

A Pragmatic Approach for Detecting nCOVID-19 using Pervasive Computing Based on Dual Diagnostic Measures

Swarnava Biswas¹, Chandranath Chakraborty^{1,*}, Riddhi Chawla², Dabosmita Paul³, Debajit Sen⁴, Niloy Sarkar¹ and Moumita Mukherjee^{5,*}

¹The Neotia University, Kolkata, West Bengal, India

²Medical School, Akfa University, Tashkent, Uzbekistan

³Mendel University, Brno, Czech Republic

⁴Robert Bosch Engineering and Business Solutions, Bangalore, Karnataka, India

⁵Department of Physics, School of Basic and Applied Sciences, Adamas University, Kolkata, West Bengal, India

Abstract: Our regular way of life has been disrupted by the COVID-19, and we have been obliged to accept the procedures that are in place under the new normal regime. It is envisaged that the standard diagnostic technique will evolve throughout the course of the procedure. As a help to this type of diagnostic technique, our research group is developing a tool. In this article, the group discusses the importance of employing two diagnostic metrics that have proven to be pivotal in many diagnoses for doctors, and how they might be used to their advantage. Together, natural language processing-based symptoms measures and a machine learning-based strategy that takes into account medical vitals can help to minimise the error percentage of detection by as much as 50%. The technique suggested in this study is the first of its type, and the authors have obtained findings that are satisfactory in terms of accuracy. A further justification for suggesting such a strategy is the manner in which a fusion algorithm might arrive at the correct results from two concurrent algorithms performing the same task. One of the group's other objectives was to give the doctor a valuable opinion in the form of such an architectural design. The suggested design may be employed at any point of care facility without the need for any additional infrastructure or escalation of the current amenities to accommodate the proposed architecture.

Keywords: Natural Language Processing, Text Analytics, Support Vector Machine, Ensemble Learning, Body Vitals, Symptoms, COVID-19 Detection, COVID-19 Cluster Identification, Decision Fusion.

1. INTRODUCTION

It is estimated that more than fifty lakhs people have been killed by COVID-19 [1] since it was first discovered. Each government has issued an open notification to maintain social separation, the use of masks, and the use of hand sanitisers in order to combat the crisis [2]. In any event, it was extremely difficult to raise awareness among the whole public, particularly in nations where the population comes from a wide range of cultural and linguistic origins. The most difficult problem that any government had to deal with was how to conduct random testing, identify the ill, and separate them from the rest of the population [3]. The government should have the necessary infrastructure in place to provide competent medical treatment and assistance to those who have been affected by the disaster. As a responsible citizen, it is becoming increasingly important to put into practise the best preventative measures that have been identified or

expressed in order to bring this transmission rate to a close. It is also critical to understand the first week of symptoms since the same is extremely significant given the fact that the first vaccination to enter the market will not be available for another few months at this point. According to different publications, research, and ongoing public awareness campaigns, it is already recognised that the first week of infection might be the most telling period in determining the potentiality of a viral strain [4].

A thermal gun is being used to screen visitors in any business or residential building, as well as in healthcare facilities, and the individual is being asked to complete a series of questionnaires. The procedure can assist in the identification of some questionable instances at the point of entry. Our goal is to eliminate the human aspect from the equation and, as a result, to decrease the amount of manual mistakes. Based on temperature [5] and oxygen saturation percentage (SpO₂) [6], we utilised an AI-based classification algorithm that we trained on the provided data set to distinguish between COVID patients and 'Normal' patients in this study.

*Address correspondence to these authors at the Neotia University, Kolkata, West Bengal, India; E-mail: chandranath.chakraborty@tnu.in

Department of Physics, School of Basic and Applied Sciences, Adamas University, Kolkata, West Bengal, India; E-mail: moumita.mukherjee@adamasuniversity.ac.in

Natural Language Processing (NLP) is a subfield of computer science that is primarily concerned with linguistics and Artificial Intelligence (AI) models [7-11]. Natural language processing (NLP) is used to understand and process the contributions of people's natural language for a range of applications [8-11]. NLP is used to comprehend and process their contributions for a number of applications. To train the algorithms, millions of samples of text words, phrases, and paragraphs produced by people are studied using different Machine Learning training methods [12].

In this research, the authors provide for the first time a classification solution based on Natural Language Processing that provides us with an early diagnostic knowledge of the existence of COVID-19. This straightforward tool, when used in conjunction with the AI-enabled body sensor-based diagnostic analysis, serves as the initial step in the screening procedure.

The machine learning model created on medical vitals such as temperature and saturation percentage of oxygen (SpO₂) for diagnosing the existence of the virus strain serves as the foundation for our software design. By utilising text analytics, we may include extra information from symptomatic data such as the existence of diabetes, a lung problem, or a fever into the design of our software architecture. Our software design, when combined with other factors, predicts the existence of the virus in the people who have been infected.

The implications of such a development are numerous and significant. According to the findings of several other research, the majority of patients present to hospitals at a later stage of the illness, resulting in a dismal prognosis. The suggested tool is therefore extremely vital for routine detection and isolation of patients, which will ultimately result in a better prognosis for those suffering from the condition. We have developed a model to identify COVID-affected individuals based on their pre-existing and current medical illnesses, as well as their various bodily vitals. We have also classified the subjects into clusters ranging from moderate to severe in severity. Six major subgroups of infections were identified for our investigation, which was ranked according to their criticality.

2. MATERIALS AND METHODOLOGY

The authors' goal is to enable all diagnostic centres, including those without the necessary equipment to do

RT-PCR testing, to swiftly identify and isolate the afflicted persons by utilising the intelligent architecture that has been outlined. It is possible to forecast the beginning of a viral infection in a patient by taking a sample of various symptoms detected during the first week and combining them with other body sensor readings. Our AI-based architecture may also be able to assist in determining the real possibility of COVID-19 becoming harmful and a patient being admitted to the hospital as a result of the incident.

2.1. Data Collection and Pre-Processing

North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, India, provided both qualitative and quantitative information in the form of its data collection efforts. Separation of the aggregated data into two different data sets according to the medical vitals and symptoms was accomplished. The two sets were then subdivided into two groups: the training set and the test set. Both training sets were labelled with pre-defined labels, however, the test set did not have any pre-defined labels connected with it. Pre-processing steps included cleaning the data with a normalisation approach and generating the missing values for the quantitative set as a part of the overall process. When it came to the qualitative data, the writers used a different method. The data set was entirely cleaned up by removing all of the missing values. The fact that both data sets were to be utilised independently for training the respective models meant that any disparity in size between the two data sets would not result in any ambiguity in the final model.

2.2. A Machine Learning Model for COVID Detection that is Based on Natural Language Processing

To appropriately identify between COVID and healthy patients, this stage comprises the use of a text analytics algorithm that is based on symptoms such as the existence of diabetes, respiratory distress, or a history of pneumonia.

Although COVID-19 is often characterised by symptoms such as fever, cough, and shortness of breath, the illness can manifest itself with stomach discomfort and other symptoms as well as being asymptomatic in certain cases [13-15]. In other studies of individuals who had a positive test for COVID-19, the overall death rate was 2 to 3 percent [16]. The severity of the illness might range from minor to serious. The relevance of text analytics can assist us in categorising people based on their likelihood of developing the

disease. Despite the fact that just a very preliminary suggestive estimation is achieved in the first screening stage, it obviously serves as an enabler for some of the confirmatory tests performed in the subsequent and subsequent screening phases.

Using symptomatic text descriptions from patient reports to classify patients and completing an initial level classification using a conventional machine learning model known as Support Vector Machine (SVM) [17], we have included a unique approach to classification in our suggested technique.

Text data is naturally organised in a sequential fashion. A piece of text is a collection of words that may or may not be related to one another in any way. Support vector machines were utilised to learn and categorise a sequence of text data, and the relationships between the text data sequences were retrieved using the tokenized document and Bag of words concepts.

The following are the four stages that were completed in order to train with the SVM model:

1. Import the data that has been pre-processed.
2. Using the tokenized document and bag of words, convert the words into numeric sequences.
3. Develop and train an SVM model.
4. Classify fresh symptomatic text data using the SVM model that was previously trained.

2.3. Comparison and Contrast of the Support Vector Machine Model Described above with Two Deep Learning Models

The Long Short-Term Memory (LSTM) network [18] is a Deep Learning-based model for classifying text descriptions that is widely used. An LSTM network is a form of Recurrent Neural Network (RNN) that can learn long-term dependencies between time steps of sequence data [19]. It is a type of RNN that can learn long-term dependencies between time steps of sequence data. Text input to an LSTM network must be transformed into numeric sequences before it can be used. This may be accomplished through the use of a word encoding system that converts documents into sequences of numeric indices. In order to achieve better outcomes, it is possible to incorporate a word embedding layer into the current network. Word embeddings, as opposed to scalar indices, map words

in a lexicon to numeric vectors instead of scalar indices. Because these embeddings preserve semantic characteristics of the words, words with comparable meanings will have vectors that are similar to one another.

The following are the four steps involved in training using the LSTM network:

1. Import the data that has been pre-processed.
2. Convert the words into numeric sequences by employing a word encoding technique.
3. Develop and train an LSTM network with a word embedding layer, as shown in Figure 3.
4. Classify fresh text input using the LSTM network that has been trained.

The convolutional neural network, or CNN for short, is another common model for classifying text input that has recently gained popularity [20].

Before convolutions can be used to categorise text data, the text data must first be translated into pictures. Specifically, the observations required to be padding or shortening so that they were all of the same length S , and the documents were to be turned into sequences of word vectors of length C using a word embedding technique. Image representation of the document must be in the form of a 1-by- S -by- C grid (an image with height 1, width S , and C channels).

The network is trained with 1-D convolutional filters of variable widths, which are used to train the network. When a filter is applied, its width correlates to how many words it is capable of detecting (the n -gram length). Because the network has many branches of convolutional layers, it can accommodate a variety of n -gram lengths.

The CNN Architecture is described in detail in the following steps:

1. 1 by S by C input size, where S denotes the length of the sequence and C indicates the number of features (the embedding dimension).
2. In step 2, blocks of layers are built for the n -gram lengths 2, 3, 4, and 5, each of which has a convolutional layer, batch normalisation layer, a ReLU layer, a dropout layer, and a max-pooling layer.

3. For each block, 200 convolutional filters of size 1-by-N and pooling areas of size 1-by-S are defined, where N is the length of the n-grams in question.
4. Using a depth concatenation layer, the input layers of each block are joined and concatenated with the outputs of the blocks, completing the circuit.
5. In order to categorise the outputs, a fully connected layer with output size K, a softmax layer, and a classification layer are added, where K is the number of classes to be classified.

Figure 1 illustrates the network design.

2.4. Algorithm for Decision Fusion

Any Decision Fusion Algorithm [21] is critical in multi-hypothesis situations and is thus generally applicable to AI-assisted clinical diagnostic processes. Multi-hypothesis will operate as a beneficial second opinion for health practitioners, therefore enabling any diagnostic-based strategy. In our example, we developed a two-stage screening procedure for identifying and isolating COVID participants and their associated clusters based on symptoms and vital signs. To elaborate, our initial Verification step includes two sub diagnostic measures: sensor vitals and symptomatic examination of participants. This is more

of a preliminary study than a physical examination, which doctors often undertake as the first step in their diagnostic method. Due to the fact that we created a two-step verification from two distinct sets of data and derived the corresponding inferences from two distinct algorithms, we are proposing an algorithm based on a "Maximization Rule" and a probabilistic measure for correctly recognising the subject in decision conflict scenarios. Finally, our approach analyses the sets of inferences generated by the various AI engines and returns the output of the class that has been inferred the most.

Our methodology's equation is as follows:

C1 denotes the projected output class for problematic data by the machine learning method.

S1 denotes the probability score for the C1 class.

C2 denotes the output class for radiographic pictures predicted by the deep learning algorithm.

S2 denotes the output class's probability score. C2

Assume that X reflects the one output provided, rather than two distinct outputs generated by two distinct algorithms.

For all output classes, $X = \text{argmax}(S1, S2)$.

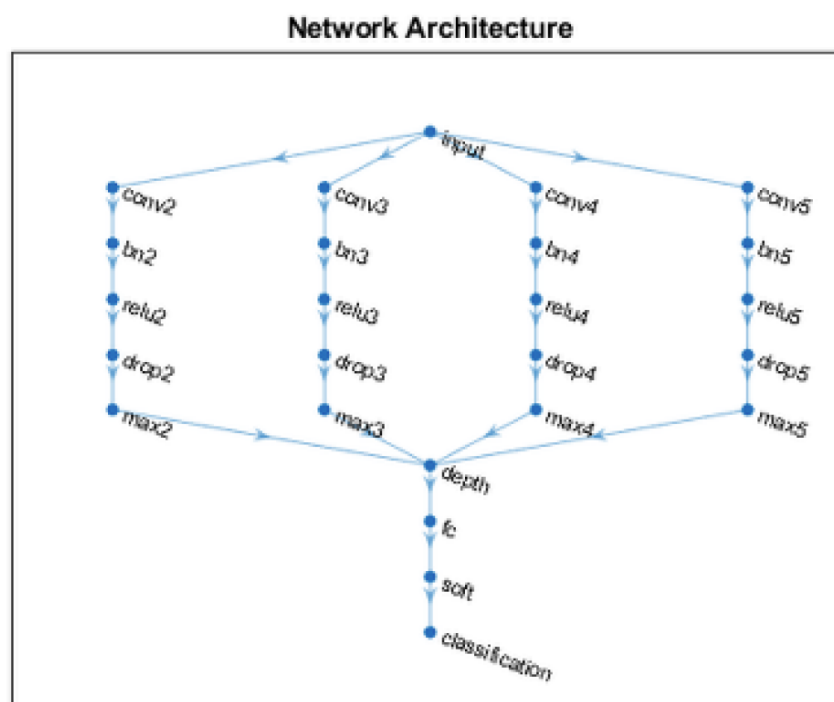


Figure 1: The Network Architecture for CNN Model.

The entire procedure is depicted in Figure 2.

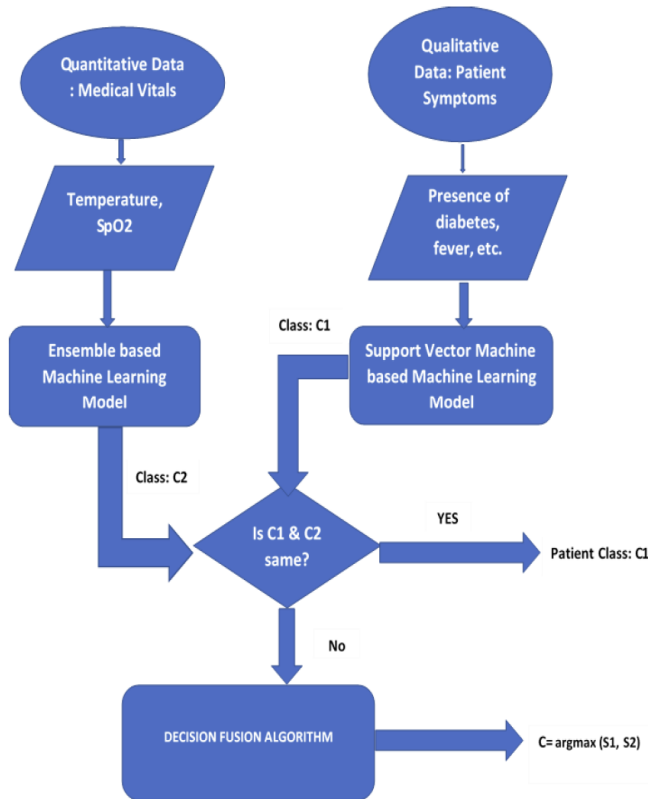


Figure 2: Flow Chart for Decision Fusion Algorithm.

2.5. A Machine Learning Model for COVID CLUSTER Identification Based on Natural Language Processing (NLP)

Once a COVID-19 subject has been found, the next logical step is to classify the subject into one of the six clusters based on the symptoms associated with each cluster, which is done in a systematic manner. To clarify further, the COVID cluster identification approach is a way by which the authors have attempted to contribute to the segmentation of patients into clusters based on the degree of symptoms they were experiencing.

The symptomatic text data was first translated into numeric sequences and then used as an input feed to test our support vector machine (SVM) trained model for further clustering of the COVID subjects before being used as an input feed for further clustering of the COVID subjects. This is accomplished by the use of word encoding, which converts texts into sequences of numeric indices. Vector arithmetic is used to describe the connections between words in this case as well. The support vector machine is a widely popular Machine Learning Model that is favoured above other

approaches for Medical Diagnosis, which is one of the primary reasons for our pick of the model.

The four phases that were carried out for Training using the SVM model are very similar to the Classification Approach for COVID Segmentation. They are as follows:

1. Determination of a COVID-19 subject's identity.
2. Using the tokenized document and bag of words, convert the words into numeric sequences. A critical phase in the process of cluster segmentation is based on specific symptoms.
3. Create and train an SVM model for grouping the COVID Subjects in order to better understand them.
4. Using the SVM Model that was previously trained, cluster the fresh symptomatic text data.

The group has attempted to categorise the COVID Subjects into six clusters in order to elucidate on their methodology further.

Cluster-I is considered to be the mildest kind of infection, and the symptoms experienced by those who are infected with it include flu-like symptoms without a fever, cold-like symptoms, sore throat, blocked nose, chest discomfort, muscular pain, loss of smell, and headache. Only a small number of people experience upper respiratory tract discomfort as a result of an elevated viral load.

With the presence of a flu-like illness and fever, the Cluster-II becomes a little more difficult to manage. Patients who fell into this group reported experiencing symptoms such as persistent fever, lack of appetite, and hoarseness in the voice, which are often associated with dry cough.

The participants in Cluster-III suffer from a greater number of problems, including gastrointestinal infection. Patients who belonged to this cluster experienced symptoms that interfered with their digestion and gastrointestinal function, as described above. Despite the fact that cough was not a significant symptom in this cluster, infected patients complained of nausea, lack of appetite, vomiting, and diarrhoea, which were all far more prevalent. Headache and chest discomfort are two less frequent symptoms to be aware of.

The people classified under Cluster-IV can be classified as Severe Level-1, with exhaustion being the

most prevalent symptom experienced by them. The symptoms identified in this cluster of infections were connected to energy loss and tiredness, which were brought on by a slowdown in the immune system. Patients in this category experienced symptoms such as weariness, headache, loss of smell and taste, sore throat, fever, and chest discomfort, which were considered to be warning signs for severe COVID-19.

Cluster-V can be classified as Severe level-2 with confusion, in which case the affected persons begin to experience neurological symptoms as well as physical symptoms. The sort of symptoms in this cluster were more severe than those in level 1, and they had an influence on nerve functioning. This was the beginning of the long-term impact that COVID may have on the brain in the long run. Headache, loss of smell, lack of appetite, cough, fever, hoarseness, disorientation, sore throat, chest discomfort, weariness, confusion, and muscular soreness were some of the symptoms that were noted.

When it comes to stomach and respiratory discomfort, Cluster-VI individuals might be classified as Severe Level 3. This is the scariest and severe type of symptoms that patients experience during the first week of their illness. The people who belonged to this cluster were much more likely to be hospitalised than the general population and to require ventilation and oxygen support because they were experiencing symptoms such as confusion, sore throat, chronic fever, loss of appetite, headache, diarrhoea, shortness of breath, muscle and abdominal pain.

3. OBSERVATIONS AND DISCUSSIONS

The quantitative data set contained age, gender, temperature, and oxygen saturation as parameters for the correct classification of COVID individuals. The Classification Learner App in MATLAB® was used to train the COVID individuals, and the Model with the highest accuracy was automatically returned after training. For better comprehension, the specifics of the trained model, as well as the various performance evaluation visualisations, are presented. Table 1 lists the parameters of the Machine Learning model that was used.

Bag, Ada Boost, and RUS Boost were all evaluated independently, and the ensemble technique was shown to be effective. It was tested with a range of learners ranging from 10 to 500, with the Learning Rate varying from 0.001 to 1, with the number of splits ranging from 1 to 560, and with the number of predictors ranging from 1 to 4. The default value for the cost matrix for misclassification was set to zero. The final model that was obtained is returned in the manner described above.

The following is the outcome of our model for the provided data set:

As shown in Figures 3-5, a Machine Learning Model with a confusion matrix, receiver operating characteristic (ROC) plot, and Misclassification Error Plot is used to classify data.

Table 1: Machine Learning Model Parameters

MODEL: OPTIMIZABLE ENSEMBLE	
<i>Learner Type</i>	Decision Tree
<i>Accuracy</i>	81.5%
<i>Total Misclassification Cost</i>	104
<i>Prediction Speed</i>	~790 observations/sec
<i>Training Time</i>	312.59 sec
OPTIMIZED HYPERPARAMETERS	
<i>Ensemble Method</i>	BAG
<i>Max No of Splits</i>	556
<i>No of Learners</i>	58
<i>No of Predictor Samples</i>	4
OPTIMIZER OPTIONS: BAYSEAN OPTIMIZATION	
<i>Iterations</i>	30
<i>Training Time</i>	300 s (True)

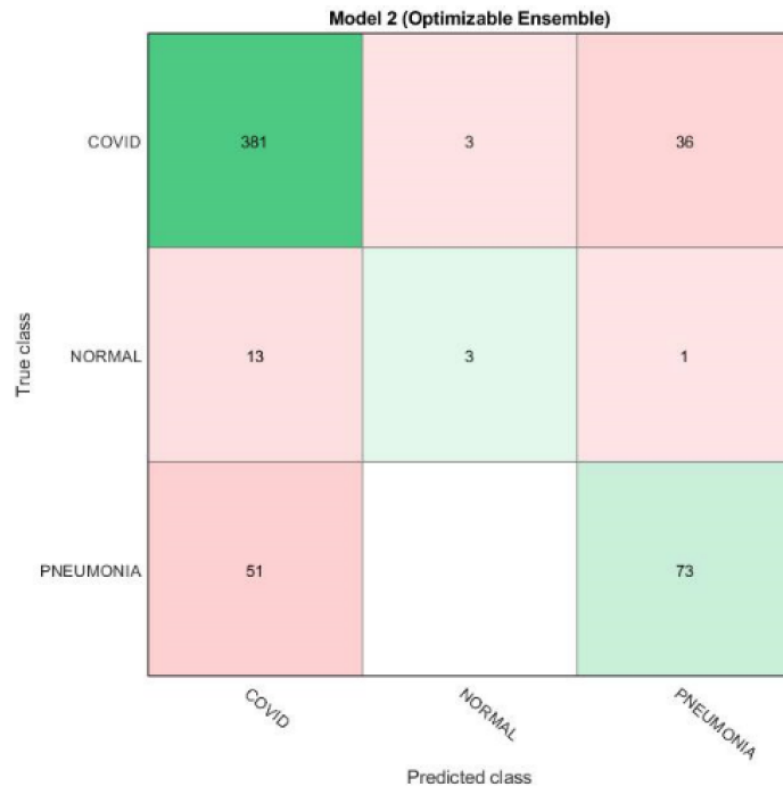


Figure 3: Confusion Matrix of Machine Learning Model.

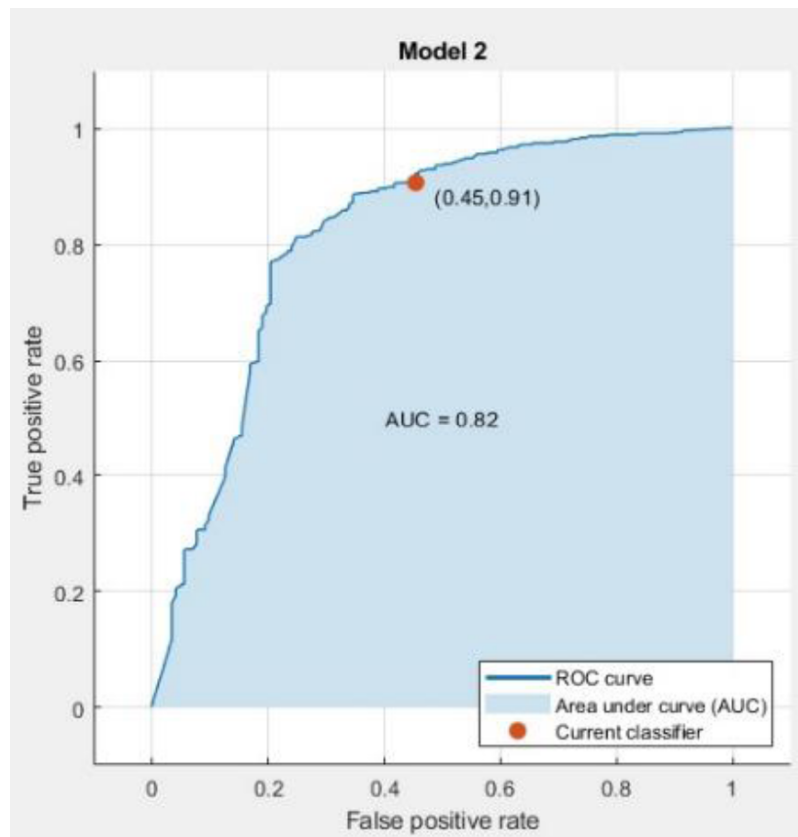


Figure 4: Receiver Operating Characteristics Plot of Machine Learning Model.

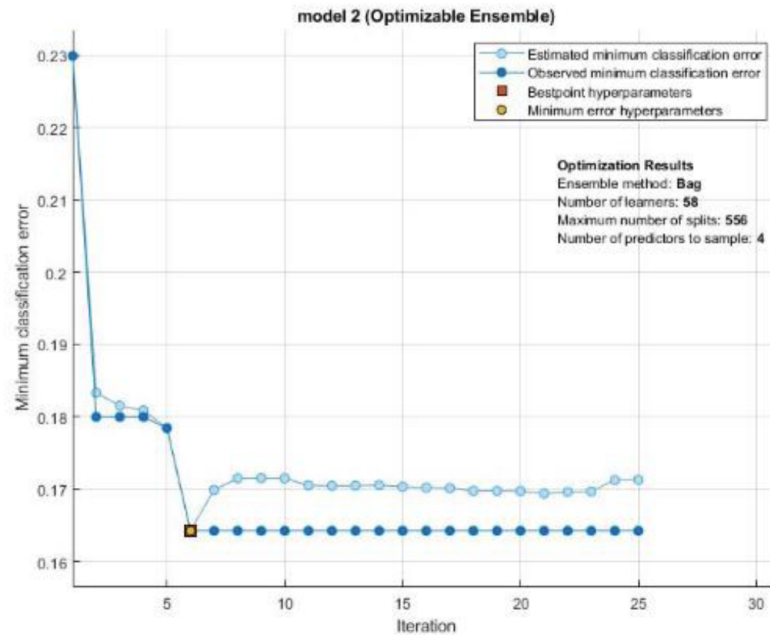


Figure 5: Misclassification Error Plot of Machine Learning Model.

The following are some of the most important mathematical inferences that can be drawn from Figure 4:

- (i) In terms of sensitivity (COVID), it is 90.71 percent
- (ii) Specificity (for healthy patients): 17.65 percent
- (iii) In the case of pneumonia patients, specificity is 58.87 percent.
- (iv) The AUC of the ROC Plot is calculated to be 0.82. The region covered by the model increases in accuracy as the size of the area increases. As a result, it is very suggestive that the model obtained is quite resilient.

It was also tuned to be within the permissible limit for the misclassification (Figure 5) or the Error Value (0.165).

The sensitivity and specificity of a model are the two most important elements to consider when determining which model to use. Putting it simply, and as an explanation for the aforementioned findings (confusion matrix), we discover that the sensitivity is defined by the class of COVID (target class) observed as the output class of COVID. It has been noted that the sensitivity test for 'COVID' has a greater accuracy return than the other tests performed. In addition, the number of false negatives in the case of the applicant who is projected to be a 'Normal Patient' is minimal.

The other false negative that corresponds to the 'Pneumonia Patient' component is of little consequence because it also necessitates the implementation of corrective medical procedures.

Specificity is the other component that is equivalent to sensitivity. We have defined the same thing in two different ways. One who is defined by the term 'Normal' (target class) and is discovered as such (output class). In Nature, we discover that the sensor data does not provide a great deal of detail. The 'Normal' discovered as 'COVID' is significant, but not to a higher level than the 'COVID' since they do not pose a threat to the spread of the disease. The number of false positives may reduce the accuracy of our model, but from the perspective of medical criticality, this is a minor concern compared to the number of false negatives.

The second specialised class is referred to as 'Pneumonia.' When it comes to the accuracy of the same, we find that it is around 59 percent. However, none of the candidates has been incorrectly labelled as 'Normal.' Despite the fact that a significant section of the population has been misclassified as 'COVID,' the individuals might again be subjected to additional confirmatory exams to get a more accurate conclusion. Given that the total accuracy of an AI-based system is reliant on the number of true positives and false negatives detected, the whole model was designed to account for these two factors. The precision metric, which accounts for the false positives among the genuine positives, is an acceptable measure of effectiveness.

Table 2: Comparative Table to Identify Superior Model

Model	Parameters	Scores in %
Support Vector Machine (SVM)	Sensitivity (COVID)	97.39
	Specificity (Normal)	83.33
	Specificity (Pneumonia)	68.85
	Precision	90.56
	Accuracy	9.43
Long Short-Term Memory (LSTM)	Sensitivity (COVID)	98.44
	Specificity (Normal)	0
	Specificity (Pneumonia)	54.10
	Precision	85.91
	Accuracy	86.72
Convolutional Neural Network (CNN)	Sensitivity (COVID)	97.65
	Specificity (Normal)	0
	Specificity (Pneumonia)	53.28
	Precision	85.62
	Accuracy	85.94

The qualitative data was utilised as an input feed to evaluate our Support Vector Machine (SVM) trained model, which was trained using support vector machines. On the following parameters, a comparative table is provided in Table 2 to help you choose which model is the preferable one.

Support vector machine, a machine learning model, surpasses the two deep learning models on the majority of performance assessment parameters, including accuracy. The sensitivity measure was somewhat higher for the two Deep Learning Models, but the support vector machine was determined to be

better than the other two Deep Learning Models in terms of the other assessment metrics, which are accuracy and sensitivity.

Table 3 presents the results of a symptomatic examination of a small number of individuals, as well as their respective cluster segmentation:

The symptoms outlined in the six clusters can offer us suggestive inferences as to how COVID-19 influences different categories of individuals and can serve as a warning system for the kind of symptoms to expect in different groups of people. Among all patients, fever and cough were persistent in all groups

Table 3: The Symptomatic Analysis of a Sample and their Respective Segmentation

SI. No	Symptoms	Actual Labels	Predicted Labels	Cluster
1	"Hypertension, Type 2 diabetes, Coronary Heart Disease, Lung cancer, Low-grade Fever and Fatigue, Travel History"	COVID	COVID	Cluster-III
2	" Hypertension, Type 2 Diabetes, Coronary Heart Disease, Low-grade Fever and Fatigue. Travel History."	COVID	COVID	Cluster-II
3	" Hypertension, Type 2 Diabetes, Coronary Heart Disease, Low-grade Fever and Fatigue. Travel History."	COVID	COVID	Cluster-II
4	" Fever and Non-Productive cough, Travel History, Controlled Hypertension, Fever, No Respiratory Distress"	COVID	COVID	Cluster-III
5	" High-grade Fever, Cough, and Fatigue for 1 week. "	COVID	COVID	Cluster-II
6	"Sore Throat, Dry Cough, Fatigue, and Low-grade Subjective Fever, history of Hypothyroidism, Travel History"	COVID	COVID	Cluster-III

for three to four days before subsiding. The sense of smell was something that was stimulated in the patients only after the 4th day after the infection had occurred. In accordance with the findings of the different analyses, the difference in the severity of illness could only be noticed after 4-5 days of infection. As a result, the segmentation is susceptible to change and should not be regarded until after the next 4-5 days.

4. CONCLUSION

For starters, the authors of this research have presented a solution pipeline that is more composite and resilient than previous work since it involves a two-step verification method. We want to convert our algorithm into software as part of the next phase of our development, and it will serve as a useful tool for health practitioners and other medical staff members, as well as providing a vital second opinion during the early diagnosis process. Our suggested technique may aid in the acceleration of the screening process and, as a result, the rise in the number of isolated cases each day, thanks to the efficiency of the identification and clustering processes that we offer. Our training data and validation data are consistent with Asian origin. The results/observations are in accordance with an upgraded mathematical model that has more accuracy than the previous model. An expanded number of clinical validations is being sought in order to promote acceptance and adaptability while also seeking to improve the overall quality of the product. Patients in the severe or high-risk categories are more likely than those in the mild or moderate cluster to experience a symptom such as weariness during the first week of their illness. Some of the crucial signs were seen as early as the first day of the experiment. Breathlessness, weariness, and stomach discomfort were among the symptoms experienced by patients. As a result of their research, they discovered that those who were assigned to clusters 4, 5 or 6 were often older and frailer and that they were also more likely to be overweight and have more severe pre-existing medical disorders than those who were assigned to clusters 1, 2 or 3. Researchers observed that just 1.5 percent of persons in Cluster-I, 4.4 percent in Cluster II, and 3.3 percent in Cluster III needed oxygen, which was thought to be a mask for the disease's progression. The identification of clusters can help people recognise how important it is to monitor symptoms and to offer priority care to individuals who may require it more than others. It can also help people understand how to provide the appropriate tools to

prevent a second wave in certain people. The characteristics and particular symptoms might also prove to be a breakthrough, allowing doctors to identify the people who are most at risk for developing the disease. This would aid in the making of quick choices and the saving of lives. To summarise, this article demonstrates the need of tracking symptoms over time in order to generate accurate forecasts about individual risk and to anticipate more nuanced and accurate outcomes in future research. The strategy we took would eventually aid in understanding the developing tale of this disease in each individual patient over time, allowing them to get the best possible therapy.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the Department of Radiology and Imaging at the North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences in Shillong, Meghalaya, India, for providing data for large-scale clinical validation. Additionally, the authors thank The Neotia University and Adamas University for providing critical research infrastructure.

DECLARATIONS

Funding

Provided by The Neotia University and Adamas University, Kolkata, India.

Conflicts of Interest/Competing Interests

Not applicable.

Availability of Data and Material

Not applicable.

Code Availability

Not applicable.

Authors' Contributions

Mr. Swarnava Biswas and Mr. Debajit Sen contributed equally for this research paper.

Ethics Approval

Not applicable.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

REFERENCES

- [1] Organization WH. Coronavirus disease 2019 (COVID-19): situation report 2020; 82.
- [2] Hopman, Allegranzi B, Mehtar S. Managing COVID-19 in low-and middle-income countries. *JAMA* 2020; 323(16): 1549-1550. <https://doi.org/10.1001/jama.2020.4169>
- [3] Güner HR, Hasanoğlu I, Aktaş F. COVID-19: Prevention and control measures in community. *Turkish Journal of Medical Sciences* 2020; 50(SI-1): 571-577. <https://doi.org/10.3906/sag-2004-146>
- [4] Tian H, *et al.* An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 2020; 368(6491): 638-642. <https://doi.org/10.1126/science.abb6105>
- [5] Hanson DF. Fever, temperature, and the immune response. *Annals of the New York Academy of Sciences* 1997; 813: 453-464. <https://doi.org/10.1111/j.1749-6632.1997.tb51733.x>
- [6] Davis DP, Hwang JQ, Dunford JV. Rate of decline in oxygen saturation at various pulse oximetry values with prehospital rapid sequence intubation. *Prehospital Emergency Care* 2008; 12(1): 46-51. <https://doi.org/10.1080/10903120701710470>
- [7] Cambria E, White B. Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine* 2014; 9(2): 48-57. <https://doi.org/10.1109/MCI.2014.2307227>
- [8] Selvapandian D, Mary RUN, Karthikeyan C. Artificial intelligence in online shopping using natural language processing (NLP). *Journal of Critical Reviews* 2019; 7(4): 2020. <https://doi.org/10.31838/jcr.07.04.165>
- [9] Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing *IEEE Computational Intelligence Magazine* 2018; 13(3): 55-75. <https://doi.org/10.1109/MCI.2018.2840738>
- [10] Wang Z, Wu Q. An integrated deep generative model for text classification and generation. *Mathematical Problems in Engineering* 2018; 2018. <https://doi.org/10.1155/2018/7529286>
- [11] Abdel-Hamid O, Mohamed A-R, Jiang H, Deng L, Penn G, Yu D. Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 2014; 22(10): 1533-1545. <https://doi.org/10.1109/TASLP.2014.2339736>
- [12] Sang K, Sharma A. Natural Language Processing: An Approach to Aid Emergency Services in COVID-19 Pandemic. *International Journal of Innovative Research in Computer Science & Technology (IJRCST)* ISSN, 2020; pp. 2347-5552.
- [13] Gostic K, Gomez AC, Mummah RO, Kucharski AJ, Lloyd-Smith JO. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *Elife* 2020; 9: e55570. <https://doi.org/10.7554/eLife.55570>
- [14] Yan CH, Faraji F, Prajapati DP, Boone CE, DeConde AS. Association of chemosensory dysfunction and Covid-19 in patients presenting with influenza-like symptoms. in *International forum of allergy & rhinology: Wiley Online Library* 2020. <https://doi.org/10.1002/alr.22579>
- [15] Pan L, *et al.* Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China: a descriptive, cross-sectional, multicenter study. *The American Journal of Gastroenterology* 2020; 115. <https://doi.org/10.14309/ajg.0000000000000620>
- [16] Rafiq D, Suhail SA, Bazaz MA. Evaluation and prediction of COVID-19 in India: A case study of worst hit states. *Chaos, Solitons & Fractals* 2020; 139: 110014. <https://doi.org/10.1016/j.chaos.2020.110014>
- [17] Cauwenberghs G, Poggio T. Incremental and decremental support vector machine learning. in *Advances in Neural Information Processing Systems* 2001; pp. 409-415.
- [18] Graves A. Long short-term memory. in *Supervised sequence labelling with recurrent neural networks: Springer*, 2012; pp. 37-45. https://doi.org/10.1007/978-3-642-24797-2_4
- [19] Sak H, Senior AW, Beaufays F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling 2014. <https://doi.org/10.21437/Interspeech.2014-80>
- [20] Hu B, Lu Z, Li H, Chen Q. Convolutional neural network architectures for matching natural language sentences, in *Advances in neural information processing systems* 2014; pp. 2042-2050.
- [21] Ertunc H, Loparo K. A decision fusion algorithm for tool wear condition monitoring in drilling. *International Journal of Machine Tools and Manufacture* 2001; 41(9): 1347-1362. [https://doi.org/10.1016/S0890-6955\(00\)00111-5](https://doi.org/10.1016/S0890-6955(00)00111-5)

Received on 24-10-2021

Accepted on 30-11-2021

Published on 01-12-2021

<https://doi.org/10.6000/1929-6029.2021.10.17>

© 2021 Biswas *et al.*; Licensee Lifescience Global.

This is an open access article licensed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution and reproduction in any medium, provided the work is properly cited.