



Neural Network Based Mental Depression Identification and Sentiments Classification Technique From Speech Signals: A COVID-19 Focused Pandemic Study

Syed Thouheed Ahmed¹, Dollar Konjengbam Singh², Syed Muzamil Basha³,
Emad Abouel Nasr^{4*}, Ali K. Kamrani⁵ and Mohamed K. Aboudaif^{4*}

¹ School of Computing and Information Technology, REVA University, Bengaluru, India, ² ICT for Internet and Multimedia, University of Padua, Padua, Italy, ³ School of Computer Science and Engineering, REVA University, Bengaluru, India, ⁴ Industrial Engineering Department, College of Engineering, King Saud University, Riyadh, Saudi Arabia, ⁵ Industrial Engineering Department, College of Engineering, University of Houston, Houston, TX, United States

OPEN ACCESS

Edited by:

Celestine Iwendi,
School of Creative Technologies
University of Bolton, United Kingdom

Reviewed by:

Parvaiz Ahmad Naik,
Xi'an Jiaotong University, China
Ebuka Ibeke,
Robert Gordon University,
United Kingdom

*Correspondence:

Emad Abouel Nasr
eabdelghany@ksu.edu.sa
Mohamed K. Aboudaif
maboudaif@ksu.edu.sa

Specialty section:

This article was submitted to
Digital Public Health,
a section of the journal
Frontiers in Public Health

Received: 23 September 2021

Accepted: 08 November 2021

Published: 06 December 2021

Citation:

Ahmed ST, Singh DK, Basha SM,
Abouel Nasr E, Kamrani AK and
Aboudaif MK (2021) Neural Network
Based Mental Depression
Identification and Sentiments
Classification Technique From Speech
Signals: A COVID-19 Focused
Pandemic Study.
Front. Public Health 9:781827.
doi: 10.3389/fpubh.2021.781827

COVID-19 (SARS-CoV-2) was declared as a global pandemic by the World Health Organization (WHO) in February 2020. This led to previously unforeseen measures that aimed to curb its spread, such as the lockdown of cities, districts, and international travel. Various researchers and institutions have focused on multidimensional opportunities and solutions in encountering the COVID-19 pandemic. This study focuses on mental health and sentiment validations caused by the global lockdowns across the countries, resulting in a mental disability among individuals. This paper discusses a technique for identifying the mental state of an individual by sentiment analysis of feelings such as anxiety, depression, and loneliness caused by isolation and pauses to the normal chains of operations in daily life. The research uses a Neural Network (NN) to resolve and extract patterns and validate threshold trained datasets for decision making. This technique was used to validate 2,173 global speech samples, and the resulting accuracy of mental state and sentiments are identified with 93.5% accuracy in classifying the behavioral patterns of patients suffering from COVID-19 and pandemic-influenced depression.

Keywords: sentiment extraction, speech signal processing, COVID-19, mental depression, neural network

INTRODUCTION

The world is at present facing an uncertain time due to the global pandemic caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), also known as COVID-19. The pandemic has forced nations to exercise lockdown as a preventive measure to slow the spread of the virus. The pandemic has resulted in economic failure and disruptions in the supply chain all over the world. There was a race for a vaccine among modern drug and research organizations. The pandemic has caused major adverse effects such as mental depression, isolation, anxiety, and loneliness besides respiratory disorders and aligned symptoms. Depression and other mental health issues have been caused by lockdown and restrictions in travel and work, with a new normal social life now conducted via technological platforms.

The pandemic has bought a sense of maturity and adverse implications concerning psychosocial behaviors and mental health implications such as depression, anxiety, and loneliness. In this research, a systematic evaluation was conducted on the behavior of users based on speech signals, which were recorded using a machine learning technique to extract keywords and classify data using sentimental analysis techniques (1). The research in this article also focuses on identifying the user’s mental state via the speech signals recorded over technological platforms used for virtual meetings and other gatherings (2). The research aims to provide schematic evaluation and validation approaches and classify patients based on medical conditions.

LITERATURE SURVEY

The global pandemic situation under COVID-19 has left traces of various adverse effects on people, which are the result of isolation, lockdown, mental health destabilization, and much more. The authors Singh et al. (3) and Naik et al. (4) have discussed the impact caused by COVID-19 and lockdown on the younger

generation, focussing on children’s behavior and reactions to the new normal. The study shows the overall implications of isolation on children and adolescents. Pfefferbaum and North (5) have discussed the impact and relationship of mental stress caused due to the global pandemic situation, with detailed insights into public health emergencies and the influence of the pandemic on looming health conditions. Furthermore, a discussion on the challenges faced by health care workers (HCW) and their state of mental stress is documented by Spoorthy et al. (6). The HCWs are frontline attributes and hence require assistance in evaluating and validating mental health via the main mode of communication now used, i.e., speech signals through digital media platforms and applications and a similar discussion is highlighted in other studies (7, 8).

Some studies are focused on the terms of technological solutions for the mental distress caused by the pandemic. These solutions have outlined the use of a telemedicine approach for reaching the maximum and remote population of a developing country like India. A study by Ahmed et al. (9) discussed

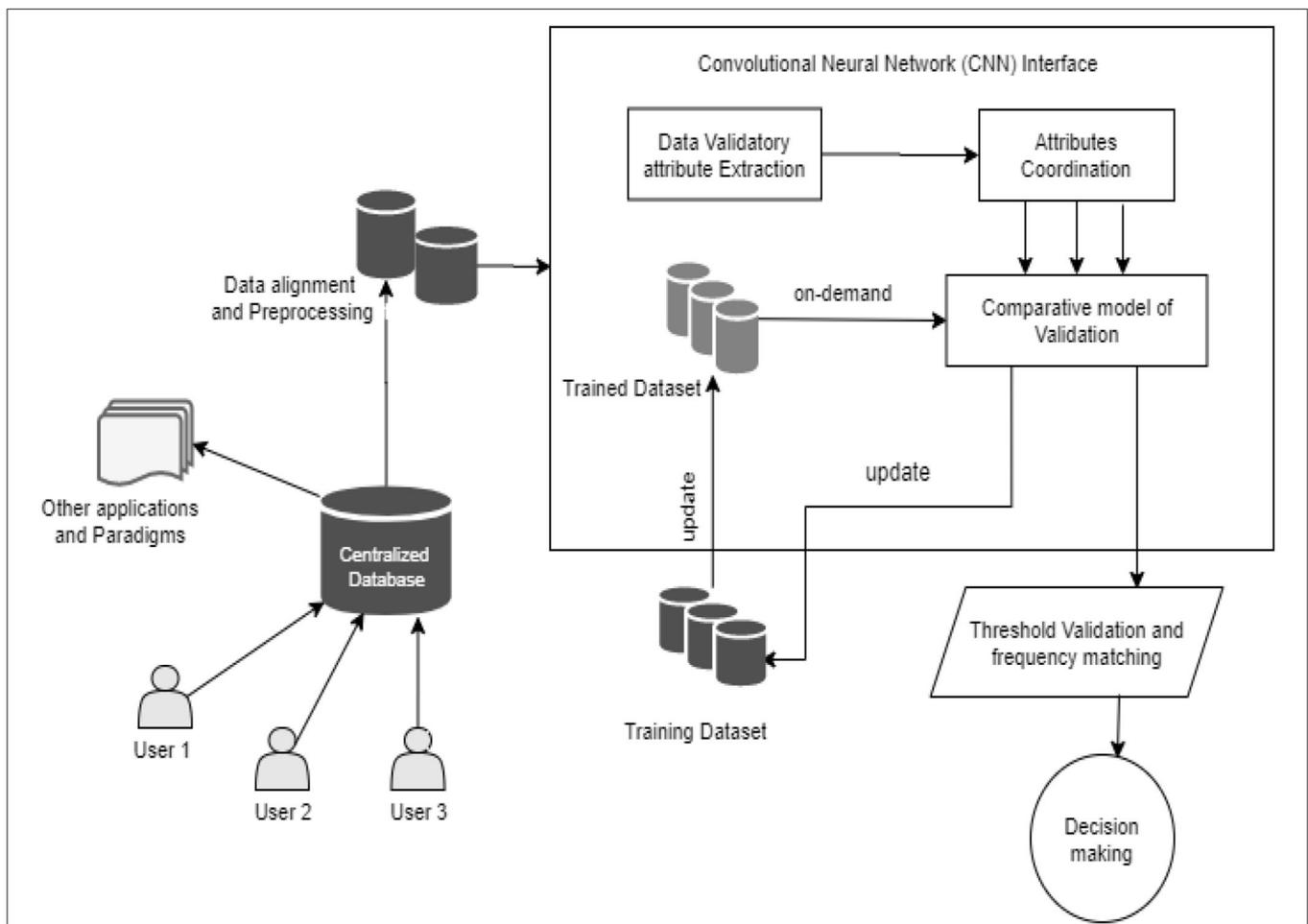


FIGURE 1 | Architectural diagram of the proposed technique toward decision making in speech signals.

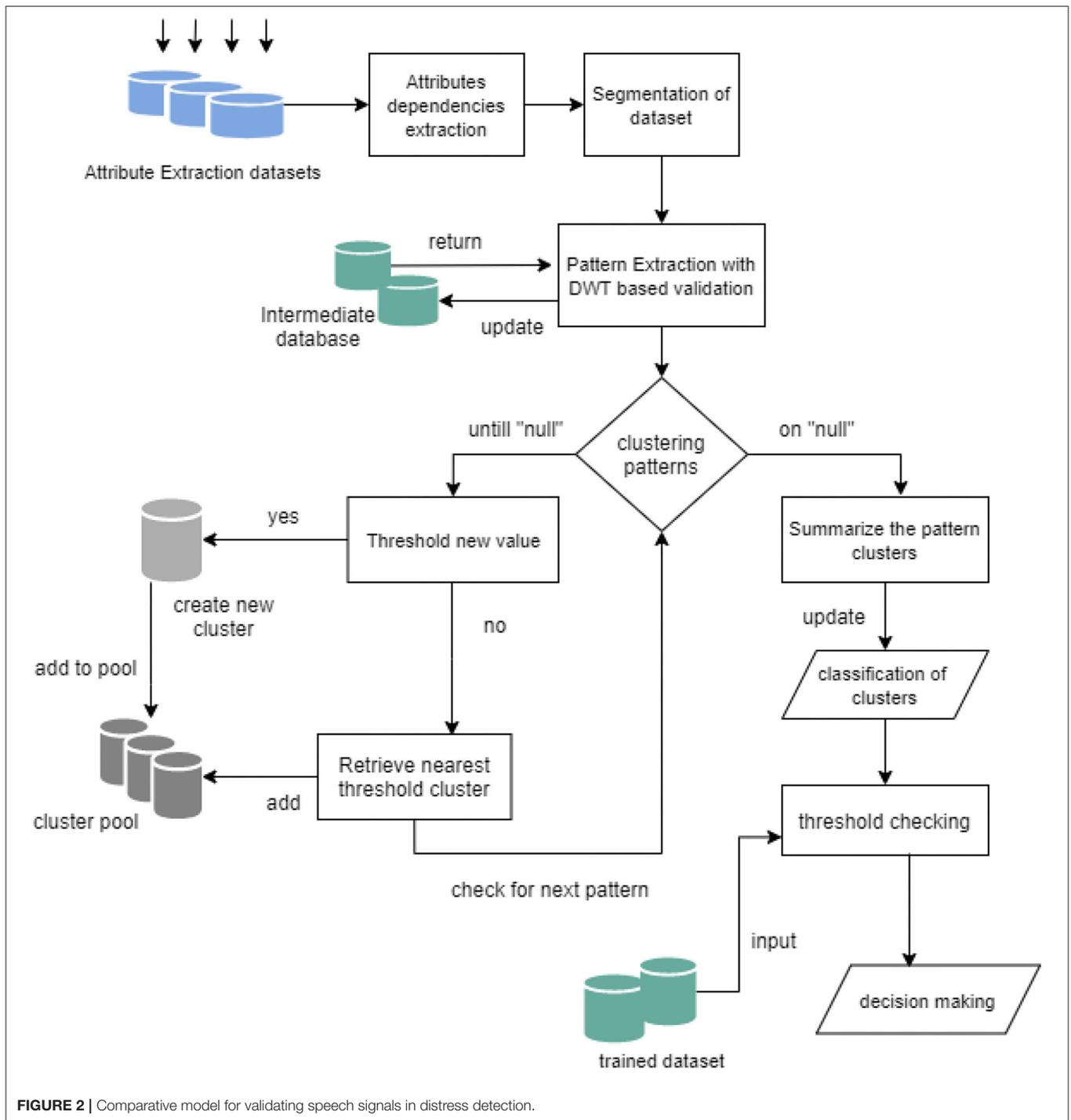
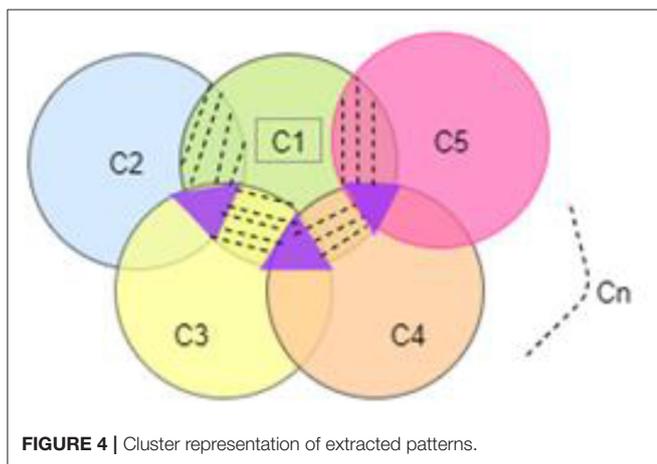
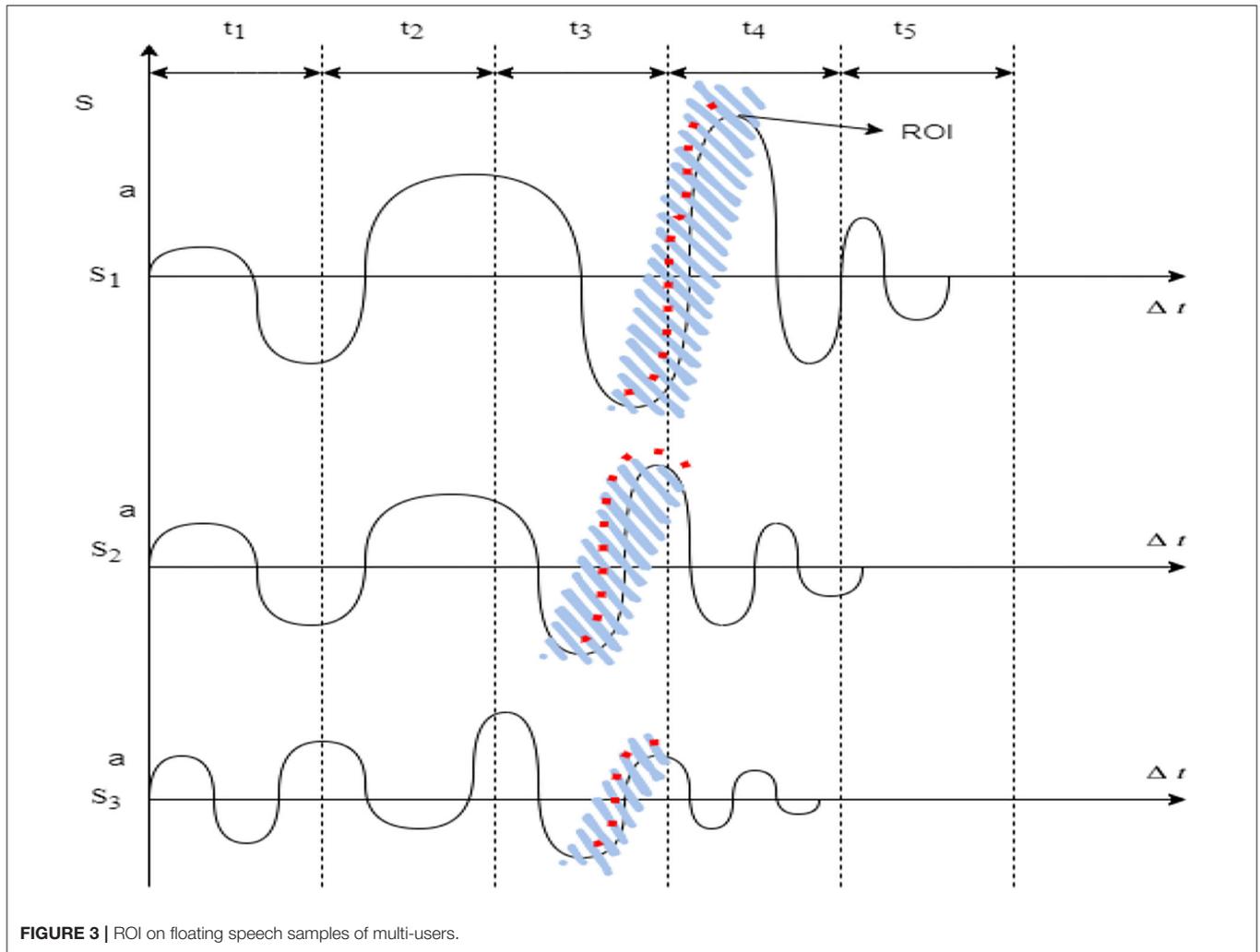


FIGURE 2 | Comparative model for validating speech signals in distress detection.

Multidimensional Optimal Medical Dataset processing under a telemedicine channel. These MooM datasets include a signal processing unit for a standardized approach and can be used for intimated processing in the proposed study, with a supported algorithm from (10). The method of detecting and validating speech signals is also proposed in this article, based on the influence of telemedicine approaches with a numerical clustering validation by (11) and (12).

The latest findings in the survey are recorded with real-time datasets as discussed in (4). This approach aims to validate treatment and handling, focusing on pandemic control and coordination. The prediction and modeling of the pandemic are discussed by Iwendi et al. (13) and Ngabod et al. (14), who propose a technique for classifying pandemic growth in smart cities. Under the process processing state, this dedicated networking model can be utilized, as discussed by Ahmed et al.,



under a dynamic user cluster grouping approach (15–17). These developments have provided a reliable solution for handling pandemic data using text mining and decision support. The

classification of Covid-19 studies and surveys are reported and validated by (18, 19).

METHODOLOGY

The proposed methodology aims to focus on the detection and validation of speech signals via a depression and mental disorder identification based on speech signal processing using a neural network (NN). The process is defined using mass datasets from 2,173 speech samples, as discussed in the architecture model in **Figure 1**. The agenda of the proposed technique is to restore a correlation with trained datasets in extracting and evaluating the samples of speech and classifying on demand. These speech signals are interdependent and have a higher order of distinction in recovering and validating the sample of COVID-19 patients' mental stability and sentiments (20).

The processing datasets are computed in a centralized database with user-to-user interface coordination, thereby generating a pool of databases consisting of raw and unprocessed data from the users. The process is initiated with data alignment

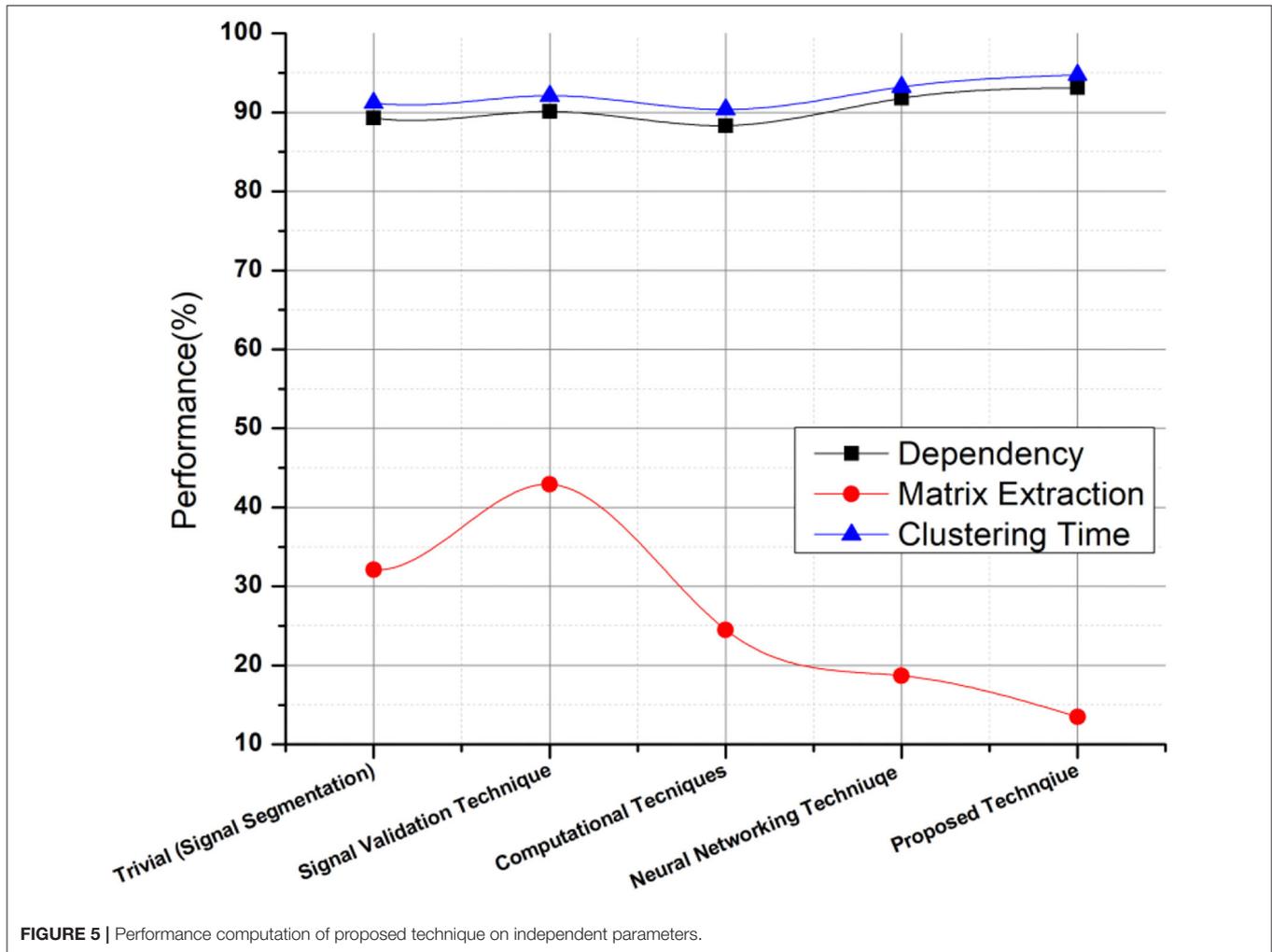


FIGURE 5 | Performance computation of proposed technique on independent parameters.

and pre-processing techniques, as discussed in the mathematical modeling of the proposed technique. The process is designed with a trained database of the speech signals with a heap address of thresholds relating to global attributes such as country, location, gender, age, and professional practice.

The trained datasets provide the threshold process for the extracted attributes of the user input signals. The process is designed with a comparative validation model to assure the process execution, as demonstrated in Figure 2. The comparative model evaluates the detailed execution process, such as the pattern extraction and clustering of signal samples in the form of JPEG intermediate files and a dedicated intermediate database, generating a cluster pool for segregation of samples based on ROI as demonstrated in Figure 3. This results in a threshold value comparison and thus provides a single decision and classification of the user’s mental condition.

MATHEMATICAL REPRESENTATION

The computation of the speech signals and detection of mental stress is achieved under the processed instruction architecture,

as demonstrated in Figure 1. The process aims to validate the signals into coordination datasets with a synchronization approach of proving learning and pooling clusters of similar patterns, as demonstrated in Figure 3. The mathematical approach is discussed in this section.

Attribute Extraction and Dependencies Validation

Consider a dataset (D) with a raw calibrated ecosystem of attributes (A) where each of attributes $A=\{A_1,A_2,A_3,\dots,A_n\}$ such that each attribute (A_i) resembles the paradigm of operation, as in Equation (1).

$$A_i = \int_0^\infty \frac{\partial (A_i)}{\partial (D)} \tag{1}$$

Where, each of the i^{th} attributes, has a correlated paradigm of operation and process extraction. Thus, the extracted attributes (A_e) are as shown in Equation (2).

$$A_e = \int_0^\infty \left(\Delta D_z \cdot \sum_{i=0}^n \frac{\partial A_i}{\partial D} \right) \cdot \Delta T \tag{2}$$

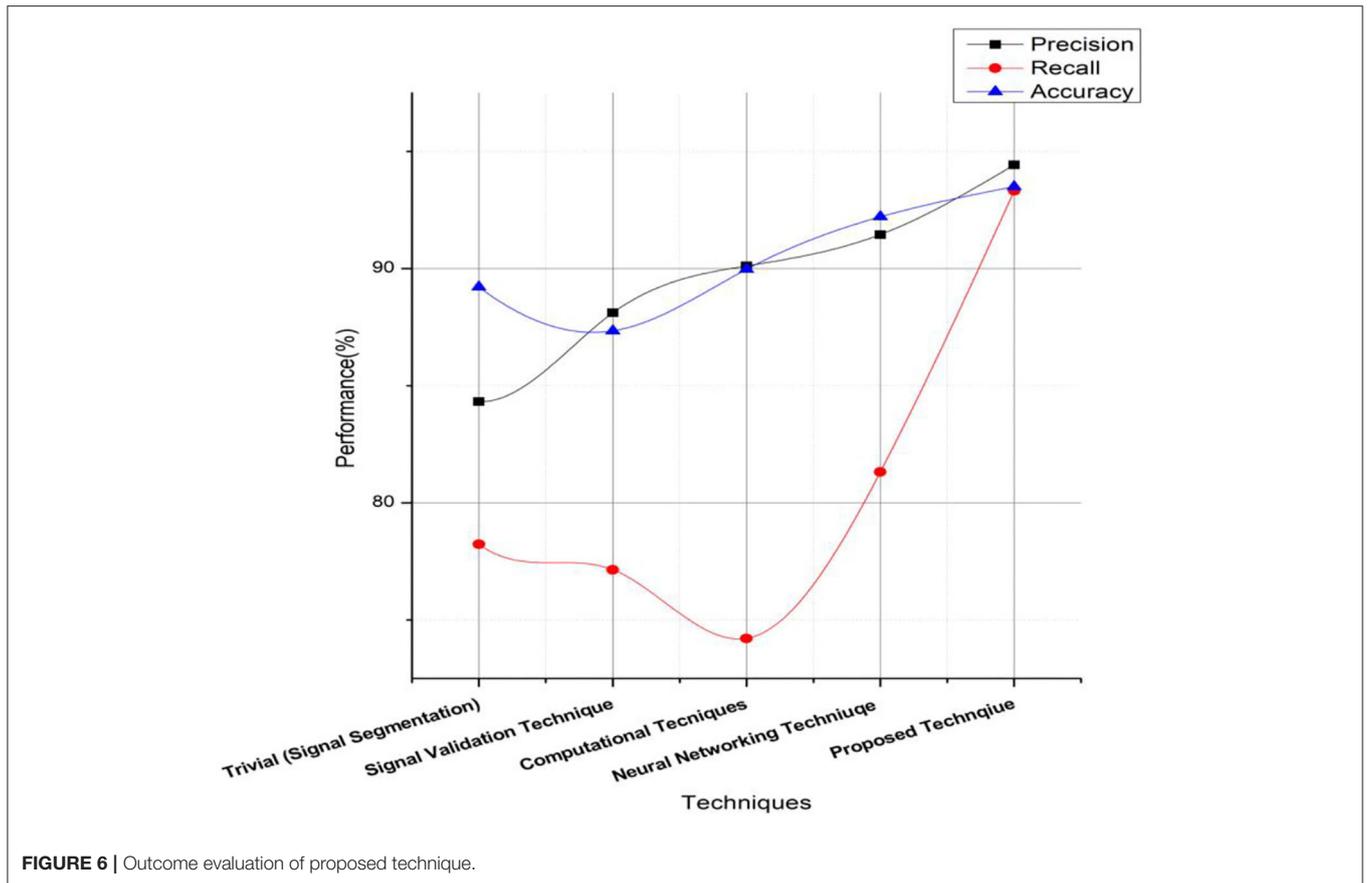


FIGURE 6 | Outcome evaluation of proposed technique.

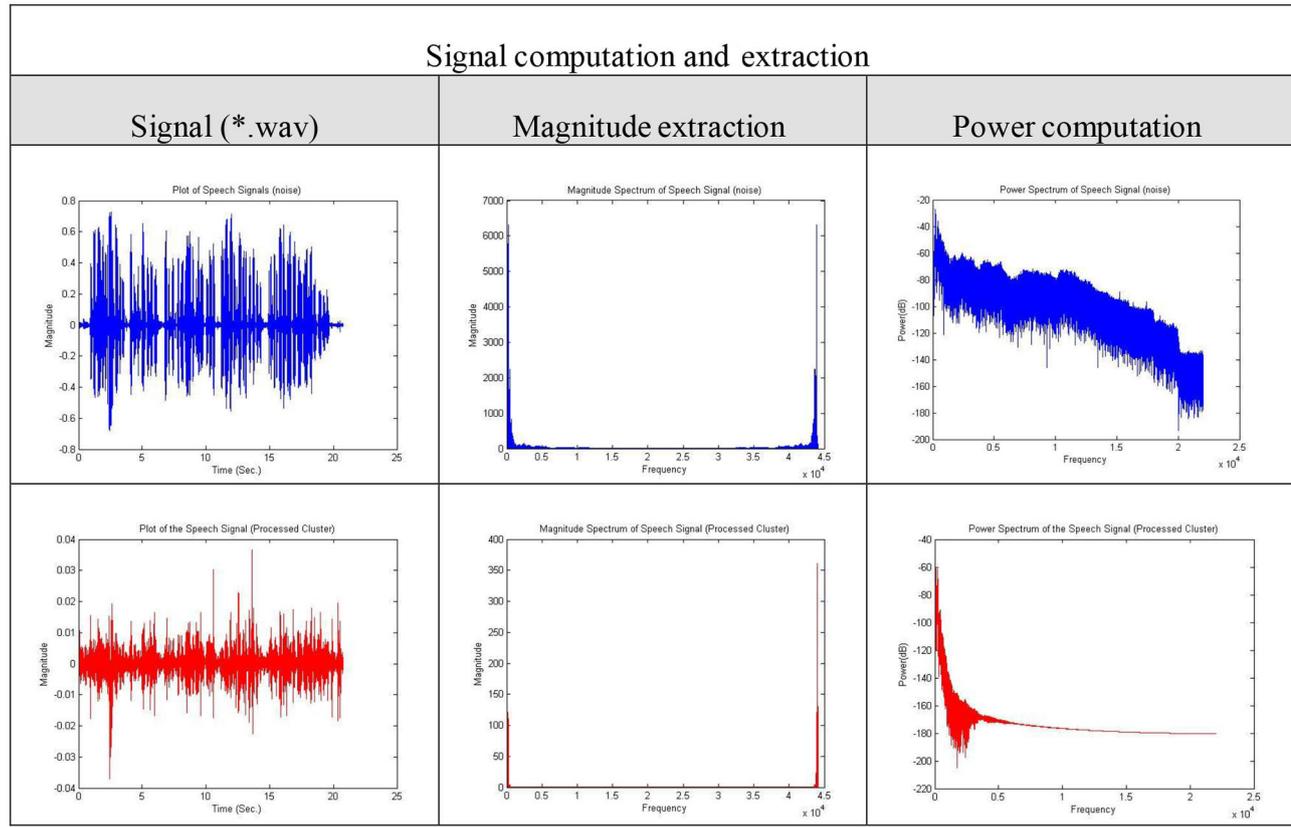
TABLE 1 | Decision support and evaluation parameters.

Patient relevance (sentiments)	Cluster range	Occurrence difference (Avg. sec)	Pattern differences	Decision accuracy (%)
COVID-19 (positive)	Average	3.211	0.327	97.23
Post COVID-19 (positive)	High	1.432	0.129	94.92
Loneliness	High	1.328	0.091	96.91
Anxiety	Average	2.114	0.181	92.17
Depression	High	0.994	0.021	97.28
Normal (Non-COVID19)	Low	5.251	0.448	95.39

TABLE 2 | Performance matrix for speech signal in mental distress validation.

Sample (Age group) (yrs)	Approach	Precision (%)	Recall (%)	Detection-score (%)	Accuracy (%)
5-9	Unclassified	92.43	89.23	89.3	90
10-15	Unclassified	95.72	93.22	93.08	95
16-25	Classified	97.12	97.48	98.1	98.48
26-35	Classified	97.78	97.53	98.9	98.97
36-55	Classified	98.3	97.12	99.12	99.23
56-60	Classified	95.33	91.23	97.12	94.19
61-75	Unclassified	96.12	89.23	91.07	88.32
76-100	Unclassified	92.31	88.7	87.2	81.2

TABLE 3 | Signal processing and analysis appendix.



Where the extracted attribute (A_e) is processed over the raw attributes set, in extracting the most relevant threshold attributes such as the peak frequency of a word or a repeated phrase of a sentence with a dilution of ΔD_z and mapping with ΔT as a threshold paradigm in validating all processing attributes (A_e) in the speech signal.

Segmentation of Samples

Samples are primarily divided into extracted attributes (A_e) sets, such that each of the attribute Region of Interest (ROI) is highlighted and marked in the entire speech signal, as shown in **Figure 3**.

Consider the segmentation (S) of the overall input signal (speech signal) with a highlighted extracted attribute (A_e). On consideration, each attribute in the signal has an occupancy time (Δt) in operating, and thus, a reflective ratio of division is processed based on CNN's evaluation paradigms.

The signal (S) of an independent sample (S_i) tends to occur in ROI in an independent location of the time matrix (Δt). Hence, the segmentation of signal (S) is as shown in Equation (4).

$$S = \frac{2\pi}{\Delta R} \int_0^\infty \left[\int_n^i \left(\frac{\Delta S_i}{\Delta t} \right) \right] \cdot \frac{\Delta t}{\Delta s} \quad (3)$$

$$\therefore S = \frac{2\pi}{\Delta R \cdot \Delta S} \left(\int_0^\infty \left(\frac{\partial S_i}{\partial t} \right) \cdot \frac{\partial (A_e)}{\partial t} \right) \cdot \Delta t \quad (4)$$

Where each signal strength is measured in ΔR with a signal time Δt , for all regional attributes extraction; hence, for segmentation to be processed completely, the schematics of each attribute signal strength (ΔR) is then computed with an exhausted peak of ROI from the signal as shown in **Figure 3**.

Pattern Extraction and Schema Alignment of Datasets

The process of pattern extraction is calibrated with the internally divided segments of datasets. These datasets are processed with a frequency (f) such that the internal segments (S) = $\{S_1, S_2, S_3, \dots, S_n\}$ has calibrated frequencies (f) = $\{f_1, f_2, f_3, \dots, f_n\}$. Thus, the inter-correlated frequencies can be defined and associated as (S_f) = $\{S_{f1}, S_{f2}, S_{f3}, \dots, S_{fn}\}$.

The process of pattern with speech signals is internally correlated to the amplitude of the signal (amp). Where it is represented as $f_{amp} = \{f_{amp1}, f_{amp2}, f_{amp3}, \dots, f_{ampn}\}$. The amplitude of each frequency feature can be represented and extracted as shown in Equation (5).

$$S_i(n) = \int \left\{ \sum_{k=0}^n [\alpha_i(k) \cdot amp(f)] \right\} \quad (5)$$

Where each signal pattern (S_n) represents the overall coordination in speech signals, and the ‘’ represents band filters of the speech signal with a coefficient of amplitude and frequency. On extraction of patterns from those correlated in Equation (5), the frequency patterns can be sorted by independent bandwidth as shown in Equation (6).

$$P_i = \frac{2\pi}{\Delta R.\Delta S} \left\{ \left(\log(G(f)) \cdot \int_0^\infty \frac{\partial(S_i(n))}{\partial t} \right)^2 \right\} \quad (6)$$

$$P_i = \frac{2\pi}{\Delta R.\Delta S} \left\{ \left(\log(G(f_{amp})) \right)^2 \cdot \int_0^\infty \frac{\partial^2(S_i(n))}{\partial t^2} \right\} \quad (7)$$

The ‘ P_i ’ on Equation 7 is the pattern of repeated learning from the CNN framework. The internal arrangements can be represented as the frequency (f) under the operation of amplitude, (mode) is represented as f_{amp} , further graded into the Gaussian constant (G). The process in Equation (7) is then concluded, as shown in Equation (8).

$$P_i = \frac{2\pi}{\Delta R.\Delta S} \left\{ \left(2\log(G(f_{amp})) \cdot \int_0^\infty \frac{\partial^2 \left(\sum_{j=0}^n \sum_{k=0}^n \alpha_i(j) \cdot amp(f_k) \right)}{\partial t^2} \right) \right\} \quad (8)$$

$$P_i = \frac{4\pi}{\Delta R.\Delta S} \left\{ \left(\log(G(f_{amp})) \cdot \int_0^\infty \frac{\partial^2 \left(\sum_{j=0}^n \left(\sum_{k=0}^n \alpha_i(j) \cdot amp(f_k) \right) \right)}{\partial t^2} \right) \right\} \quad (9)$$

$$P_i = \frac{4\pi \cdot \log G(f_{amp})}{\Delta R.\Delta S} \cdot \left\{ \left(\int_0^\infty \frac{\partial^2 \left(\sum_{j=0}^n \left\{ \sum_{k=0}^n \left(\alpha_i(j) \cdot amp(f_k) \right) \right\} \right)}{\partial t^2} \right) \right\} \quad (10)$$

Thus, Equation (10) represents the coordinates of the pattern extracted and validated for the pattern with respect to a segment (S_i). Thus, on summarisation, the representation can be as $P = \{P_1, P_2, P_3, \dots, P_n\}$ co-related to coordination of segment as $S_p = \{S_{p1}, S_{p2}, S_{p3}, \dots, S_{pn}\}$, where ‘ n ’ is the last segment of given input signal.

Clustering and Classification of Datasets

Equation (10) retrieves the pattern of individual segments, and thus the coefficient of such segments are summarized and represented in $S_p = \{S_{p1}, S_{p2}, S_{p3}, \dots, S_{pn}\}$. Hence the clustering is shown in Figure 4.

The cluster (C) is retained from a group of values and its corresponding coefficients for the value re-compensation. The clusters are internally evaluated with the focus of associating.

$$C_i = \int_0^\infty \left\{ \sum_{i=0}^n \left[\sum_{j=i}^n \left(\frac{\partial(P_i)_j}{\partial t} \right) \cdot \Delta T_i \right] \right\} \quad (11)$$

Where each cluster (C_i) is validated with a corresponding pattern coefficient and a threshold value (ΔT). The internal Threshold value (ΔT) is validated and evaluated. In summary, the clusters (C) = $\{C_1, C_2, C_3, \dots, C_n\}$. These clusters have an association of common patterns, for example, represented as $\{(C_i \cap C_j) \cap C_k\}$, and these associations are subjected to attribute validation, as shown in Figure 4.

Threshold Validation and Decision Making

The clusters and classification of speech signals using clusters are validated and approved for processing into decision making. The decision-making approach is termed with a threshold value consultation, i.e., the overall technique extracts the validated pattern coefficient and thereby synchronizes it with a relatively more and likely approach of matching and schema validation. The proposed approach typically validates the decision of signal segmentation using the threshold value toward segregating the dataset of speech signals based on emotions. These emotion-based evaluations are rather computational, and hence a most likely decision is processed.

RESULTS AND DISCUSSIONS

The proposed technique has successfully retrieved the signal attributes and the prediction ratio for evaluation. The input signals from the users via a remote connecting platform are uploaded to a centralized database in a cloud computing ecosystem using AWS-sponsored services. The datasets are processed and validated according to a multidimensional approach.

The variation of predicting the sentiments is based on the information designed and developed via clustering datasets. The prediction ratio is summarized in Figures 5, 6, respectively, with a comparative evaluation from previous systems. Table 1 shows the parameters related to the mental stress and paradigms to provide decision support. The table highlights the evaluation parameters such as the occurrence delay of a keyword in clustering, as shown in Equation (11). The supported approach thus classifies the pattern of these keyword occurrence sequences for decision making.

The results of data/signal processing and decision-making are shown in Table 2. The results show promising outcomes in proving a precision of 90% and higher in various users across the language and location. The results of processing a single sample are included in Table 3. The processing signal magnitude and the power spectrum computation demonstrate a higher order of signal clarity in analysis and validation.

CONCLUSION

The technique proposed in the present study uses neural networking terminology to learn and develop a pool of clusters and patterns to provide a systematic and reliable decision to categorize speech signals. The processing system is based on open database processing to validate the mental health conditions of users during the ongoing isolation and lockdowns caused

by the COVID-19 pandemic. The results show a promising outcome with a precision of 90% and higher accuracy across various users. The approach has a projected accuracy of 93.5% under the open validation platform on a computational evaluation. The proposed technique could be included in classifying and categorizing patients' behavior in future, with supervised approaches to keyword extraction and classification in dynamic signals.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

REFERENCES

- Ahmed ST, Basha SM, Arumugam SR, Kodabagi MM. *Pattern Recognition: An Introduction*. MileStone Research Publications (2021).
- Lin C, Ibeke E, Wyner A, Guerin F. Sentiment–topic modeling in text mining. *Wiley Interdisc Rev.* (2015) 5:246–54. doi: 10.1002/widm.1161
- Singh S, Roy MD, Sinha CPT, Parveen CPTMS, Joshi CPTG. Impact of COVID-19 and lockdown on mental health of children and adolescents: a narrative review with recommendations. *Psychiatry Res.* (2020) 2020:113429. doi: 10.1016/j.psychres.2020.113429
- Naik PA, Owolabi KM, Zu J, Naik MUD. Modeling the transmission dynamics of COVID-19 pandemic in caputo type fractional derivative. *J Multisc Model.* (2021) 20:1–20. doi: 10.1142/S1756973721500062
- Pfefferbaum B, North CS. Mental health and the Covid-19 pandemic. *N Engl J Med.* (2020) 383:510–2. doi: 10.1056/NEJMp2008017
- Spoorthy MS, Pratapa SK, Mahant S. Mental health problems faced by healthcare workers due to the COVID-19 pandemic—A review. *Asian J Psychiatry.* (2020) 51:102119. doi: 10.1016/j.ajp.2020.102119
- Cullen W, Gulati G, Kelly BD. Mental health in the Covid-19 pandemic. *QJM Int J Med.* (2020) 113:311–2. doi: 10.1093/qjmed/hcaa110
- Iwendi C, Mahboob K, Khalid Z, Javed AR, Rizwan M, Ghosh U. Classification of COVID-19 individuals using adaptive neuro-fuzzy inference system. *Multi Syst.* (2021) 1–15. doi: 10.1007/s00530-021-00774-w
- Ahmed ST, Sankar S, Sandhya M. Multi-objective optimal medical data informatics standardization and processing technique for telemedicine via machine learning approach. *J Amb Intell Hum Comput.* (2020) 12:5349–58. doi: 10.1007/s12652-020-02016-9
- Ahmed ST, Sandhya M, Sankar S. An optimized RTSRV machine learning algorithm for biomedical signal transmission and regeneration for telemedicine environment. *Proc Comp Sci.* (2019) 152:140–9. doi: 10.1016/j.procs.2019.05.036
- Ayoub A, Mahboob K, Javed AR, Rizwan M, Gadekallu TR, Abidi MH, et al. Classification and categorization of covid-19 outbreak in Pakistan. *Comput Mater Continua.* (2021) 69:1253–69. doi: 10.32604/cmc.2021.015655
- Kumar SS, Ahmed ST, Vigneshwaran P, Sandeep H, Singh HM. Two phase cluster validation approach towards measuring cluster quality in unstructured and structured numerical datasets. *J Ambient Intell Hum Comput.* (2020) 12:7581–594. doi: 10.1007/s12652-020-02487-w
- Iwendi C, Bashir AK, Peshkar A, Sujatha R, Chatterjee JM, Pasupuleti S, et al. Jo O. COVID-19 patient health prediction using boosted random forest algorithm. *Front Public Health.* (2020) 8:357. doi: 10.3389/fpubh.2020.00357

AUTHOR CONTRIBUTIONS

Conceptualization and writing: SA, DS, SB, EA, and MA. Methodology: SA, DS, SB, AK, and MA. Investigation and programming: SA and DS. Resources: SB, EA, MA, and AK. Review: EA and AK. All authors contributed to the article and approved the submitted version.

FUNDING

The authors extend their appreciation to King Saud University for funding this work through Researchers Supporting Project number (RSP-2021/164), King Saud University, Riyadh, Saudi Arabia.

- Ngabo D, Dong W, Ibeke E, Iwendi C, Masabo E. Tackling pandemics in smart cities using machine learning architecture. *Math Biosci Eng.* (2021) 18:8444–61. doi: 10.3934/mbe.2021418
- Ahmed ST, Sandhya M, Sankar S. TelMED: dynamic user clustering resource allocation technique for MooM datasets under optimizing telemedicine network. *Wireless Pers Commun.* (2020) 112:1061–77. doi: 10.1007/s11277-020-07091-x
- Al-Shammari NK, Syed TH, Syed MB. An Edge–IoT framework and prototype based on blockchain for smart healthcare applications. *Eng Technol Appl Sci Res.* (2021) 11:7326–31. doi: 10.48084/etasr.4245
- Ahmed ST and Sankar S. Investigative protocol design of layer optimized image compression in telemedicine environment. *Proc Comput Sci.* (2020) 167:2617–22. doi: 10.1016/j.procs.2020.03.323
- Bhattacharya S, Maddikunta PKR, Pham QV, Gadekallu TR, Chowdhary CL, Alazab M, et al. Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey. *Sustain Cities Soc.* (2021) 65:102589. doi: 10.1016/j.scs.2020.102589
- Dhanamjayulu C, Nizhal UN, Maddikunta PKR, Gadekallu TR, Iwendi C, Wei C, et al. Identification of malnutrition and prediction of BMI from facial images using real-time image processing and machine learning. *IET Image Process.* (2021) doi: 10.1049/IPR2.12222
- Reddy PK, Reddy TS, Balakrishnan S, Basha SM, Poluru RK. Heart disease prediction using machine learning algorithm. *Int J Innov Technol Expl Eng.* (2019) 8:2603–6. doi: 10.35940/ijitee.J9340.0881019

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Ahmed, Singh, Basha, Abouel Nasr, Kamrani and Aboudaif. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.