Adaptive Neuro Fuzzy Inference System for Diagnosing Coronavirus Disease

2019 (COVID-19)

Kingsley C. Ukaoha*
Oluwadamilola Ademiluyi
Juliana Ndunagu
Stephen S. Daodu
Frank Osang

Department of Computer Science, University of Benin, Benin City, Nigeria
Department of Computer Science, University of Benin, Benin City, Nigeria
Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria
Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria

Received 2020-10-30; Revised 2020-11-6; Accepted 2020-11-9
Available online 2020-12-26

Abstract

Coronaviruses which are positively sensed single-stranded Ribonucleic Acid (RNA) viruses are causing serious threat to global public health due to the widespread of the virus and no one having immunity to the virus. Timely diagnosis of the disease has become a major challenge due to the limitation associated with the present methods used in diagnosing of COVID-19 and a limited number of COVID-19 test kits available in hospitals due to the increasing number of cases daily. There is a need to propose a model that can provide timely, differential and alternative diagnosis option to prevent COVID-19 spreading among people. In this study an ANFIS based model was proposed for diagnosing COVID-19, the model was trained and tested
using 600 COVID-19 dataset. The ANFIS model had accuracy of 96.6% for predicting and diagnosing COVID-19.

**Keywords:** Coronavirus; COVID-19; Diagnose; ANFIS

1.0 Introduction

The world has experienced several wide spread diseases causing serious threat to global public health affecting a large proportion of the population, including the 2002 Severe Acute Respiratory Syndrome (SARS) epidemic that caused 774 deaths out of about 8000 cases, the 2009 H1N1 influenza virus pandemic with 18500 deaths, the 2012 Middle East Respiratory Syndrome (MERS) epidemic that caused 800 deaths out of 2500 cases, the 2014 Ebola virus outbreak with 28616 cases and 11310 deaths, and the recent Coronavirus Disease 2019 (COVID-19) pandemic [1].

Coronaviruses (CoVs) are positively sensed single-stranded Ribonucleic Acid (RNA) viruses that belong to the order Nidovirales, family Coronaviridae, and subfamily Orthocoronavirinae with four (4) species; alpha, beta, delta and gamma coronaviruses [2]. Alpha CoVs and beta CoVs originated from bats and rodents while delta CoVs and gamma CoVs have their origins from avian species [3]. COVID-19 can be transmitted from human-to-human by respiratory droplets from sneezing, coughing and aerosols, with symptomatic people being the major source of transmission. It has a dynamic incubation period of about 5 to 14 days [4].
CoVs have been reported to be the major causes of about 5% to 10% acute respiratory infections [5]. The first outbreak of COVID-19 was reported in Wuhan, China in December, 2019 with the reported confirmed cases standing at over 29.9 million and recorded deaths 942,735 worldwide which is still counting day by day. COVID-19 diseases continuing to infect and reduce human populations which is due to the fact that no one has immunity to COVID-19, which means thousands to millions of people are likely to be more vulnerable to this viral infection and severe disease. COVID-19 affects different people in different ways, the World Health Organization (WHO) listed the symptoms of COVID-19 from the less common symptoms to serious symptoms [6]. The listed symptoms of COVID-19 is tabulated in Table 1:
### Table 1: COVID-19 Symptoms

<table>
<thead>
<tr>
<th>S/N</th>
<th>Less common</th>
<th>Most common</th>
<th>Serious symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aches and pains</td>
<td>Fever</td>
<td>Difficulty breathing or shortness of breath</td>
</tr>
<tr>
<td>2</td>
<td>Sore throat</td>
<td>Dry cough</td>
<td>Chest pain or pressure</td>
</tr>
<tr>
<td>3</td>
<td>Diarrhoea</td>
<td>Tiredness</td>
<td>Loss of speech or movement</td>
</tr>
<tr>
<td>4</td>
<td>Conjunctivitis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Headache</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Loss of taste or smell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Rash on skin, or discoloration of fingers or toes.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source:** WHO, 2020

For the diagnosis of COVID-19, detection of viral RNA in the secretions from the respiratory tract of infected patients by Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test is currently the standard method for diagnosis of COVID-19. This method have some weaknesses ranging from long turnaround time usually around 2–4 hour to the requirement of specialized facilities to carry out the test. Scientists around the world have been devoting effort to developing improved nucleic acid-based, simpler and faster methods. The United States (US) FDA issued an emergency-use authorization to Cepheid’s Xpert Xpress SARS-CoV-2 test, which became the first point-of-care COVID-19 diagnostic test to receive this designation in the US. The test is designed to use the company’s automated GeneXpert Systems and has a turnaround time of approximately 45 min [7]. Timely and differential diagnosis of disease are today becoming a necessity in medicine and health systems [8], [9]. In this regard, proposing and developing expert systems based on machine learning method for accurate prediction, efficient diagnosis and effective management of diseases has drawn significant research attention among
Adaptive Neuro Fuzzy Inference System for Diagnosing Coronavirus Disease 2019 (COVID-19) researchers and physicians [8-14]. However, this study is to propose a model based on Adaptive Neuro-Fuzzy Inference System (ANFIS) for diagnosing COVID-19.

ANFIS is a branch of Artificial Intelligence (AI) which plays a major role in prediction, modeling and inference. ANFIS is a Fuzzy Inference System (FIS) implemented in the framework of adaptive networks. It integrates both Neural Networks (NN) And Fuzzy Logic (FL) principles into a single framework with learning capability to approximate non-linear functions and works as a universal estimator [15]. It is based on the Takagi–Sugeno FIS, developed in the early 1990s. The learning networks in this model are based on mathematical computations capable of solving complex problems. ANFIS based predicting models which contains the knowledge and experience of an expert can accurately predict diseases [11], [16].

2.0 Related Works

There are number of numerous studies which have been carry out on detecting, predicting, classifying and diagnosing of COVID-19 using Artificial Intelligence (AI).

Alile and Otokiti, proposed and simulated a Bayesian Belief Network (BBN) model to predict COVID-19 using 47 disease ailments dataset and each ailment has a value which represents the probability of such disease ailment causing COVID-19. The proposed model was trained and tested and it was reported to have an accuracy of 99% in predicting COVID-19 with its symptoms [17].

Narin et al. proposed a three different Convolutional Neural Network (CNN) based models (ResNet50, InceptionV3 and Inception-ResNetV2) for the detection of coronavirus pneumonia infected patient using 100 chest X-ray radiographs. The chest X-ray radiographs consists of 50 normal healthy people and 50 COVID-19 patients. The three CNN models; ResNet50,
InceptionV3 and Inception-ResNetV2 were evaluated using five-fold cross validation and it was reported that ResNet50 model had the best detection accuracy of 98% in detection of COVID-19. While the other two models reported 97% accuracy for InceptionV3 and 87% accuracy for Inception-ResNetV2 [18].

Sethy and Behera proposed a model for detecting COVID-19 by extracting features from chest X-ray images using a deep-learning algorithm and classified the images as either infected or non-infected using a Support Vector Machine (SVM). 11 deep-learning models was employed; AlexNet, VGG16, VGG19, GoogLeNet, ResNet18, ResNet50, ResNet101, InceptionV3, Inception-ResNetV2, DenseNet201, and XceptionNet. Two categories of dataset was collected; the first containing chest X-ray images of 25 infected patients and 25 non-infected patients and the other containing chest X-ray images of 133 infected patients (e.g. MERS, SARS and Acute Respiratory Distress Syndrome (ARDS) patients) and 133 non-infected patients. They performed separate feature extractions on each dataset using various models and achieved a 95.38% accuracy with ResNet50 and SVM [19].

Gozes et al. proposed a system to automatically identify COVID-19 patients and examine the disease burden quantification by employing a deep-learning approach on Computed Tomography (CT) scans using a dataset of CT scans from 157 foreign patients from China and the United States of America (USA). The proposed system analyses the CT scan at two distinct levels; subsystems A and B. Subsystem A performs a 3D analysis and subsystem B performs a 2D analysis of each segment of the scan to identify and locate broader diffuse opacities including ground-glass infiltrates which have been clinically identified as representative of COVID-19. For the system evaluation ResNet50 was applied to subsystem B and reported an accuracy of 98.2% and specificity of 92.2% respectively [20].
Fatima and Samy designed and implemented an expert system to detect and diagnose symptoms of COVID-19 using Clips and Delphi expert system languages. The main sources of the knowledge for this expert system are medical and specializes websites for COVID-19. The knowledge captured are converted into the Knowledge Base (KB) using CLIPS shell. The diagnosis of COVID-19 by the expert system was accomplished by displaying all symptoms of COVID-19 and selecting from the symptoms in the list and then analyze to diagnosis the day of recognizing symptoms, survival and spread, favorable conditions and snapshot of the status. The expert system was assessed by doctors and they were satisfied and accepted with its quality of performance [21].

Fu et al. proposed a classification system based on ResNet50 to detect COVID-19 and some other infectious lung diseases (pulmonary tuberculosis and bacterial pneumonia). 60,427 CT scans from 918 patients were dataset collected for the proposed system. 14,944 of these CT scans were from 150 COVID-19 patients and 15,133 from 154 non-COVID-19 viral pneumonia patients. Several tests was carried out for numerous number of lung diseases. The achieved accuracy, sensitivity and specificity of the system were 98.8%, 98.2% and 98.9% respectively [22].

3.0 Research Methodology

3.1 Dataset

In this study, the dataset used in training, testing and diagnosing COVID-19 are 600 data from patients who visited Hospital Israelita Albert Einstein, Sao Paulo, Brazil who samples were collected to perform the COVID-19 test. The data was retrieved from Kaggle dataset open source repository. The chest dataset consists of 300 negative cases and 300 positive cases of COVID-19 patients. In Figure 2 and Figure 3 the sample dataset are shown:
<table>
<thead>
<tr>
<th>S/N</th>
<th>Fever</th>
<th>Dry Cough</th>
<th>Headache</th>
<th>Pharyngitis</th>
<th>Sore Throat</th>
<th>Rash or Skin</th>
<th>Difficulty in Breathing</th>
<th>Pneumonia</th>
<th>Rash on Skin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.1618</td>
<td>2.1434</td>
<td>3.0737</td>
<td>2.5456</td>
<td>2.8221</td>
<td>3.8667</td>
<td>2.1503</td>
<td>2.1434</td>
<td>2.1434</td>
</tr>
<tr>
<td>2</td>
<td>3.5005</td>
<td>2.8881</td>
<td>3.0777</td>
<td>2.8881</td>
<td>2.8881</td>
<td>3.8667</td>
<td>2.1503</td>
<td>2.1434</td>
<td>2.1434</td>
</tr>
<tr>
<td>3</td>
<td>2.8881</td>
<td>2.1434</td>
<td>3.0777</td>
<td>2.8881</td>
<td>2.8881</td>
<td>3.8667</td>
<td>2.1503</td>
<td>2.1434</td>
<td>2.1434</td>
</tr>
</tbody>
</table>

COVID-19 Dataset (Negara Cases)
Adaptive Neuro Fuzzy Inference System for Diagnosing Coronavirus Disease 2019 (COVID-19)
Adaptive Neuro Fuzzy Inference System for Diagnosing Coronavirus Disease 2019 (COVID-19)

<table>
<thead>
<tr>
<th>Figure 2: COVID-19 Dataset (Negative Cases)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S/N</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
</tbody>
</table>
Adaptive Neuro Fuzzy Inference System for Diagnosing Coronavirus Disease 2019 (COVID-19)
Adaptive Neuro Fuzzy Inference System for Diagnosing Coronavirus Disease 2019 (COVID-19)
Figure 3: COVID-19 Dataset (Positive Cases)
3.2 **Adaptive Neuro Fuzzy Inference System (ANFIS)**

ANFIS is a Fuzzy Inference System (FIS) implemented in the framework of adaptive networks. It integrates both Neural Networks (NN) and Fuzzy Logic (FL) principles into a single framework with learning capability to approximate non-linear functions and works as a universal estimator. It is based on the Takagi–Sugeno FIS [15]. The ANFIS architecture is made up of six layers, the layers are described below:

**Input Layer:** The symptoms for diagnosing COVID-19 are the inputs used in this layer. These symptoms include: fever, dry cough, tiredness, aches and pains, sore throat, diarrhoea, conjunctivitis, headache, loss of taste or smell, a rash on skin or discolouration of fingers or toes, difficulty breathing or shortness of breath and chest pain or pressure. The input layer is represented mathematically as shown in equation 1:

\[
0_i^1 = x_i \quad (1)
\]

Where:
- \(0_i^1\) is the \(i^{th}\) neuron output from the input layer
- \(x\) is value for each parameters

**Membership function Layer:** It connects the symptoms from the input layer to a fuzzy set. In this layer each symptom is mapped using the Gaussian membership function to a membership set. The Gaussian membership function is shown in equation 2:
\[ \mu(x) = \exp\left(-\frac{(c_i - x)^2}{2\alpha_i^2}\right) \]  \hspace{1cm} (2)

Where;

- \(c_i\) is the centre of the \(i^{th}\) fuzzy set
- \(a_i\) is the width of the \(i^{th}\) fuzzy set
- \(x\) is the value for each node input
- \(\mu(x)\) is the membership function of \(x\).

Rule Layer: in this layer the fuzzified symptom values are combined using Takagi-Sugeno inference rule to generate an outcome for each case. Takagi-Sugeno is a fuzzy inference technique that develops a systematic approach to generate fuzzy rules from a given input – output data set. Takagi-Sugeno has fuzzy inputs and a crisp output. Takagi-Sugeno uses weighted average to compute the crisp output. It is computationally efficient and suitable to work with optimization and adaptive techniques, so it is very adequate for control problems, mainly for dynamic nonlinear systems [17]. It can be represented mathematically as shown in equation 3:

\[ O_i^3 = \mu(x) \ast \mu(y) \]  \hspace{1cm} (3)

Where;

- \(O_i^3\) is the \(i^{th}\) neuron output from the rule layer
- \(x\) and \(y\) are the inputs to node \(i\)
- \(\mu(x)\) and \(\mu(y)\) is the membership function of \(x\) and \(y\) respectively.

Normalization Layer: each neuron in this layer links to exactly one neuron in the rule layer and it computes the firing strength of each rule. It can be denoted mathematically as shown in equation 4:

\[ O_i^4 = \frac{O_i^3}{O_1^3 + O_2^3 + \ldots + O_n^3} \]  \hspace{1cm} (4)
Where;

\( O_i^4 \) is the \( i^{th} \) neuron output for normalization layer

\( O_i^3 \) is the \( i^{th} \) neuron output from the rule layer

\( n \) is the total number of neurons in normalization layer.

Defuzzification Layer: consists of a single neuron to which all the neurons from the normalization layer are linked. The defuzzification layer output is determined by multiplying the firing strength of a rule by its subsequent parameters. The defuzzification method used in this layer is wtaver. It can be represented mathematically as shown in equation 5:

\[
O_i^5 = O_i^4 (p_ix + q_iy + r) \quad (5)
\]

Where;

\( O_i^5 \) is the \( i^{th} \) neuron output for defuzzification layer

\( O_i^4 \) is the \( i^{th} \) neuron output for normalization layer

\( x \) and \( y \) are the inputs to node \( i \)

\( p_i \) and \( q_i \) are the consequent parameters

\( r \) is the bias

Output Layer: the neurons in this layer determined the total output of the ANFIS. The input into this layer is received from the defuzzification layer and it creates its output by adding the inputs from the defuzzification layer. It can be represented mathematically as shown in equation 6:

\[
O_i^6 = \sum_{i}^{n} O_i^5 \quad (6)
\]

Where;

\( O_i^6 \) is the total output of the ANFIS

\( O_i^5 \) is the \( i^{th} \) neuron output from defuzzification layer.
Matrix Laboratory (MATLAB) version 7.5.0.342 (R2007b) was used to implement the ANFIS model on Windows 10 Operating System (OS), running on Intel Celeron.

4.0 Experiment, Result and Discussion

The dataset used for the experiment was divided into two independent datasets with 67% (400) of the dataset was used for training of the ANFIS model and 33% (200) of dataset for testing of the model. The 400 dataset used for training of the ANFIS model consists of 200 dataset from the negative cases and another 200 dataset from the positive cases, the 200 dataset used for testing of the ANFIS model consists of 100 dataset from the negative cases and another 100 dataset from the positive cases. The structure of the loaded dataset for training and testing is shown in Figure 4 and Figure 5 respectively:

![Figure 4: Training Data Structure](image-url)
The ANFIS model was trained for 20 epochs utilizing a hybrid optimization method with an error tolerance of 0.5. Subtractive clustering (Sub-clustering) was used to produce the FIS of the ANFIS model. The idea behind the Sub-clustering method is to divide the data space into fuzzy clusters, each representing a particular part of the system behaviour. Subtractive clustering is one-pass algorithm for estimating the number of clusters. The model parameters are updated in the training process utilizing hybrid optimization learning algorithm, the hybrid optimization learning algorithm is a combination of two optimization methods which are; gradient descent (backward pass) and least squares methods (forward pass). The least squares method (forward
pass) is used to optimize the consequent parameters and the gradient descent method (backward pass) is used to optimize the premise parameters.

The structure of the ANFIS model and the FIS is shown in Figure 6 and Figure 7 respectively:

Figure 6: ANFIS Model Structure
The ANFIS model was designed using Gaussian membership function. The membership function for each symptoms has three (3) linguistic labels which are; mild, moderate and severe. The membership function for one of the symptom “fever” is shown in Figure 8:
The ANFIS model was trained using the training dataset, the result achieved from training of the model showed that the system had a training error of 3.2876e-006 at epoch 1. The structure of the training model is shown in Figure 9:
The system had 400 fuzzy rules in the rule layer, the structure of the rule is shown in Figure 10:
Rule Viewer: Diagnosis

over = 4.0
in2 = 5
in3 = 5
in4 = 4.0
in5 = 5.0
in6 = 5.0
in7 = 5.0
in8 = 5.0
in9 = 5.0
in10 = 5.0
in11 = 5.0
in12 = 5.0
in13 = 5.0

K.C. Ukaoha et al.
After training of the ANFIS model, the model was tested using the test dataset. The result achieved from testing of the model showed that the system had an average testing error of 3.4254 on the test dataset. The result generated from the ANFIS model indicated that the model was able to accurately classify approximately 96.6% of the test dataset. The structure of the testing model is shown in Figure 11:
In this study, symptoms were used for diagnosing COVID-19 patients. The ANFIS model was trained and tested on the symptoms dataset collected. The ANFIS model had a prediction accuracy of 96.6%. The higher prediction accuracy achieved by the model can be traced to the learning capability of ANFIS to easily approximate non-linear functions and works as a universal estimator, the Takagi–Sugeno FIS of ANFIS and also the range of dataset used for the training and testing of the model [15]. The ANFIS model does not have any feature extraction or selection.
5.0 Conclusion

Detecting people infected with COVID-19 early and timely from the large number of population is imperative to prevent the spread of the disease to the masses. ANFIS model was proposed in this study for diagnosing COVID-19, the model reported high accuracy of 96.6%. This ANFIS model is cost effective, reduces uncertainties, processing time and will be beneficial in improving healthcare systems. With this model, timely and differential diagnosis of COVID-19 can be achieved and which in return helps doctors make clinical decisions easily and also reduce diagnostic errors. In subsequent studies, the ANFIS model can be implemented and also the model can be trained and tested by increasing the number of data instances of the dataset.

References


