%$
	Covid-19 Negative	$96.8 \pm 0.1\%$	$97.6 \pm 0.3\%$	$97.8 \pm 0.4\%$	$97.3 \pm 0.2\%$	$95.6 \pm 0.3\%$
	Overall	$\mathbf{96.8 \pm 0.5\%}$	$96.3 \pm 0.4\%$	$96.2 \pm 0.7\%$	$96.5 \pm 0.3\%$	$96.5 \pm 0.5\%$

Model performance metrics across different models on Task 3 Using Using Both

Results (Multi-Class Classification : Task 4)

	F1-score	Precision	Sensitivity	Specificity	Accuracy
Covid-19 Positive	$86.38 \pm 0.03\%$	$81.88 \pm 0.01\%$	$91.39 \pm 0.04\%$	$97.49 \pm 0.03\%$	$96.81 \pm 0.05\%$
Covid-19 Negative	$92.16 \pm 0.01\%$	$95.09 \pm 0.02\%$	$89.41 \pm 0.08\%$	$98.64 \pm 0.05\%$	$96.55 \pm 0.11\%$
Bronchitis	$92.85 \pm 0.04\%$	$97.70 \pm 0.05\%$	$88.45 \pm 0.05\%$	$98.08 \pm 0.12\%$	$93.46 \pm 0.02\%$
Asthma	$83.88 \pm 0.13\%$	$75.46 \pm 0.03\%$	$94.41 \pm 0.04\%$	$93.10 \pm 0.01\%$	$93.34 \pm 0.05\%$
Overall	$90.09 \pm 0.17\%$	$90.92 \pm 0.09\%$	$90.41 \pm 0.14\%$	**$96.83 \pm 0.06\%$**	$95.04 \pm 0.18\%$

Model performance metrics across four different diseases

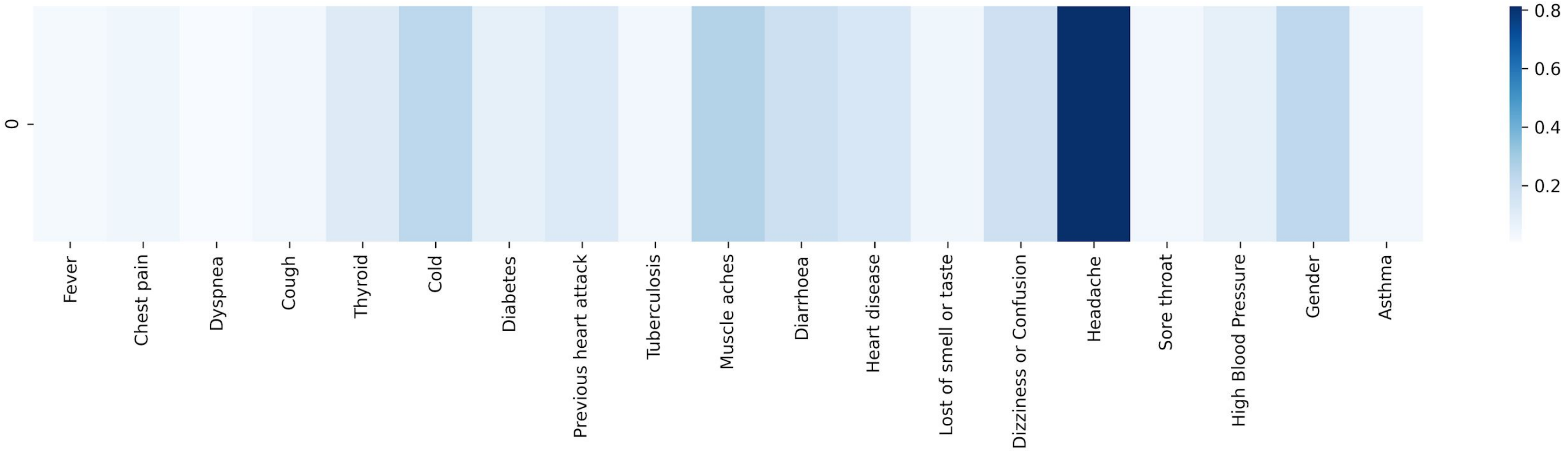
Interpretability

The Clinical selection of an algorithm depends on two main factors,

- its clinical usefulness, and
- trustworthiness.

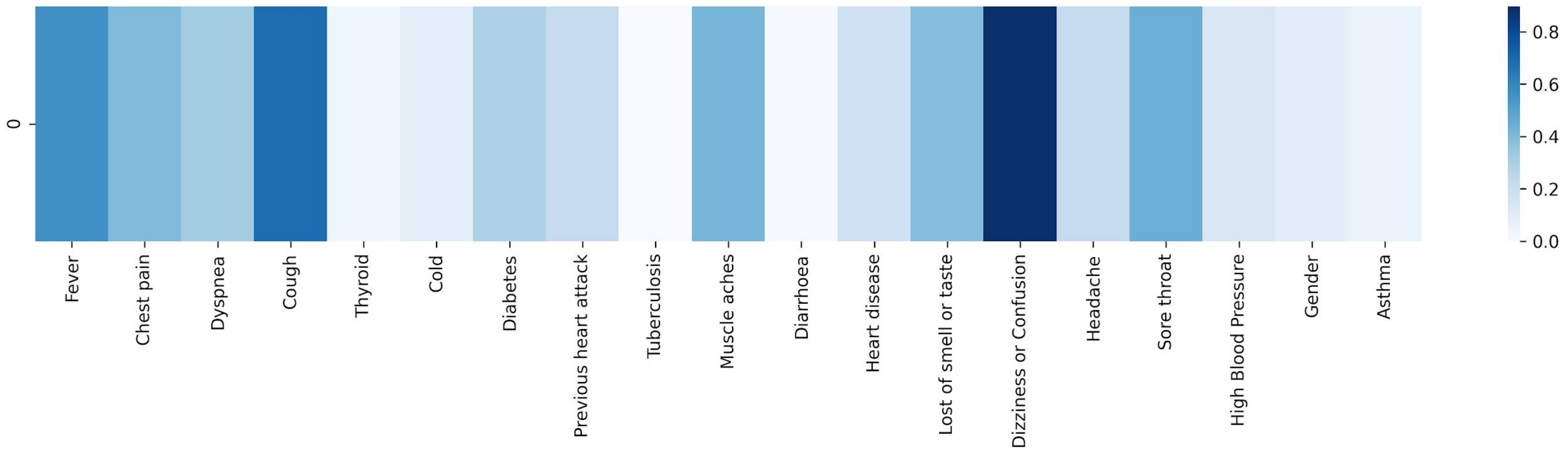
When the prediction does not directly explain a particular clinical question, its use is limited

Interpretability



Attention distribution over the Symptoms of a random Healthy(COVID-19 Negative) person. The color depth expresses the seriousness of a symptom.

Interpretability



Attention distribution over the Symptoms of a random COVID-19 infected person, The color depth expresses the seriousness of a symptom.

In-Depth Analysis (Different Types of cough)

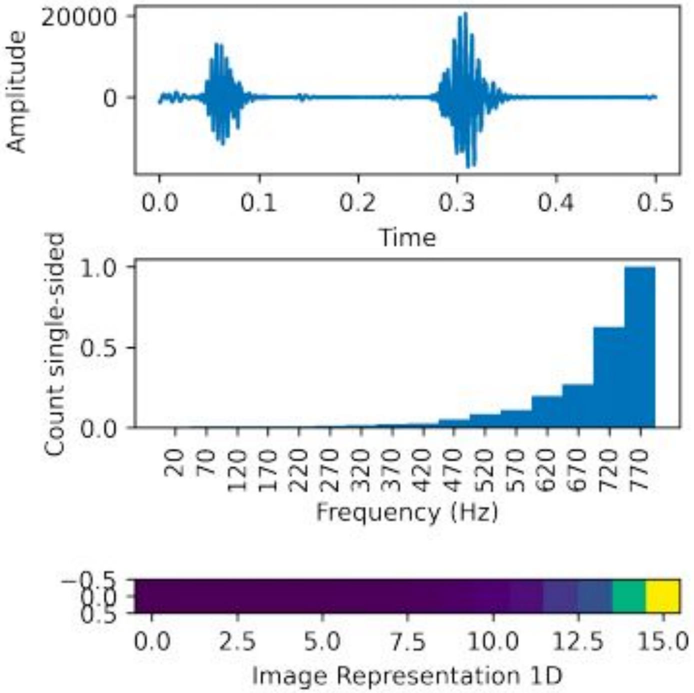


Figure 1: Healthy Cough

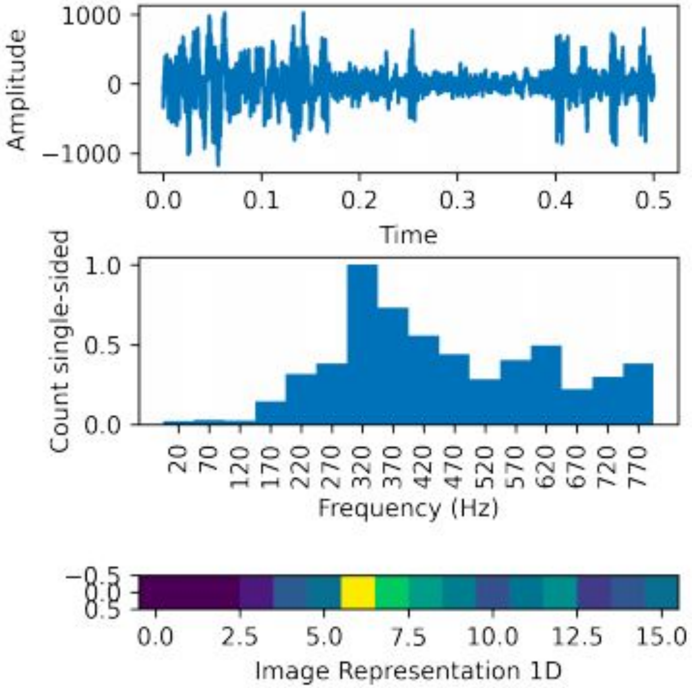


Figure 2: Asthma Cough

Healthy & Asthma Cough with their original sound, FFT output, and 1D image representation

In-Depth Analysis (Different Types of cough)

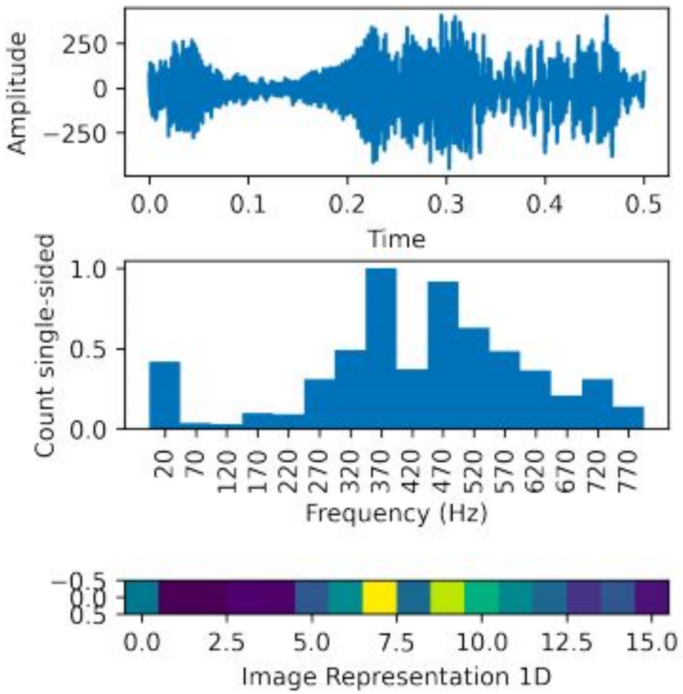


Figure 3: Bronchitis Cough

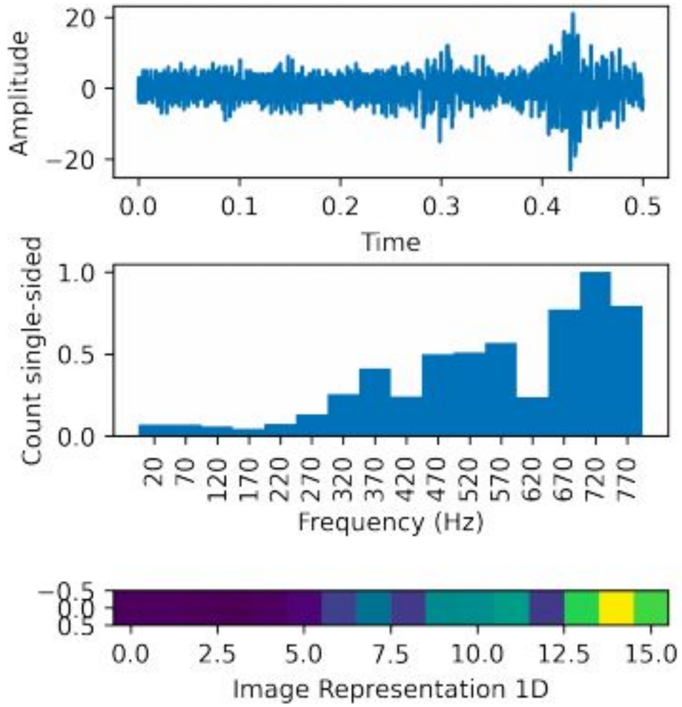


Figure 4: COVID-19 Cough

Bronchitis & COVID-19 Cough with their original sound, FFT output, and 1D image representation

In-Depth Analysis (Peak Analysis)

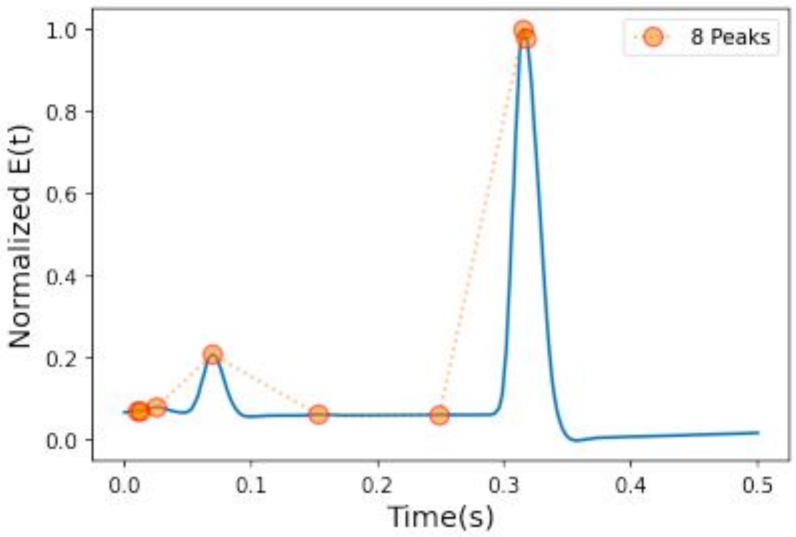


Figure 1: Peaks in Healthy Cough

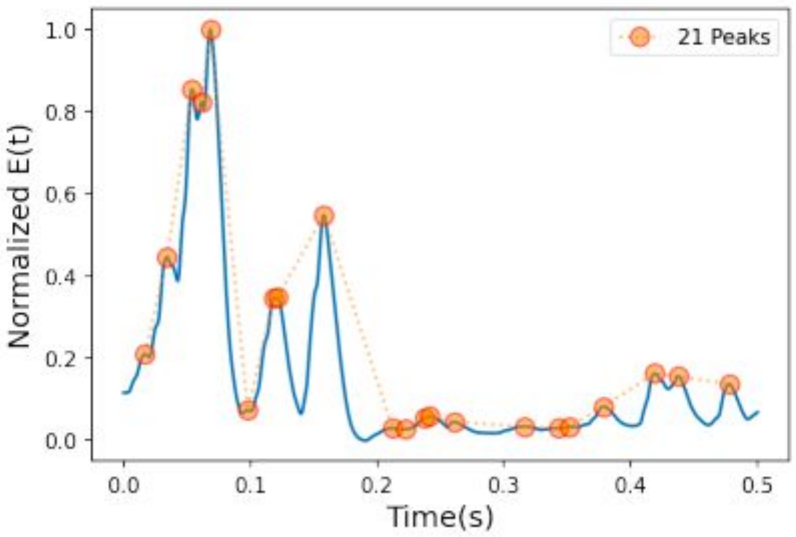


Figure 2: Peaks in Ashtma Cough

Peaks analysis in Healthy & Asthma cough. Where y-axis represents Normalized Energy envelope E(t), obtained from applying band-pass filter followed by second-order Butterworth low pass filter, and the x-axis represents the time in seconds.

In-Depth Analysis (Peak Analysis)

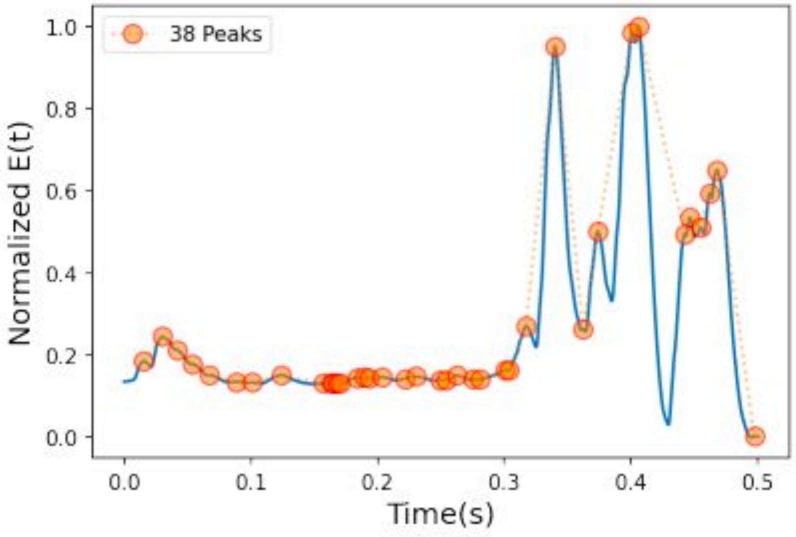


Figure 3: Peaks in Bronchitis Cough

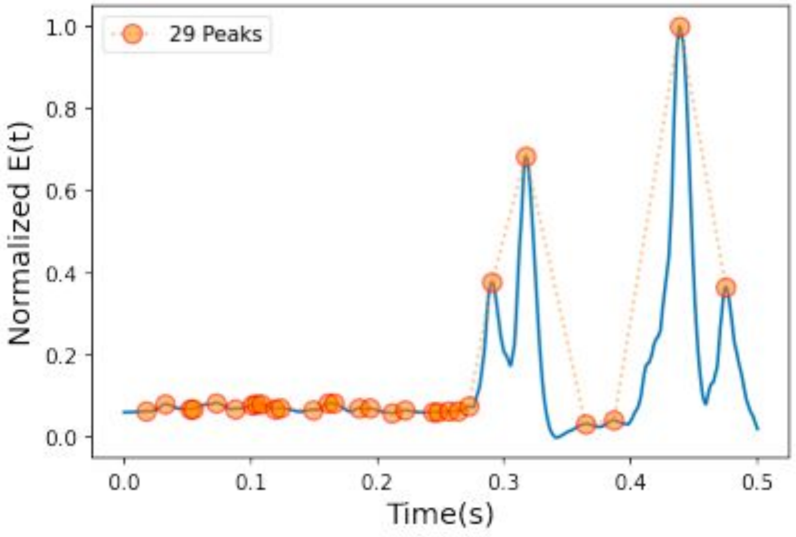
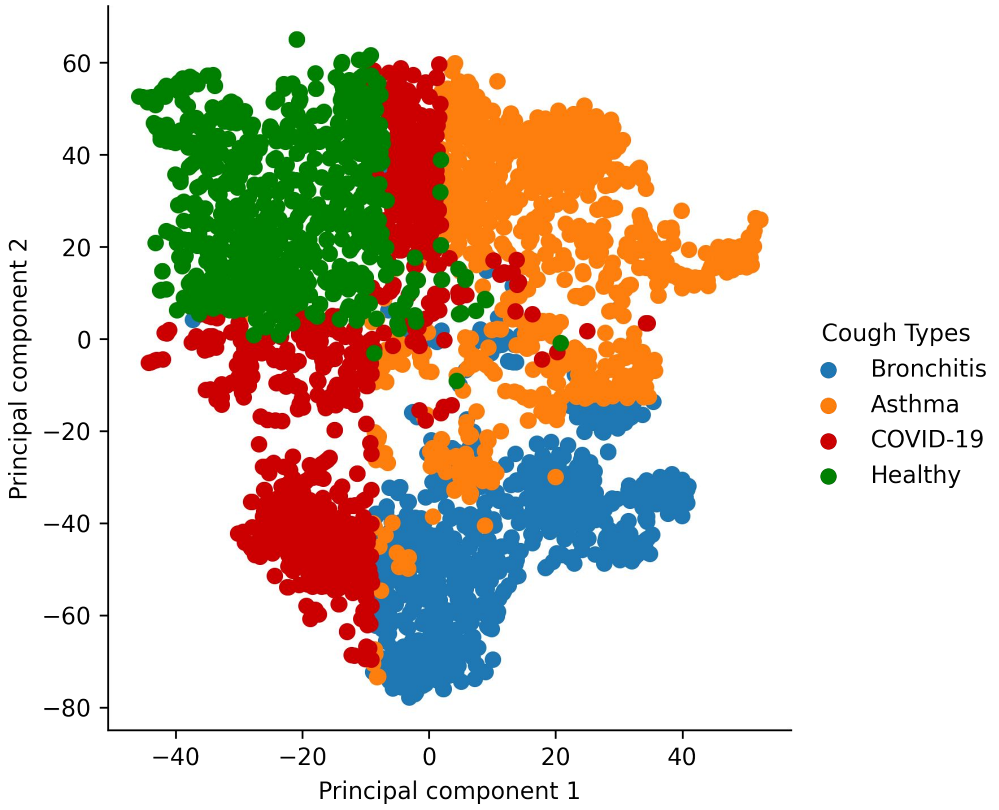


Figure 4: Peaks in COVID-19 Cough

Peaks analysis in Bronchitis & COVID-19 cough. Where y-axis represents Normalized Energy envelope E(t), obtained from applying band-pass filter followed by second-order Butterworth low pass filter, and the x-axis represents the time in seconds.

In-Depth Analysis (t-SNE visualization)



t-SNE visualization of four types of Cough features

Scope & Future work

- Low cost, rapid and interpretable AI-based diagnostic tool
- Large-scale COVID-19 disease screening and areas where healthcare facilities are not easily accessible
- Experiments will be carried out in the future by incorporating additional voice data features such as breathing sound, counting sound (natural voice samples), and sustained vowel phonation

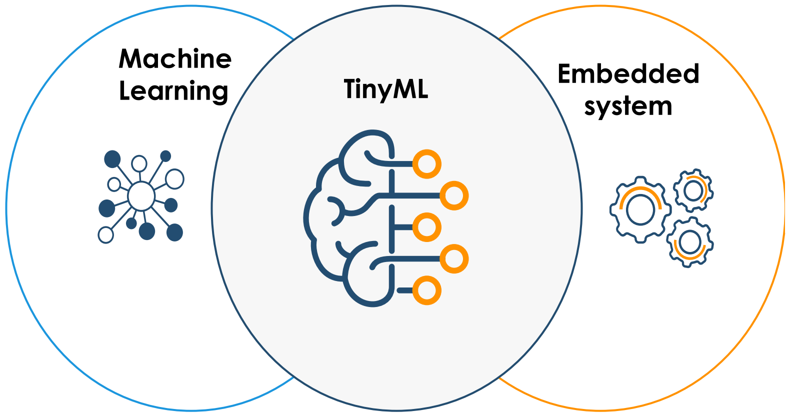
Take-Home Messages

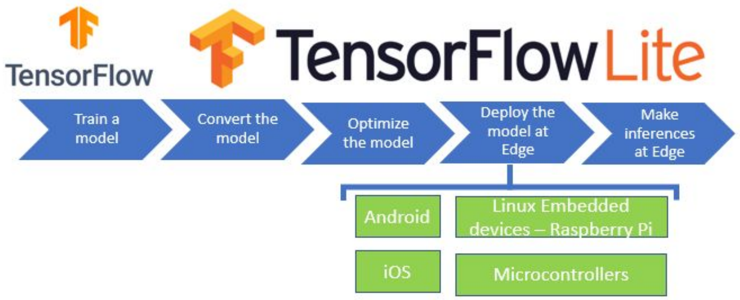
- It isn’t a replacement or a medically approved diagnostic tool; instead, it is a first-level diagnosis method for anyone, anywhere, anytime
- Cough detection and classification is fast, affordable, and accurate using AI-based methods
- We utilized different types of patient information; The results prove to be transparent, interpretable, and multi-model learning in cough classification research
- Additional clinical data such as case trajectory, patient characteristics, diagnosis, symptoms, comorbidities, and outcomes & EHR data could be utilized to capture better symptoms & cough features to build a rapid, accurate, and easy-to-access mobile diagnosis tool for different diseases

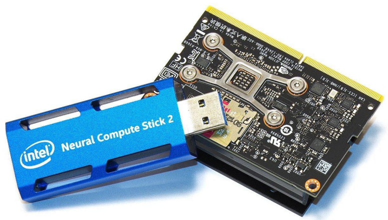
Edge deployment of models (on device and close-to-endpoint server)

Using Tiny ML, Intel's Neural Compute Stick

TinyML is a type of machine learning that shrinks deep learning networks to fit on tiny hardware.







- Trained edge level machine learning models using Intel's Neural Compute Stick
- Utilized the TinyML to train models on raspberry pi and deployed internally

Questions?

Thanks !

ankit.pal@saama.com

aadityaura.github.io
twitter : @aadityaura