



**HAL**  
open science

# Towards cough sound analysis using the Internet of things and deep learning for pulmonary disease prediction

Ajay Kumar, Kumar Abhishek, Muhammad Ghalib, Pranav Nerurkar, Kunjal Shah, Madhav Chandane, Sunil Bhirud, Dhiren Patel, Yann Busnel

► **To cite this version:**

Ajay Kumar, Kumar Abhishek, Muhammad Ghalib, Pranav Nerurkar, Kunjal Shah, et al.. Towards cough sound analysis using the Internet of things and deep learning for pulmonary disease prediction. Transactions on emerging telecommunications technologies, In press, 10.1002/ett.4184 . hal-03035289v1

**HAL Id: hal-03035289**

**<https://imt-atlantique.hal.science/hal-03035289v1>**

Submitted on 2 Dec 2020 (v1), last revised 18 Jan 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Towards cough sound analysis using the internet of things and deep learning for pulmonary disease prediction <sup>†</sup>

Ajay Kumar\*<sup>1</sup> | Kumar Abhishek<sup>1</sup> | Muhammad Rukunuddin Ghalib<sup>2</sup> | Pranav Nerurkar<sup>3,4</sup> | Kunjal Shah<sup>3</sup> | Madhav Chandane<sup>3</sup> | Sunil Bhirud<sup>3</sup> | Dhiren Patel<sup>3</sup> | Yann Busnel<sup>5</sup>

<sup>1</sup>Dept. of CSE, NIT Patna, Bihar, India

<sup>2</sup>School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Tamil Nadu, India

<sup>3</sup>Dept. of CE & IT, VJTI Mumbai, Maharashtra, India

<sup>4</sup>Dept. of Data Science, MPSTME, NMIMS, Mumbai, Maharashtra, India

<sup>5</sup>SRCD department, IMT Atlantique, IRISA, Rennes, France

## Correspondence

\*Ajay Kumar, Dept. of CSE, NIT Patna.  
Email: ajayk.phd18.cs@nitp.ac.in

## Present Address

Dept. of CSE, NIT Patna, Bihar, India

## Abstract

Cough is a symptom in over a hundred respiratory diseases. The audio features in cough signals contain erudition about the predicament of the respiratory system. Using deep learning or signal processing, these features can be used to build an effective disease prediction system. However, cough analysis remains an area that has received scant attention from machine learning researchers. This can be attributed to several factors such as inefficient ancillary systems, high expenses in obtaining datasets, or difficulty in building classifiers. This paper categorized and reviewed the current progress on cough audio analysis for the classification of pulmonary diseases. It also explored potential future issues in research. Additionally, it proposed a model for the classification of ten serious pulmonary ailments commonly seen in Indian adolescents. The proposed model is evaluated against four existing state of the art techniques in the literature.

## KEYWORDS:

Cough classification; Machine learning; Pulmonary diseases; Internet of Things; Deep Learning

## 1 | INTRODUCTION

Diseases that affect respiratory pathways of the human body create secretions in the airways. Consequentially as a protective measure exhaling such secretion leads to coughing episodes in patients<sup>1,2,3</sup>. Several pulmonary maladies such as TB or Chronic Obstructive Pulmonary Disease (COPD) can lead to exacerbations or hospitalizations and even death if not treated promptly<sup>4</sup>. Clinical investigations consider cough severity as an important indicator to understand the progression and presence of such respiratory diseases<sup>5,6,7,8</sup>. Hence, there is a need for research on disease prediction using cough analysis, which shall lead to preventive healthcare applications and provide timely care to patients. However, limited progress has been made on this front<sup>9</sup>. The reasons being inefficient automated cough detection systems, lack of datasets on cough sounds, and limited research on developing disease prediction models using cough analysis<sup>10</sup>.

Automatic cough detection systems surveyed in the literature need a labeled audio dataset for training a machine learning classifier that discriminates cough from non-cough sounds. Such a dataset is obtained by monitoring volunteers during their daily schedules or monitoring them in a clinical setting and recording their audio. After the audio is recorded, it is labeled manually into cough and non-cough events<sup>10</sup>. This makes data collection a laborious and cost-intensive exercise. Due to the high cost in building datasets, the majority of the studies conclude using data from 12 – 22 patients. A drawback of this process

<sup>†</sup>Deep learning in E-health systems

<sup>0</sup>**Abbreviations:** COPD, Chronic Obstructive Pulmonary Disease; TB, Tuberculosis; EEG, Electroencephalography; ECG, Electrocardiography;

<sup>0</sup>**Authorship:** A. Kumar and P. Nerurkar contributed equally.

is that the data can only be collected from a limited number of volunteers to save costs. This may affect machine learning results as the probabilistic model may not be suitable for a generalized application. To save time and cost in data collection, subjective techniques such as questionnaires and self-reported scales are widely used for understanding cough severity<sup>1</sup>. However, these fail to provide an accurate description of cough features as the patient could have forgotten or misremembered<sup>9</sup>.

Data collection and labeling costs can be saved by asking patients to produce cough samples voluntarily in a clinical setting by placing audio recording equipment in suitable positions. Commonly used devices are microphones<sup>11</sup>, electrocardiography sensors, chest belts<sup>12,13</sup>, accelerometers<sup>14,15,16</sup>, smartwatches<sup>17</sup>, wearable patches<sup>18</sup>, breath analysis<sup>19,20</sup> and thermistors placed near the nose. T Drugman *et al.* concluded that non-contact microphones were most effective<sup>16</sup> amongst these for recording cough sounds.

Once cough and non-cough effects are labeled to produce a dataset either through manual labeling or automated systems, the focus shifts towards building a suitable machine learning model for the prognostication of diseases. In this domain, the direction of research is more inclined towards signal processing centered approaches, as was observed from the literature review<sup>21,6,1,5,22,23,24,25,26,27,28,29,30,31</sup>. Even in signal processing literature, two categories were present: cough analysis through feature extraction<sup>21,6,1,5,22,23,25,26,27,28,29</sup> and data-driven cough analysis<sup>6,24,32</sup>. The former category used feature engineering on cough signals to train classifiers, whereas the latter used classifiers that learn discriminating features directly from cough signals eliminating extensive feature engineering.

Machine learning and deep learning have brought paradigm shifts in healthcare analytics fields such as X-ray diagnosis, EEG/ECG diagnosis, etc.<sup>33</sup>. However, limited availability of datasets due to difficulty in data collection and labeling have hampered the achievement of the state of the art results in cough diagnosis<sup>1,30</sup>. Unlike X-ray and EEG/ECG diagnosis that can be obtained at any time, cough episodes are occasional occurrences. Hence, a cough identifier should be with minimal type I error while having maximal recall<sup>21</sup>. The second difficulty in cough analysis is that there is high variability in cough sounds of different people and low variability in cough sounds of the same person<sup>26</sup>, making the learning of discriminative features difficult. While there exist preliminary studies on disease prediction using cough audio analysis (pertussis<sup>34</sup>, childhood pneumonia<sup>35</sup>), there is a need for extending this. It is known that there is a variation in sound signals of coughs (timbre and tone) generated by patients suffering from different diseases<sup>36,10,37</sup>. These variations can inform about the state of the respiratory system<sup>29,38</sup>.

## 1.1 | Contributions

Despite an extensive literature review, a review of the existing state of the art in cough audio analysis for pulmonary disease classification was not traceable. Hence, to highlight the direction of research in this field and identify research gaps, the current survey is drawn. The progress of research on the analysis of cough audio signals in the last two decades is provided. State of the art is tabulated in Table 1 to make available a useful starting point for beginners.

Additionally, the current paper aims to build a probabilistic model that classifies diseases using cough sounds. Targeted diseases are the common cold, tonsillitis, adenoids, laryngotracheal bronchitis, croup, asthma, pneumonia, pleural effusion, TB, COPD, and whooping cough. To train such a model, voluminous data is needed, and cough audio datasets collected for all the diseases mentioned above were unavailable in the literature. Section 3 details the procedure for data collection followed in this study to overcome this drawback. In the literature, a model that targets such a broad spectrum of pulmonary diseases was not found. Section 4 describes the mathematical model of the proposed cough classification approach that aims to detect the diseases mentioned above using cough audio analysis. Since tropical countries, especially in Asia, are known as hot-beds for respiratory ailments, the current paper restricts the scope to the Indian population for study. Datasets collected from Indian population is also a contribution of this study. An experimental study with results and discussions are highlighted in Section 5. The paper concludes in Section 6, and future works suggest that a majority of patients continue their treatment in their homes, it is essential to produce a low cost, simple to utilize the accurate device for diagnosis of cough<sup>22</sup>.

## 1.2 | Outline of the paper

The remaining paper is organized into Section 2 that traces the related works in the past two decades. Section 1 is additionally segregated into Subsections 2.1 and 2.2 that review two types of proposed techniques viz. Machine learning on audio signals and Deep learning on audio signals, respectively. Finally, the conclusion and future works are presented in Section 6.

## 2 | REVIEW OF LITERATURE

Cough classification systems aim to map cough sounds to different diseases, and cough detection systems map audio signals as cough and non-cough events. Although the outcomes are different, cough detection systems require analysis of cough audio signals, which is also the backbone of cough classification systems. Consequentially, a review of cough detection systems was imminent. Cough detection systems surveyed in the literature could be broadly categorized in this segment into two categories. Their differences are illustrated in Figure 1 .

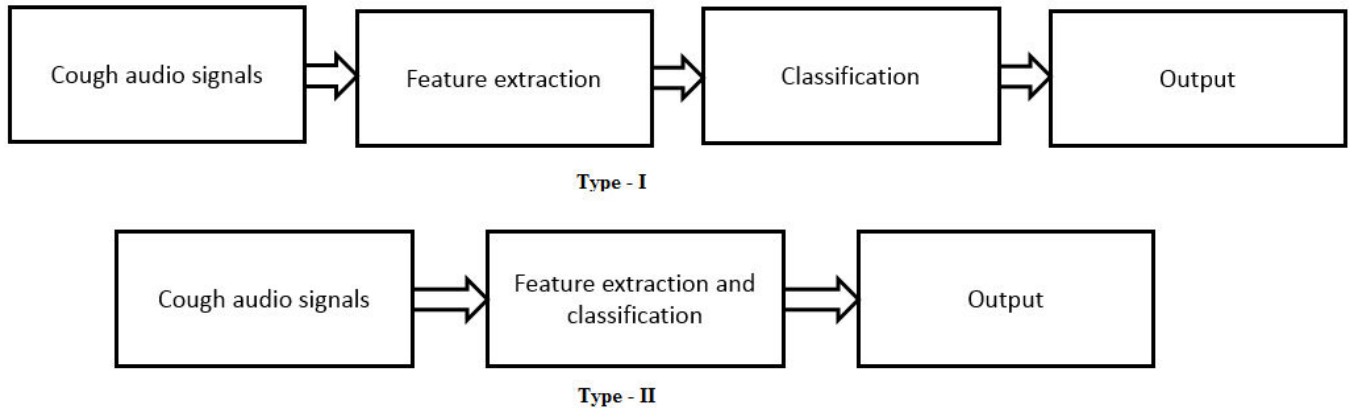


FIGURE 1 Categories of cough detection systems

### 2.1 | Type - I: Machine learning on audio signals

The raw audio is preprocessed using signal processing techniques to obtain the structured dataset. These datasets are used to train learning models for a classification task (discriminating cough and non-cough events). Signal processing techniques prevalent in the literature are Mel Frequency Cepstral Coefficients and Short Time Fourier Transform. Learning models used are neural networks, Support vector machines, decision trees, among others.

J Liu *et al.* suggested a pretrained deep learning model for cough detection. The authors collected 3873 cough audio signal samples from 22 patients using a collar fitted digital audio recorder. Features from the audio signals were extracted using 39D Mel Frequency Cepstral Coefficients (MFCC), and a structured dataset was created<sup>1</sup>. The proposed model achieved 90% on specificity and 85% on specificity. J Amoh *et al.* proposed a two-layer CNN architecture for the detection of cough events in acoustic data<sup>21</sup>. The authors used a wearable chest fitted sensor to collect lung sounds of 14 adults (7 of either sex) and created a dataset of 627 samples. Each sample (cough and non-cough) is of 64 ms window on which a 128-bin Short Time Fourier Transform (STFT) is applied to obtain a 64×16 spectral segments which are provided as inputs to the two-layered CNN for binary classification, i.e., cough or non-cough.

In contrast with the line of previous work, M Solinski *et al.* focused on feature extraction from the airflow signal of raw spirometry curve data. The authors speculated that the huge NHANES database collected from patients during spirometry tests<sup>19,39,40</sup> could be useful for developing a generic model of cough detection. The authors proposed a two-layered ANN trained using six features from airflow signals (Sensitivity - 0.852, Specificity - 0.961, Accuracy - 0.904). R Xaviero *et al.* chose a logistic regression classifier trained on a feature set comprising of three spectral features extracted from cough signals<sup>5</sup>. The authors achieved a sensitivity of 90.31%, specificity of 98.14%, and F-score of 88.70% on a dataset of sound signals 1980 (cough: 980, non-cough: 1000). A Windmon *et al.* focused on COPD detection using cough audio signal analysis. The authors built a dataset of 13 spectral features extracted from cough audios of 39 patients (COPD: 23, Non-COPD:16) and trained a random forest model on it (Recall - 85.5%, Precision - 85.6%, F-score - 85.4% and Accuracy - 85.4%)<sup>22</sup>.

L Perna *et al.* preferred to collect cough audios from remote and unobtrusive patient monitoring instead of collecting voluntarily produced cough samples taken from a controlled setting<sup>41</sup>. 12 MFCC of cough and non-cough audio were collected to build a dataset for training SVM, KNN, and XGBOOST (AUC - 0.916 ± 0.027) models of which XGBOOST gave most promising

results. H Hee *et al.* experimented with the use of cough audio analysis for building a classifier for asthmatic and normal cough sounds<sup>23</sup>. Signal processing techniques such as MFCC and Constant-Q Cepstral Coefficients (CQCC) were applied on asthmatic cough and non-asthmatic cough samples of 89 children to extract features. The training set (asthmatic cough-1192, healthy cough-1140) built from this procedure was used to train a Gaussian Mixture Model—Universal Background Model (GMM-UBM) that achieved specificity - 82.81% and sensitivity - 84.76%. P Kadambi *et al.* calculated 13 MFCCs, 13 MFCC delta features, 13 MFCC delta-delta features, log energy within the 13 MFCCs, 13 MFCC delta features, and 13 MFCC delta-delta features of a cough and non-cough audio signal<sup>25</sup>. The features are used to train a three-layered ANN for binary classification.

P Klco *et al.* used Octonionic Neural Network architecture (ONN) for the segregation of cough and non-cough samples. The authors argued that sensitivity 96.8% and specificity 98.4% was obtained by ONN on the dataset of 95200 audio samples. Training of ONN was done using Fast Fourier Transformation (FFT) and MFCC spectral characteristics<sup>26</sup>. E Larson *et al.* used the PCA on an audio spectrogram (generated through FFT) of cough and non-cough samples to train a random forest model for classification. The authors found ten principal components sufficient for optimal classification (true positive rate of 92% and false positive rate of 0.5%)<sup>27</sup>. In<sup>28</sup>, 12th order MFCC and the 0th order energy, along with their first and second temporal derivatives of cough and non-cough sounds, are used as features to train an ANN for cough classification.

## 2.2 | Deep learning on audio signals

J Amoh *et al.* experimented with 2D Convolutional neural networks (2dCNN) and recurrent neural networks (RNN) for cough detection<sup>6</sup>. The authors avoided handcrafted features from cough audio signals for distinguishing them from non-cough audio signals and preferred to learn discriminative features from data. The models proposed by the authors yielded a specificity of 92.7% (2dCNN) and sensitivity of 87.7% (RNN). H Wang *et al.* investigated optimal methods for encoding audio signals into images for training CNNs<sup>24</sup>. The authors used five techniques viz. original spectrum, RASTA-PLP power spectrum, RASTA-PLP cepstrum, 12th order PLP power spectrum without RASTA, and 12th order PLP cepstrum without RASTA. The authors found RASTA-PLP cepstrum as most suitable encoding method (accuracy- 0.9965, F-score- 0.9768 ). In<sup>32</sup> aimed to empirically determine the efficacy of STFT, MFB, and MFCCs for feature engineering and deep neural networks, CNN and LSTM for classification. The authors concluded that allowing the data-driven feature learning would give a better performance than hand engineering features.

## 2.3 | Discussions

Tracing the direction followed by the scientific literature, it was observed that sparse attention had been received by this field compared to other areas of machine learning-based diagnosis such as detection of breast cancer, EEG, ECGs, etc. However, the existing techniques can be grouped into feature engineering-based methods or deep learning-based methods. In the early days of machine learning research, as datasets were limited and computation power or storage was expensive, feature engineering was possible and also gave good results.

However, as large datasets emerged coupled with a decrease in the cost of computation power and storage, the trend emerged towards the use of deep learning models<sup>33</sup>. An advantage of deep learning techniques was that they eliminated the need for extensive feature engineering. Due to deep learning, the attention brought into focus use of sensors and wearable gadgets that could ease the load of doctors and technicians in recording body vitals and storing data. Sensor technology aligned with cloud storage had already brought paradigm shifts in diagnosing ailments such as tumors or cardiac-related. The same was not observed in pulmonary disease diagnosis<sup>45,46</sup>.

From the literature review in Section 2, it is concluded that the signal processing of cough audio signals was used reliably to classify cough sounds from non-cough events. Comparatively, the domain of cough classification was relatively unexplored. In this respect, in Section 3, a cough classification model was implemented.

## 3 | MATERIALS AND METHODS

### 3.1 | Subjects

The data for the research is a corpus of audio samples that are voluntarily generated cough samples by patients aged 3-73 visiting a paediatric clinic in Mumbai city of India. Additionally, physical indicators are given in Table 3 were also collected from the

**TABLE 1** Summary of State of the Art techniques in literature

| Authors                                 | Cough data collection                | Feature extraction   | Models used  | Total samples |
|---|--------------------------------------|--|--|---------------|
| <i>J Liu et al.</i> <sup>1</sup>        | Collar fitted mic                    | MFCC   | Neural network pretrained using Deep belief networks | 3873          |
| <i>J Amoh et al.</i> <sup>21</sup>      | Wearable acoustic sensor             | Spectrogram obtained from Short Time Fourier Transform (STFT)      | Two layered CNN model                                | 627           |
| <i>M SoliÅłski et al.</i> <sup>19</sup> | NHANES database of spirometry curves | Feature extraction from signal processing                          | Two layered ANN model                                | 10874         |
| <i>R Xaviero et al.</i> <sup>34</sup>   | Multiple repositories <sup>5</sup>   | Feature extraction from STFT                                       | Logistic regression model                            | 1980          |
| <i>A Windmon et al.</i> <sup>22</sup>   | Smart phone mic                      | 13 features extracted using STFT, MFCC                             | Random Forest model                                  | 165           |
| <i>J Amoh et al.</i> <sup>6</sup>       | Wearable acoustic sensor             | Cough audio (RNN), Spectr-temporal image of cough audio using STFT | RNN and 2dCNN model                                  | 627           |
| <i>L Perna et al.</i> <sup>41</sup>     | Microphone                           | 12 MFCC  | SVM, KNN, XGBOOST model                              | 31754         |
| <i>H Hee et al.</i> <sup>23</sup>       | Microphone                           | CQCC and MFCC  | GMM-UBM model  | 2332          |
| <i>P Kadambi et al.</i> <sup>25</sup>   | Microphone                           | 13 MFCC  | ANN model  | 5670          |
| <i>E Larson et al.</i> <sup>31</sup>    | Smart phone mic                      | FFT, PCA   | RF model   | *             |
| <i>J Alvarez et al.</i> <sup>42</sup>   | Smart phone mic                      | Spectral features  | SVM model  | *             |
| <i>K Nguyen et al.</i> <sup>15</sup>    | Smart phone mic, accelerometer       | MFCC, Chroma feature analysis, zero crossing rate                  | SVM, ANN, RF   | 800           |
| <i>R Sharan et al.</i> <sup>7</sup>     | Smart phone mic                      | MFCC   | SVM  | 3262          |
| <i>S Matos et al.</i> <sup>43,44</sup>  | Microphone                           | MFCC, linear predictive coding, filter banks output                | HMM  | 4268          |

patients. Experts from Defence Army Hospital Research and Referral experts working at a private nursing home in Mumbai were involved. Voluntarily produced cough samples were preferred in the study as they avoid the need for manual labeling and have been found efficacious in previous literature<sup>16,47</sup> for testing cough monitors and clinical settings. Consistent with the clinical environment in<sup>16,47</sup>, subjects were kept in a quiet room with a recording device. Clinical studies revealed the age groups that are affected by pulmonary diseases that manifest cough symptoms<sup>10</sup> (Table 2 ).

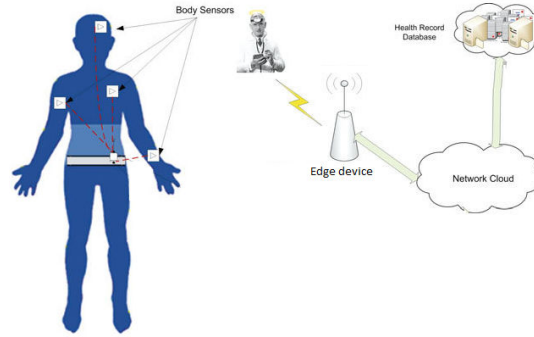
### 3.2 | Data collection

Data of vitals signs of the subjects was collected through examination by an experienced doctor using handheld sensors and questionnaires. While diagnosing the subjects, the doctor would dictate the vitals of the subject to an assistant. The assistant used a web-based application to enter the readings. This data was then stored into a SQL database (Table 3 ).

The diagnosis of the doctor was added to the dataset as a categorical attribute named Disease (no disease=0, common cold=1, tonsillitis=2, adenoids=3, foreign body in throat=4, laryngothreacheo bronchitis=5, croup=6, asthma=7, Pneumonia=8, Plural effusion=9, COPD=10, tuberculosis=11, whooping cough=12). Cough signals were recorded using the microphone of Xiaomi ME17S with recording quality configured at 44.1 kHz frame rate, wav format, two audio channels, and no compression.

**TABLE 2** Common pulmonary diseases and age groups affected

| Disease                    | Age group<br>in years |
|----------------------------|-----------------------|
| Common cold                | 0+                    |
| Tonsillitis                | 2-15                  |
| Adenoids                   | 6-16                  |
| Foreign body in throat     | 0+                    |
| Laryngotheacheo bronchitis | 0.5-3                 |
| Croup                      | 0.5-3                 |
| Asthma                     | 5-16                  |
| Pneumonia                  | 0+                    |
| Pleural effusion           | 21-30                 |
| COPD                       | 30+                   |
| Tuberculosis               | 0+                    |
| Whooping cough             | 0-18                  |

**FIGURE 2** Data collection

### 3.3 | Data preprocessing

Preprocessing included converting text data to categorical values. The attributes that needed conversion are indicated in Table 3. This data conversion was done to obtain body vitals in the format required for the application of machine learning techniques.

## 4 | MATHEMATICAL MODEL: PULMONARY DISEASES DETECTION

### 4.1 | Proposed approach

Each training sample  $(x^{(i)}, y^{(i)})$  of dataset  $\mathcal{D}$  is with  $m$  features and total  $n$  samples are present in the dataset. Hence,  $\mathcal{D} = \{(x_i, y_i)\}$  ( $|\mathcal{D}| = n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R}$ ). The proposed tree ensemble model uses  $K$  additive functions to predict the output.

$$\hat{y}_i = \phi(\mathbf{x}_i) = \sum_{k=1}^K f_k(\mathbf{x}_i), \quad f_k \in \mathcal{F}, \quad (1)$$

where

- $\mathcal{F} = \{f(\mathbf{x}) = w_{q(\mathbf{x})}\} (q : \mathbb{R}^m \rightarrow T, w \in \mathbb{R}^T)$ : set of regression trees
- $q$ : tree structure mapping an  $x^{(i)}$  to its leaf index
- $T$ : leaves count in the tree

**TABLE 3** Description of the dataset

| Physical indicator                     | Collection mechanism | Attributes  |
|--|----------------------|---|
| Days since cough symptoms started      | Questionnaire        | Numeric   |
| Onset of cough                         | Questionnaire        | Categorical (none=0, sudden=1, gradual=2)   |
| Exacerbating factors for cough         | Questionnaire        | Categorical (none=0, food=1, dust=2, smoke=3, exercise=4, sleep=5)  |
| Relieving factors for cough            | Questionnaire        | Categorical (none=0, vomiting=1, sputum=2)  |
| Type of cough                          | Questionnaire        | Categorical (none=0, wet=1, barking=2, dry=3, whoop=4)  |
| Type of cold associated with cough     | Questionnaire        | Categorical (none=0, running nose=1, water discharge=2, mucupurulent=3, sneezing=4, watery eyes=5)                            |
| Irritation or pain in throat           | Questionnaire        | Categorical (yes=0, no=1)   |
| Weightloss                             | Questionnaire        | Categorical (yes=0, no=1)   |
| Difficulty in swallowing               | Questionnaire        | Categorical (yes=0, no=1)   |
| Loose motions                          | Questionnaire        | Categorical (yes=0, no=1)   |
| Weight in kg                           | Strain gauge         | Numeric   |
| Height in cm                           | Scale                | Numeric   |
| Respiratory rate per minute            | Stethoscope          | Numeric   |
| Blood pressure (mmHg)                  | Sphygmomanometer     | Numeric   |
| Pulse rate per minute                  | Pulse oximeter       | Numeric   |
| Body temperature (°C)                  | Thermometers         | Numeric   |
| Peripheral capillary oxygen saturation | Pulse oximeter       | Numeric   |
| Color                                  | Observation          | Categorical (none=0, pink=1, pale=2, cynosis=3)   |
| Respiratory distress                   | Questionnaire        | Categorical (none=0, flaring=1, subcostal=3, retraction=4, intercostal retraction=5, supraclavialar retraction=6, grunting=7) |
| Air entry in lungs                     | Stethoscope          | Categorical (good=0, bad=1)   |
| Foreign sounds in chest                | Stethoscope          | Categorical (crepitation=0, ronchii=1, clear=2, conducted sounds=3)   |
| Percussions                            | Stethoscope          | Categorical (dull=0, normal=1)  |
| Heart sounds                           | Stethoscope          | Categorical (normal=0, murmurs=1)   |
| Liver                                  | Stethoscope          | Categorical (palpable=0, nonpalpable=1)   |
| Spleen                                 | Stethoscope          | Categorical (palpable=0, nonpalpable=1)   |
| Ascitis                                | Stethoscope          | Categorical (present=0, absent=1)   |
| Central Nervous System                 | Stethoscope          | Categorical (drowsy=0, normal=1)  |
| Cough audio                            | Microphone           | Numeric timeseries  |

- $f_k$ :  $q$  having leaf weights  $w$
- $w_i$ : score on  $i^{th}$  leaf

For each  $x^{(i)}$  the decision rules of  $q$  classify it into the leaf nodes and calculate the final prediction by  $\sum w$  i.e. summing up the score in the corresponding leaves. The obtain the optimal model  $\phi$ , the loss  $\mathcal{L}(\phi)$  is minimized by following *regularized*



objective.

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (2)$$

where  $\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$

where,

- $l$ : loss function
- $\hat{y}_i$ : prediction
- $y_i$ : target
- $\Omega$ : regularization term

## 4.2 | Optimization

As the loss function given in Eq. 2 cannot be optimized using standard optimization techniques, the model is trained in an additive manner specified in<sup>48</sup>. Eq. 2 is modified as Eq 3.

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t) \quad (3)$$

where,

- $\hat{y}_i^{(t)}$ : prediction of the  $i$ -th instance at the  $t$ -th iteration
- $f_t$ :  $q$  having leaf weights  $w$

Using second-order approximation for optimizing Eq 3,

$$\mathcal{L}^{(t)} \simeq \sum_{i=1}^n [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i)] + \Omega(f_t)$$

where  $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$  and  $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$  are first and second order derivatives on the loss function. Removing the constant terms in Eq 4.2.

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^n [g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i)] + \Omega(f_t) \quad (4)$$

By defining  $I_j = \{i | q(\mathbf{x}_i) = j\}$  to be instance set of leaf  $j$  and expanding  $\Omega$ , rewriting Eq 4 as Eq 5

$$\begin{aligned} \tilde{\mathcal{L}}^{(t)} &= \sum_{i=1}^n [g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i)] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \\ &= \sum_{j=1}^T [(\sum_{i \in I_j} g_i) w_j + \frac{1}{2} (\sum_{i \in I_j} h_i + \lambda) w_j^2] + \gamma T \end{aligned} \quad (5)$$

After obtaining  $q(\mathbf{x})$  computation of optimal weights  $w_j^*$  of leaf  $j$  by Eq 6

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}, \quad (6)$$

leads to optimal value as Eq 7

$$\tilde{\mathcal{L}}^{(t)}(q) = -\frac{1}{2} \sum_{j=1}^T \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \quad (7)$$

The proposed model uses Eq 7 to measure quality (impurity score) of  $q$ . A greedy algorithm to estimate the optimal  $q$ , initiates the tree from a single leaf node by iteratively adding branches to it. Given the leaves in left and right nodes, loss reduction after the split is given by Eq 8,

$$\mathcal{L}_{split} = \frac{1}{2} \left[ \frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \quad (8)$$

where,

- $I_L$  and  $I_R$ : instance sets of left and right nodes after the split
- $\mathcal{L}_{split}$ : loss reduction after the split

### 4.3 | Algorithm

1. Initialization of weights  $\theta$  and bias  $b$
2. Configure model to minimize Loss
3. Add activation function: rectified linear unit to model
4. Set epochs, batch size
5. Optimize the loss function  $L$  using gradient descent

## 5 | EXPERIMENTAL STUDY

### 5.1 | Experimental setting

Proposed model in Section 4 is compared with models proposed by R Xavier *et al.*<sup>5</sup>, A Windmon *et al.*<sup>22</sup> and J Alvarez *et al.*<sup>42</sup>. Execution environment was IPython notebook hosted on Google colaboryatory connected to Google compute engine Python 3 (GPU) with 12.72GB RAM.

### 5.2 | Metrics

Given the true positives  $t_p$ , true negatives  $t_n$ , false positives  $f_p$  and false negatives  $f_n$  obtained from observing  $(\hat{y}_i, y_i)$ , following metrics were used.

$$Sensitivity = \frac{t_p}{t_p + f_n} \quad (9)$$

$$Specificity = \frac{t_n}{t_n + f_p} \quad (10)$$

$$Accuracy = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \quad (11)$$

$$Prevalence = \frac{t_p + f_n}{t_p + t_n + f_p + f_n} \quad (12)$$

$$PositivePredictionValue(PPV) = \frac{sensitivity * prevalence}{((sensitivity * prevalence) + ((1 - specificity) * (1 - prevalence)))} \quad (13)$$

$$NegativePredictionValue(NPV) = \frac{sensitivity * (1 - prevalence)}{((1 - sensitivity) * prevalence) + ((specificity) * (1 - prevalence))} \quad (14)$$

$$Detectionrate = \frac{t_p}{t_p + t_n + f_p + f_n} \quad (15)$$

$$Detectionprevalence = \frac{t_p + f_p}{t_p + t_n + f_p + f_n} \quad (16)$$

$$BalancedAccuracy = \frac{sensitivity + specificity}{2} \quad (17)$$

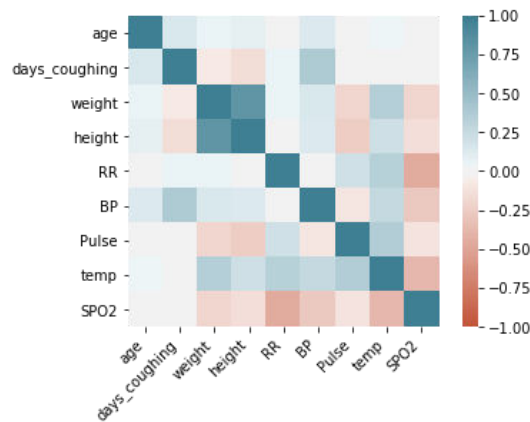
### 5.3 | Dataset

Table 4 gives the division of the attributes across the various categories. The total samples taken were from 108 subjects.

**TABLE 4** Bifurcation of the different attributes across categories

| Physical indicator                 | Categories    |                 |                     |                     |                          |                             |            |
|------------------------------------|---------------|-----------------|---------------------|---------------------|--------------------------|-----------------------------|------------|
| Gender                             | Male = 50     | Female = 58     |                     |                     |                          |                             |            |
| Onset of cough                     | None=2        | Sudden=2        | Gradual=104         |                     |                          |                             |            |
| Exacerbating factors for cough     | None=35       | Food=2          | Dust=1              | Smoke=0             | Exercise=2               | Sleep=68                    |            |
| Relieving factors for cough        | None=63       | Vomiting=23     | Sputum=22           |                     |                          |                             |            |
| Type of cough                      | None=5        | Wet=65          | Barking=3           | Dry=35              | Whooping=0               |                             |            |
| Type of cold associated with cough | None=0        | Running nose=30 | Watery discharge=24 | Mucupurulent=45     | Sneezing=20              | Watery eyes=15              |            |
| Irritation or pain in throat       | Yes=35        | No=73           |                     |                     |                          |                             |            |
| Weightloss                         | Yes=0         | No=108          |                     |                     |                          |                             |            |
| Difficulty in swallowing           | Yes=5         | No=103          |                     |                     |                          |                             |            |
| Loose motions                      | Yes=1         | No=107          |                     |                     |                          |                             |            |
| Color                              | None=3        | Pink=105        | Pale=0              | Cynosis=0           |                          |                             |            |
| Respiratory distress               | None=95       | Flaring=7       | Subcostal=0         | Retraction=0        | Intercostal retraction=6 | Supraclavialar retraction=0 | Grunting=0 |
| Air entry in lungs                 | Good=99       | Bad=9           |                     |                     |                          |                             |            |
| Foreign sounds in chest            | Crepitation=7 | Ronchii=20      | Clear=67            | Conducted sounds=14 |                          |                             |            |
| Percussions                        | Dull=11       | Normal=97       |                     |                     |                          |                             |            |
| Heart sounds                       | Normal=104    | Murmur=4        |                     |                     |                          |                             |            |
| Liver                              | Palpable=3    | Nonpalpable=105 |                     |                     |                          |                             |            |
| Spleen                             | Palpable=1    | Nonpalpable=107 |                     |                     |                          |                             |            |
| Ascitis                            | Present=0     | Absent=108      |                     |                     |                          |                             |            |
| Central Nervous System             | Drowsy=97     | Normal=11       |                     |                     |                          |                             |            |

A negative correlation exists in Figure 3 between Peripheral capillary oxygen saturation (SPO2) and weight, height, respiration rate (RR), blood pressure (BP), pulse per minute, and body temperature. Body temperature in Figure 3 shows a positive correlation with weight, height, respiration rate (RR), blood pressure (BP), and pulse per minute. There does not exist a strong correlation (>80%) between independent variables except weight and height.



**FIGURE 3** Correlation between numeric variables in the dataset

### 5.4 | Model training

The dataset was divided into training and test set with a 4:1 ratio. For hyper-parameter tuning,  $k$ -fold cross-validation method was selected ( $k = 10$ ) with upsampling to reduce the class imbalance. The final list of hyper-parameters was obtained as given,

- J Alvarez *et al.*<sup>42</sup> - degree (Polynomial Degree) = 2, scale = 0.0003422559 and C (cost) = 58.51966

- Proposed model - nrounds (Boosting Iterations)= 308, max\_depth (Max Tree Depth)= 4, eta (shrinkage)= 0.3512824, gamma (Minimum Loss Reduction)= 3.603526, colsample\_bytree (Subsample Ratio of Columns)= 0.4717281, min\_child\_weight (Minimum Sum of Instance Weight)= 4 and subsample (Subsample Percentage)= 0.2650358
- R Xaviero *et al.*<sup>5</sup> - cost = 0.01109916, loss function = L1 and epsilon (tolerance) = 1
- A Windmon *et al.*<sup>22</sup> - mtry (Randomly Selected Predictors) = 31

The Caret package of R was used for implementation. A notebook can be found at <https://www.kaggle.com/pranavn91/cough-notsound>.

## 5.5 | Results and discussion

### 5.5.1 | Overall statistics on test set

Table 5 gives a comparison of the proposed model with three baseline models. The models were evaluated on the multi-class classification problem. The dependent variable in the dataset is output or disease identified by the doctor. The disease belonged to one of the thirteen categories viz. no disease=0, common cold=1, tonsillitis=2, adenoids=3, foreign body in throat=4, laryngotheacheo bronchitis=5, croup=6, asthma=7, Pneumonia=8, Plural effusion=9, COPD=10, tuberculosis=11, whooping cough=12. However, during the period of data collection, i.e., from 1st February 2020 to 14th March 2020, no occurrence of pleural effusion, tuberculosis, and foreign body in the throat were found among the subjects. These categories were removed from the dependent variable. The average precision, accuracy, recall, and F-score of the models across thirteen disease categories is given in Table 5.

**TABLE 5** Overall statistics of performance of models on Test set

| Model                                 | Accuracy | 95% CI     | No information rate | P-value | Kappa |
|---------------------------------------|----------|------------|---------------------|---------|-------|
| J Alvarez <i>et al.</i> <sup>42</sup> | 0.41     | 0.18, 0.67 | 0.41                | 0.59    | 0.19  |
| <b>Proposed Model</b>                 | 0.41     | 0.18, 0.67 | 0.41                | 0.59    | 0.29  |
| R Xaviero <i>et al.</i> <sup>5</sup>  | 0.41     | 0.18, 0.67 | 0.41                | 0.59    | 0     |
| A Windmon <i>et al.</i> <sup>22</sup> | 0.59     | 0.33, 0.82 | 0.41                | 0.11    | 0.45  |

The proposed model achieves accuracy comparable with J Alvarez *et al.*<sup>42</sup> and R Xaviero *et al.*<sup>5</sup> but less than A Windmon *et al.*<sup>22</sup> as shown in Table 5 .

### 5.5.2 | Confusion matrix on test set

For a fine grained analysis the confusion matrix is given in Tables 6 , 7 , 8 and 9 . In these tables,  $t_p, t_n, f_p, f_n$  is given across nine disease categories for each model. Table 6 shows the SVM based model proposed by J Alvarez *et al.*<sup>42</sup> misclassified no disease, croup and asthma samples but correctly classified common cold samples.

Table 7 shows that the ensemble tree-based proposed model misclassified no disease, common cold, tonsillitis, and laryngotheacheo bronchitis and classified croup and asthma samples correctly.

Table 8 shows that the Logistic regression-based model proposed by R Xaviero *et al.*<sup>5</sup> misclassifies no disease, tonsillitis, laryngotheacheo bronchitis, croup, and asthma samples but classifies all samples of common cold accurately.

Table 9 shows that the random forest-based model proposed by A Windmon *et al.*<sup>22</sup> misclassifies no disease, common cold, tonsillitis, laryngotheacheo bronchitis and accurately classifies croup and asthma samples.

From the confusion matrix is given in Tables 6 , 7 , 8 and 9 , it is observed that the samples of diseases in the test set are limited (no disease=2, common cold=7, tonsillitis=2, adenoids=0, laryngotheacheo bronchitis=2, croup=1, asthma=3, pneumonia=0 and COPD=0). Limited samples make the dataset challenging. The proposed model could accurately classify croup, asthma and common cold in the test set.

**TABLE 6** Confusion matrix of J Alvarez *et al.*<sup>42</sup> on Test set

|                             | Reference  |             |             |          |                             |       |        |           |      |
|-----------------------------|------------|-------------|-------------|----------|-----------------------------|-------|--------|-----------|------|
| Predictions                 | no disease | common cold | tonsillitis | adenoids | laryngothreacheo bronchitis | croup | asthma | pneumonia | COPD |
| no disease                  | 0          | 1           | 0           | 0        | 1                           | 0     | 1      | 0         | 0    |
| common cold                 | 1          | 6           | 1           | 0        | 0                           | 2     | 0      | 0         | 0    |
| tonsillitis                 | 0          | 0           | 1           | 0        | 0                           | 0     | 0      | 0         | 0    |
| adenoids                    | 0          | 0           | 0           | 0        | 1                           | 0     | 0      | 0         | 0    |
| laryngothreacheo bronchitis | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| croup                       | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| asthma                      | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| pneumonia                   | 1          | 0           | 0           | 0        | 0                           | 1     | 0      | 0         | 0    |
| COPD                        | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |

**TABLE 7** Confusion matrix of Proposed model on Test set

|                             | Reference  |             |             |          |                             |       |        |           |      |
|-----------------------------|------------|-------------|-------------|----------|-----------------------------|-------|--------|-----------|------|
| Predictions                 | no disease | common cold | tonsillitis | adenoids | laryngothreacheo bronchitis | croup | asthma | pneumonia | COPD |
| no disease                  | 0          | 1           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| common cold                 | 0          | 3           | 0           | 0        | 1                           | 0     | 0      | 0         | 0    |
| tonsillitis                 | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| adenoids                    | 0          | 1           | 1           | 0        | 0                           | 0     | 0      | 0         | 0    |
| laryngothreacheo bronchitis | 0          | 1           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| croup                       | 2          | 0           | 1           | 0        | 1                           | 1     | 0      | 0         | 0    |
| asthma                      | 0          | 1           | 0           | 0        | 0                           | 0     | 3      | 0         | 0    |
| pneumonia                   | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| COPD                        | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |

**TABLE 8** Confusion matrix of R Xaviero *et al.*<sup>5</sup> on Test set

|                             | Reference  |             |             |          |                             |       |        |           |      |
|-----------------------------|------------|-------------|-------------|----------|-----------------------------|-------|--------|-----------|------|
| Predictions                 | no disease | common cold | tonsillitis | adenoids | laryngothreacheo bronchitis | croup | asthma | pneumonia | COPD |
| no disease                  | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| common cold                 | 2          | 7           | 2           | 0        | 2                           | 1     | 3      | 0         | 0    |
| tonsillitis                 | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| adenoids                    | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| laryngothreacheo bronchitis | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| croup                       | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| asthma                      | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| pneumonia                   | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |
| COPD                        | 0          | 0           | 0           | 0        | 0                           | 0     | 0      | 0         | 0    |

## 5.6 | Summary of Results: Class-wise performance statistics

Table 10 , Table 11 , Table 12 and Table 13 give the performance of the J Alvarez *et al.*<sup>42</sup>, the proposed model, R Xaviero *et al.*<sup>5</sup> and A Windmon *et al.*<sup>22</sup> models respectively, across various disease types.

None of the models could identify no disease samples in the dataset, but for croup and asthma, the proposed model achieved comparable results with the state of the art as it accurately classified all samples. However, for tonsillitis, laryngothreacheo

**TABLE 9** Confusion matrix of A Windmon *et al.*<sup>22</sup> on Test set

|                             | Reference  |             |            |          |                             |       |        |           |      |
|-----------------------------|------------|-------------|------------|----------|-----------------------------|-------|--------|-----------|------|
| Predictions                 | no disease | common cold | tonsilitis | adenoids | laryngothreacheo bronchitis | croup | asthma | pneumonia | COPD |
| no disease                  | 0          | 0           | 0          | 0        | 0                           | 0     | 0      | 0         | 0    |
| common cold                 | 0          | 0           | 1          | 0        | 1                           | 0     | 0      | 0         | 0    |
| tonsilitis                  | 0          | 0           | 0          | 0        | 0                           | 0     | 0      | 0         | 0    |
| adenoids                    | 0          | 0           | 0          | 0        | 0                           | 0     | 0      | 0         | 0    |
| laryngothreacheo bronchitis | 1          | 0           | 0          | 0        | 0                           | 0     | 0      | 0         | 0    |
| croup                       | 1          | 0           | 1          | 0        | 1                           | 1     | 0      | 0         | 0    |
| asthma                      | 0          | 1           | 0          | 0        | 0                           | 0     | 3      | 0         | 0    |
| pneumonia                   | 0          | 0           | 0          | 0        | 0                           | 0     | 0      | 0         | 0    |
| COPD                        | 0          | 0           | 0          | 0        | 0                           | 0     | 0      | 0         | 0    |

**TABLE 10** Statistics by class of J Alvarez *et al.*<sup>42</sup> on Test set

|                      | No disease | Common cold | Tonsilitis | Adenoids | Laryngothreacheo bronchitis | Croup | Asthma | Pneumonia | COPD |
|----------------------|------------|-------------|------------|----------|-----------------------------|-------|--------|-----------|------|
| Sensitivity          | 0          | 0.86        | 0.5        | NA       | 0                           | 0     | 0      | NA        | NA   |
| Specificity          | 0.8        | 0.6         | 1          | 0.94     | 1                           | 1     | 1      | 0.88      | 1    |
| PPV                  | 0          | 0.6         | 1          | NA       | NaN                         | NaN   | NaN    | NA        | NA   |
| NPV                  | 0.86       | 0.86        | 0.94       | NA       | 0.88                        | 0.94  | 0.82   | NA        | NA   |
| Prevalence           | 0.12       | 0.41        | 0.12       | 0        | 0.12                        | 0.06  | 0.18   | 0         | 0    |
| Detection rate       | 0          | 0.35        | 0.06       | 0        | 0                           | 0     | 0      | 0         | 0    |
| Detection prevalence | 0.18       | 0.59        | 0.06       | 0.06     | 0                           | 0     | 0      | 0.12      | 0    |
| Balanced accuracy    | 0.4        | 0.73        | 0.75       | NA       | 0.5                         | 0.5   | 0.5    | NA        | NA   |

**TABLE 11** Statistics by class of Proposed model on Test set

|                      | No disease | Common cold | Tonsilitis | Adenoids | Laryngothreacheo bronchitis | Croup | Asthma | Pneumonia | COPD |
|----------------------|------------|-------------|------------|----------|-----------------------------|-------|--------|-----------|------|
| Sensitivity          | 0          | 0.43        | 0          | NA       | 0                           | 1     | 1      | NA        | NA   |
| Specificity          | 0.93       | 0.9         | 1          | 0.88     | 0.93                        | 0.75  | 0.93   | 1         | 1    |
| PPV                  | 0          | 0.75        | NaN        | NA       | 0                           | 0.2   | 0.75   | NA        | NA   |
| NPV                  | 0.88       | 0.69        | 0.88       | NA       | 0.86                        | 1     | 1      | NA        | NA   |
| Prevalence           | 0.12       | 0.41        | 0.12       | 0        | 0.12                        | 0.06  | 0.18   | 0         | 0    |
| Detection rate       | 0          | 0.18        | 0          | 0        | 0                           | 0.06  | 0.18   | 0         | 0    |
| Detection prevalence | 0.06       | 0.24        | 0          | 0.12     | 0.06                        | 0.29  | 0.24   | 0         | 0    |
| Balanced accuracy    | 0.47       | 0.66        | 0.5        | NA       | 0.47                        | 0.88  | 0.96   | NA        | NA   |

bronchitis, and the common cold, it misclassified all samples. Additional training data would be suitable for improving the performance of the proposed model.

## 6 | CONCLUSION AND FUTURE WORKS

Cough audio sounds and acoustics have underlying pulmonary health knowledge. Despite this, guidelines issued by medical organizations at national and international levels fail to consider them in diagnosing pulmonary diseases. It has been speculated that by feature engineering or addition of markers within cough signals through novel applications of signal processing and deep learning, the diagnosis of pulmonary ailments can be improved, especially in developing countries. The current survey was drawn to highlight the progress made in cough audio analysis. From the literature review, it is concluded that the signal

**TABLE 12** Statistics by class of R Xaviero *et al.*<sup>5</sup> on Test set

|                             | No disease | Common cold | Tonsilitis | Adenoids | Laryngothreacheo<br>bronchitis | Croup | Asthma | Pneumonia | COPD |
|-----------------------------|------------|-------------|------------|----------|--------------------------------|-------|--------|-----------|------|
| <b>Sensitivity</b>          | 0          | 1           | 0          | NA       | 0                              | 0     | 0      | NA        | NA   |
| <b>Specificity</b>          | 1          | 0           | 1          | 1        | 1                              | 1     | 1      | 1         | 1    |
| <b>PPV</b>                  | NaN        | 0.41        | NaN        | NA       | NaN                            | NaN   | NaN    | NA        | NA   |
| <b>NPV</b>                  | 0.88       | NaN         | 0.88       | NA       | 0.88                           | 0.94  | 0.82   | NA        | NA   |
| <b>Prevalence</b>           | 0.12       | 0.41        | 0.12       | 0        | 0.12                           | 0.06  | 0.18   | 0         | 0    |
| <b>Detection rate</b>       | 0          | 0.41        | 0          | 0        | 0                              | 0     | 0      | 0         | 0    |
| <b>Detection prevalence</b> | 0          | 1           | 0          | 0        | 0                              | 0     | 0      | 0         | 0    |
| <b>Balanced accuracy</b>    | 0.5        | 0.5         | 0.5        | NA       | 0.5                            | 0.5   | 0.5    | NA        | NA   |

**TABLE 13** Statistics by class of A Windmon *et al.*<sup>22</sup> on Test set

|                             | No disease | Common cold | Tonsilitis | Adenoids | Laryngothreacheo<br>bronchitis | Croup | Asthma | Pneumonia | COPD |
|-----------------------------|------------|-------------|------------|----------|--------------------------------|-------|--------|-----------|------|
| <b>Sensitivity</b>          | 0          | 0.86        | 0          | NA       | 0                              | 1     | 1      | NA        | NA   |
| <b>Specificity</b>          | 1          | 0.8         | 1          | 1        | 0.93                           | 0.81  | 0.93   | 1         | 1    |
| <b>PPV</b>                  | NaN        | 0.75        | NaN        | NA       | 0                              | 0.25  | 0.75   | NA        | NA   |
| <b>NPV</b>                  | 0.88       | 0.89        | 0.88       | NA       | 0.88                           | 1     | 1      | NA        | NA   |
| <b>Prevalence</b>           | 0.12       | 0.41        | 0.12       | 0        | 0.12                           | 0.06  | 0.18   | 0         | 0    |
| <b>Detection rate</b>       | 0          | 0.35        | 0          | 0        | 0                              | 0.06  | 0.18   | 0         | 0    |
| <b>Detection prevalence</b> | 0          | 0.47        | 0          | 0        | 0.06                           | 0.24  | 0.24   | 0         | 0    |
| <b>Balanced accuracy</b>    | 0.5        | 0.83        | 0.5        | NA       | 0.47                           | 0.91  | 0.96   | NA        | NA   |

processing of cough audio signals can be used reliably to classify cough sounds from non-cough events. Comparatively, the domain of cough classification was relatively unexplored.

An end to end machine learning model was proposed. For evaluation of the model, a dataset collected from a pediatric facility in Mumbai, India, was used. The dataset had samples from nine mild to severe pulmonary maladies affecting adolescents in India. During the evaluation, from the results, it was observed that the proposed model could accurately identify disease samples such as croup and asthma. The dataset and model were put into the public domain for further evaluation and research.

An attempt to showcase the intuitions behind the state of the art techniques in this domain was made to provide a useful starting point for beginners. It was not an objective of the current paper to exhaustively describe the signal processing or machine learning techniques available for audio analysis. However, considering the relatively less attention received on that subject, it may be a new line of research.

## DECLARATION

### Ethical Approval and Consent to participate

Not applicable

### Consent for publication

All the authors of this paper have shown their Participation voluntarily.

### Availability of supporting data

The data will be provided based on data request by the evaluation team. Tables, figures, and latex files are available at GitHub repository <sup>1</sup>.

<sup>1</sup><https://github.com/pranavn91/blockchain>

## Competing interests

The authors of this research article declares that no conflict of interest in preparing this research article.

## Funding

This work was supported in part by the Raman Charpak Fellowship of the Indo-French Centre for the Promotion of Advanced Research Grant no: IFC/4132/RCF 2019/716. The authors thank NIT Patna, VJTI Mumbai and IMT Atlantique, France for providing the lab resources.

## Authors' contributions

The authors have proposed a probabilistic model that classifies diseases using cough sounds. Targeted diseases are the common cold, tonsillitis, adenoids, laryngotracheal bronchitis, croup, asthma, pneumonia, pleural effusion, TB, COPD, and whooping cough.

## References

1. Liu JM, You M, Wang Z, Li GZ, Xu X, Qiu Z. Cough event classification by pretrained deep neural network. *BMC medical informatics and decision making* 2015; 15(4): S2.
2. Fan Chung K, Widdicombe J. *Pharmacology and therapeutics of cough*. Springer . 2009.
3. Ma W, Yu L, Wang Y, Li X, Hanjing L, Qiu Z. Changes in health-related quality of life and clinical implications in Chinese patients with chronic cough. *Cough* 2009; 5(1): 7.
4. Hsu J, Stone R, Logan-Sinclair R, Worsdell M, Busst C, Chung K. Coughing frequency in patients with persistent cough: assessment using a 24 hour ambulatory recorder. *European Respiratory Journal* 1994; 7(7): 1246–1253.
5. Pramono RXA, Imtiaz SA, Rodriguez-Villegas E. Automatic Cough Detection in Acoustic Signal using Spectral Features. In: IEEE. ; 2019: 7153–7156.
6. Amoh J, Odame K. Deep neural networks for identifying cough sounds. *IEEE transactions on biomedical circuits and systems* 2016; 10(5): 1003–1011.
7. Sharan RV, Abeyratne UR, Swarnkar VR, Porter P. Automatic croup diagnosis using cough sound recognition. *IEEE Transactions on Biomedical Engineering* 2018; 66(2): 485–495.
8. Birring S, Fleming T, Matos S, Raj A, Evans D, Pavord I. The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough. *European Respiratory Journal* 2008; 31(5): 1013–1018.
9. Kvapilova L, Boza V, Dubec P, et al. Continuous Sound Collection Using Smartphones and Machine Learning to Measure Cough. *Digital Biomarkers* 2019; 3(3): 166–175.
10. Saba E. *Techniques for Cough Sound Analysis*. PhD thesis. University of Washington, 2018.
11. Botha G, Theron G, Warren R, et al. Detection of tuberculosis by automatic cough sound analysis. *Physiological measurement* 2018; 39(4): 045005.
12. Mahmoudi SA, Possa PDC, Ravet T, et al. Sensor-based system for automatic cough detection and classification. In: ; 2016.
13. Młynczak M, Pariaszewska K, Cybulski G. Automatic cough episode detection using a vibroacoustic sensor. In: IEEE. ; 2015: 2808–2811.
14. Mohammadi H, Samadani AA, Steele C, Chau T. Automatic discrimination between cough and non-cough accelerometry signal artefacts. *Biomedical Signal Processing and Control* 2019; 52: 394–402.



15. Nguyen KA, Luo Z. Cover your cough: detection of respiratory events with confidence using a smartwatch. In: ; 2018: 114–131.
16. Drugman T, Urbain J, Bauwens N, et al. Objective study of sensor relevance for automatic cough detection. *IEEE journal of biomedical and health informatics* 2013; 17(3): 699–707.
17. Liaqat D, Wu R, Son T, et al. Towards Ambulatory Cough Monitoring Using Smartwatches. In: American Thoracic Society. 2018 (pp. A4929–A4929).
18. Elfaramawy T, Fall CL, Morissette M, Lellouche F, Gosselin B. Wireless respiratory monitoring and coughing detection using a wearable patch sensor network. In: IEEE. ; 2017: 197–200.
19. Soliński M, Łeppek M, Kołtowski Ł. Automatic cough detection for portable spirometry system trained on large database. *arXiv preprint arXiv:1903.03588* 2019.
20. Phillips CO, Syed Y, Mac Parthaláin N, Zwigelaar R, Claypole TC, Lewis KE. Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls. *Journal of breath research* 2012; 6(3): 036003.
21. Amoh J, Odame K. DeepCough: A deep convolutional neural network in a wearable cough detection system. In: IEEE. ; 2015: 1–4.
22. Windmon A, Minakshi M, Chellappan S, Athilingam P, Johansson M, Jenkins BA. On Detecting Chronic Obstructive Pulmonary Disease (COPD) Cough using Audio Signals Recorded from Smart-Phones.. In: ; 2018: 329–338.
23. Hee HI, Balamurali B, Karunakaran A, et al. Development of Machine Learning for Asthmatic and Healthy Voluntary Cough Sounds: A Proof of Concept Study. *Applied Sciences* 2019; 9(14): 2833.
24. Wang HH, Liu JM, You M, Li GZ. Audio signals encoding for cough classification using convolutional neural networks: A comparative study. In: IEEE. ; 2015: 442–445.
25. Kadambi P, Mohanty A, Ren H, et al. Towards a Wearable Cough Detector Based on Neural Networks. In: IEEE. ; 2018: 2161–2165.
26. Klco P, Kollarik M, Tatar M. Novel computer algorithm for cough monitoring based on octonions. *Respiratory physiology & neurobiology* 2018; 257: 36–41.
27. Larson EC, Lee T, Liu S, Rosenfeld M, Patel SN. Accurate and privacy preserving cough sensing using a low-cost microphone. In: ; 2011: 375–384.
28. Liu JM, You M, Wang Z, Li GZ, Xu X, Qiu Z. Cough detection using deep neural networks. In: IEEE. ; 2014: 560–563.
29. Monge-Álvarez J, Hoyos-Barceló C, San-José-Revuelta LM, Higuera C.-d.-IP. A machine hearing system for robust cough detection based on a high-level representation of band-specific audio features. *IEEE Transactions on Biomedical Engineering* 2018; 66(8): 2319–2330.
30. Barry SJ, Dane AD, Morice AH, Walmsley AD. The automatic recognition and counting of cough. *Cough* 2006; 2(1): 8.
31. Larson S, Comina G, Gilman RH, Tracey BH, Bravard M, López JW. Validation of an automated cough detection algorithm for tracking recovery of pulmonary tuberculosis patients. *PloS one* 2012; 7(10).
32. Miranda ID, Diacon AH, Niesler TR. A Comparative Study of Features for Acoustic Cough Detection Using Deep Architectures. In: IEEE. ; 2019: 2601–2605.
33. Sangaiah AK, Arumugam M, Bian GB. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. *Artificial Intelligence in Medicine* 2020; 103: 101788.
34. Pramono RXA, Imtiaz SA, Rodriguez-Villegas E. A cough-based algorithm for automatic diagnosis of pertussis. *PloS one* 2016; 11(9).

35. Abeyratne UR, Swarnkar V, Setyati A, Triasih R. Cough sound analysis can rapidly diagnose childhood pneumonia. *Annals of biomedical engineering* 2013; 41(11): 2448–2462.
36. Korpáš J, Sadloňová J, Vrabec M. Analysis of the cough sound: an overview. *Pulmonary pharmacology* 1996; 9(5-6): 261–268.
37. Amoh J, Odame K. Technologies for developing ambulatory cough monitoring devices. *Critical Reviews in Biomedical Engineering* 2013; 41(6).
38. Abaza AA, Day JB, Reynolds JS, et al. Classification of voluntary cough sound and airflow patterns for detecting abnormal pulmonary function. *Cough* 2009; 5(1): 8.
39. Awan KM, Sherazi HHR, Ali A, Iqbal R, Khan ZA, Mukherjee M. Energy-aware cluster-based routing optimization for WSNs in the livestock industry. *Transactions on Emerging Telecommunications Technologies* 2019: e3816.
40. Sahraoui Y, Korichi A, Kerrache CA, Bilal M, Amadeo M. Remote Sensing to Control Respiratory Viral Diseases Outbreaks using Internet of Vehicles. *Transactions on Emerging Telecommunications Technologies* 2020: e4118.
41. Di Perna L, Spina G, Thackray-Nocera S, et al. An automated and unobtrusive system for cough detection. In: IEEE. ; 2017: 190–193.
42. Monge-Álvarez J, Hoyos-Barceló C, Lesso P, Higuera C.-d.-IP. Robust detection of audio-cough events using local hu moments. *IEEE journal of biomedical and health informatics* 2018; 23(1): 184–196.
43. Matos S, Birring SS, Pavord ID, Evans H. Detection of cough signals in continuous audio recordings using hidden Markov models. *IEEE Transactions on Biomedical Engineering* 2006; 53(6): 1078–1083.
44. Matos S, Birring SS, Pavord ID, Evans DH. An automated system for 24-h monitoring of cough frequency: the leicester cough monitor. *IEEE Transactions on Biomedical Engineering* 2007; 54(8): 1472–1479.
45. Sangaiah AK, Dhanaraj JSA, Mohandas P, Castiglione A. Cognitive IoT system with intelligence techniques in sustainable computing environment. *Computer Communications* 2020; 154: 347–360.
46. Sodhro AH, Pirbhulal S, Sangaiah AK. Convergence of IoT and product lifecycle management in medical health care. *Future Generation Computer Systems* 2018; 86: 380–391.
47. Barata F, Kipfer K, Weber M, Tinschert P, Fleisch E, Kowatsch T. Towards device-agnostic mobile cough detection with convolutional neural networks. In: IEEE. ; 2019: 1–11.
48. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Association for Computing Machinery; 2016; New York, NY, USA: 785–794.

