IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

DEVELOPMENT OF SOFT COMPUTING MODEL FOR EARLY DETECTION OF ACUTE RESPIRATORY DISEASES

E. G. Chukwu[1], A. G. Abasiama[2], I. J. Umoren[3] & C. E. Iniubong[4]

Department of Software Engineering, Federal University of Technology Ikot-Abasi[1]

Department of Computer Science, Federal University of Technology Ikot-Abasi[2]

Department of Cyber Security, Federal University of Technology Ikot-Abasi[3]

Department of Statistics, Federal University of Technology Ikot-Abasi[4]

Abstract: Acute respiratory Disease (ARD) is a viral respiratory disease caused by Severe acute respiratory syndrome (SARS), associated coronavirus. It was first identified at the end of February 2003 during an outbreak that emerged in China and spread to 4 other countries. WHO coordinated the international investigation with the assistance of the Global Outbreak Alert and Response Network (GOARN) and worked closely with health authorities in affected countries to provide epidemiological, clinical and logistical support and to bring the outbreak under control. ARD is an airborne virus and can spread through small droplets of saliva in a similar way to the cold and influenza. It was the first severe and readily transmissible new disease to emerge in the 21st century and showed a clear capacity to spread along the routes of international air travel. SARS can also be spread indirectly via surfaces that have been touched by someone who is infected with the virus. Most patients identified with SARS were previously healthy adults aged 25–70 years. A few suspected cases of ARD have been reported among children under 15 years. The case fatality among persons with illness meeting the current WHO case definition for probable and suspected cases of ARD is around 3%. This ARD is of different type, to determine the exert type one is suffering from amongst the various diseases is what we intend to achieve in this research. We shall search out the various types of Acute Respiratory Deseases, search out their various types of symptoms/Signs, determine their Distinctions, know their % relevance, and compare the relevant ratios, using Fuzzy Cluster Means (FCM) Algorithm for clustering of the signs, SQL which we employed for the database query and MysQl Server (PHPmyAdmin) which we used as the data backend (database).

Keywords: Fuzzy cluster means (FCM); Algorithm; Model, Computing, Architecture, SARS-COV-2; SARS-COV-2 (COVID-19); SARS-COV; MERS-COV; HCOV-NL63, Influenza

I. INTRODUCTION

Acute Respiratory Disease (ARD) is a type of respiratory (lung) failure resulting from many different disorders that cause fluid to accumulate in the lungs and oxygen levels in the blood to be too low. The person has shortness of breath, usually with rapid, shallow breathing, the skin may become mottled or blue (cyanosis), and other organs such as the heart and brain may malfunction. It is a medical emergency that may occur in people who already have lung disease or in those with previously normal lungs.

ARDS usually develops within 24 to 48 hours of the original injury or disease but may take as long as 4 or 5 days to become obvious. The person first has shortness of breath, usually with rapid, shallow breathing. Using a stethoscope, a doctor may hear crackling or wheezing sounds in the lungs. The skin may become mottled or blue (cyanosis) in light-skinned people and gray or whitish coloration in the mouth, around the eyes, and under the nails in dark-skinned people because of low oxygen levels in the blood. Other organs such as the heart and brain may malfunction, resulting in a rapid heart rate, abnormal heart rhythms (arrhythmias), confusion, and sleepiness.

To diagnose this disease, the level of oxygen in the blood is measured without taking a blood sample by using a sensor placed on a finger or an earlobe a procedure called pulse oximetry. The level of oxygen (along with carbon dioxide) in the blood can also be measured by analyzing a blood sample taken from an artery. Without prompt treatment, many people who have ARDS do not survive. However, depending on the underlying disorder, with appropriate treatment, about 60 to 75% of people with ARDS survive.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

However, the possibility of a patient surviving this ailment depends on the accuracy of the diagnosis. This is due to the reason that most patients suffering from one type of ARD are wrongly diagnosed with another type due to the similarity in the symptoms. During the COVID'19 era, most patients who suffer from other ARDS like SARS were wrongly diagnosed with COVID'19 and subsequently administered the covid'19 drugs and as a result of the wrong medications given to them, they ended up on untimely dead.

1.1 Aims and Objectives of the Study

The aim of this research is to develop an intelligent model driven by the Fuzzy Cluster Means Algorithm that will detect a specific type of Acute Respiratory disease from the Set of Diseases in the disease class.

The major objectives of this research work which serviced the above Aim are stated below:

1. Analyzing the classes of the diseases in Acute respiratory Disease (ARD) as a confusable disease syndrome
2. Reviewing the existing Methods that are used in diagnosing the Acute respiratory Disease (ARD) class
3. Design a Fuzzy Cluster Means Algorithm that will enhance the proper classification and clustering of the related symptoms in a disease class
4. Develop an early Acute Respiratory Disease (ARD) detection model
5. Simulate the algorithm developed in 3 above using the design approach in Matlab 2019 (MR 2019)
6. Implementing the Fuzzy Cluster Means Algorithm using Python Programming Language as front end programming tool, and MySQL as the Backend Tool.
7. Evaluate the Performance of the Software developed in 5 above.

II. LITERATURE REVIEW

Many theories have been proposed as it regards acute respiratory diseases and what motivates human behavior. Although the literature covers a wide variety of such theories, this review will focus on five major themes which emerge repeatedly throughout the literature reviewed. These themes are incorporation of the theoretical backgrounds, concepts, technologies, review of literatures, and research/literature gapes. Although the literature presents these themes in a variety of contexts, this paper will primarily focus on their application to self-motivation.

2.1 Theoretical Backgrounds

Acute respiratory disease is an infection that may interfere with normal breathing. It can affect just your upper respiratory system, which starts at your sinuses and ends at your vocal chords, or just your lower respiratory system, which starts at your vocal chords and ends at your lungs. This infection is particularly dangerous for children, older adults, and people with immune system disorders.

Diseases of the respiratory system can be grouped into different categories. Example categories include obstructive versus restrictive or acute versus chronic. These diseases have similar causes, symptoms, and effects. To determine signs exclusive for a particular desease necessitated this research.

Some of these deseases are itemized below:

- Respiratory Depression
- Respiratory Syncytial Virus (RSV)
- Acute Respiratory Distress Syndrome (ARDS)
- Mucociliary Clearance and Dysfunction
- Pneumonia
- Cough
- Tuberculosis (TB)
- Bronchiolitis
- Bronchitis

2.2 Literature Review Concept

This literature review outline will be based on themes, chronological order, or methodological approaches. We will write a clear and coherent narrative that synthesizes the information gathered. We will use proper citations for each source and ensure consistency in your citation style.

2.3 Literature Review Technlogy

This review will involve researching, reading, analyzing, evaluating, and summarizing scholarly literature (typically journals and articles) about a specific topic which relates our research aim.

DOI: 10.17148/IJARCCE.2024.131240

2.4 Review of Literatures

Some empirical works have been conducted or carried out by different scholars on automated methods of detecting acute respiratory diseases and their research discussed some of the diseases listed above. Some of these works shall be reviewed beneath:

James, et al (2016) proposed a fuzzy cluster means algorithm for the diagnosis of confusable disease. In their work, they argued that medical science have been overwhelmed in recent times by uncertainty of one form or the other which have greatly affected the decision making process and as such led to cases of misdiagnosis and in worst cases death. One of such forms of uncertainty is the confusability of symptomatic presentations of diseases due to the fact that they share common symptoms and as such becomes difficult for physician's to correctly diagnose them. This difficulty in diagnosis stems from the inability of physicians to quantify the amount of each disease in the confusable disease set depicted by the symptoms. The ultimate goal of medical science is good diagnosis and prevention of diseases and such it is imperative to implement a system to reduce such cases of misdiagnosis which could arise from confusability of disease symptomatic presentations. In this work an expert system driven by the fuzzy cluster means (FCM) algorithms is proposed. The system accepts symptoms as input and provides the degree of membership of each disease in any confusable disease set. Data on alcoholic liver disease were collected and used in the development of the knowledge base. Fuzzy logic and FCM algorithm propelled the inference engine. The system was implemented with CLIPS expert system shell and Java as the front end platform while Microsoft Access was used as the database application. The system gives a measure of each disease within a set of confusable disease. The proposed system had a classification accuracy of 60%.

Faith-Michael et al. (2017) proposed a framework for early differential diagnosis of tropical confusable diseases using the fuzzy cognitive map engine. Access to medical facilities is a huge source of challenge in developing countries, especially of Africa, Asia and the Middle East. Technology has been employed in various spheres of human endeavor, including health care; however, little has been done to harness the power of technology in the diagnosis of tropical (often confusable) diseases. In their study, we mined the experiential knowledge of medical experts to develop an analytic hierarchy process (AHP) model for the diagnosis of seven diseases that are prevalent in tropical (mostly developing) countries. Their system also takes cognizance of some risk factors that could pre-dispose an individual to infection. In addition, it recognizes the sematic causative relationships among symptoms, and can account for comorbidity in the seven diseases under consideration. The system is implemented in an Android environment; it recognizes the need for a friendly and simple user interface for the medical practitioner.

Fatima and Pasha (2017) conducted a survey of Machine Learning Algorithms for Disease Diagnostic. Their work provides the comparative analysis of different machine learning algorithms for diagnosis of different diseases such as heart disease, diabetes disease, liver disease, dengue disease and hepatitis disease. It brings attention towards the suite of machine learning algorithms and tools that are used for the analysis of diseases and decision-making process accordingly. In their work, the concluded that Statistical models for estimation that are not capable to produce good performance results have flooded the assessment area. Statistical models are unsuccessful to hold categorical data, deal with missing values and large data points. All these reasons arise the importance of MLT. ML plays a vital role in many applications, e.g. image detection, data mining, natural language processing, and disease diagnostics. In all these domains, ML offers possible solutions. This paper provides the survey of different machine learning techniques for diagnosis of different diseases such as heart disease, diabetes disease, liver disease, dengue and hepatitis disease. Many algorithms have shown good results because they identify the attribute accurately. From previous study, it is observed that for the detection of heart disease, SVM provides improved accuracy of 94.60%. Diabetes disease is accurately diagnosed by Naive Bayes. It offers the highest classification accuracy of 95%. FT provides 97.10% of correctness for the liver disease diagnosis. For dengue disease detection, 100% accuracy is achieved by RS theory. The feed forward neural network correctly classifies hepatitis disease as it provides 98% accuracy. Survey highlights the advantages and disadvantages of these algorithms. Improvement graphs of machine learning algorithms for prediction of diseases are presented in detail. From analysis, it can be clearly observed that these algorithms provide enhanced accuracy on different diseases. This survey paper also provides a suite of tools that are developed in community of AI. These tools are very useful for the analysis of such problems and also provide opportunity for the improved decision making process.

Obot and Inyang (2015) carried out research on an intelligent clustering based methodology for confusable diseases diagnosis and monitoring. The combination of non-specific clinical manifestations that characterize confusable tropical disease and the probable lack of expertise and experience among physicians exponentially increases the potential for misdiagnosis and subsequent increased morbidity and mortality rates resulting from these diseases. In this paper, an intelligent system driven by fuzzy clustering algorithm and Adaptive Neuro-Fuzzy Inference System for the investigation, diagnosis and management of similar and confusing symptoms of confusable diseases was developed. Data on patients diagnosed and confirmed by laboratory tests of viral hepatitis (H), malaria (M), typhoid fever (T) and urinary tract

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

infection (U) were used for training, testing and validation of the system. The system assigns patients with severity levels in all the clusters. Results on clusters validity are satisfactory. Overlapping symptoms analysis shows that symptoms of both H and T have highest degree of overlapping while symptoms common to M and U yielded the least impact. Symptoms common to M, H and T only, have equal impact with that of M, T and U only. The symptoms that are common to all the four diseases under study yielded a 12.8% contribution to the degree of severity of each of the CTD diseases.

Aniobi et al. (2020) Development of an expert system framework for the diagnosis of confusable diseases using neutrosophicahp technique. According to Aniobi et al (2020) Due to the myriad complexities in medical diagnoses, there is therefore the need to develop Expert systems that will help ameliorate decision making in confusable diseases scenarios. Diagnosis has become a difficult procedure in healthcare management due to the influence of therapeutic uncertainties that arise from confusion in disease symptoms that occur between two or more diseases. This confusion stems from the overlaps in the disease symptomatic presentation and has led to improper diagnosis with various degrees of associated costs and in worst scenarios has led to the death of patients. In this research, we designed and implemented a framework for the diagnosis of confusing diseases using Neutrosophic-Analytic Hierarchy Process (AHP) technique.

The app was implemented using Java EE and MySQL database as the knowledge store; the object oriented software engineering method used was employed in development. The developed expert framework using neutrosophic-AHP technique is indeed a breakthrough in diagnosis of confusing diseases as the results obtained from tests carried out shows that it is able to handle the indecision and uncertainties in making diagnosis/conclusion due to the presence of overlapping symptoms mapping different diseases, it equally aid in prescription upon diagnosis of ailment. In order to make apposite, rational and fitting medical decision in the diagnosis of confusable diseases, the knowledge base and the inference mechanism play an indispensable role as they are the core of clinical decision support systems. Several test cases were conducted to test the accuracy of the system; the results so far show that the introduction of Neutrosophic-AHP to medical diagnosis will truly help relieve Doctors the burden of indecisiveness and uncertainty condition they encounter at the course of diagnosis confusable diseases.

Toğaçar et al (2019) used the lung X-ray images. They employed the convolutional neural network as feature extractor and used some existing convolutional neural network models like AlexNet, VGG-16 and VGG-19 to realize a specific task. The number of deep features was reduced using deep model. Then a step of classification was done using decision tree, KNN, linear discriminant analysis, linear regression, and SVM. The obtained results showed that the deep features provided robust and consistent features for respiratory disease detection. In

Liang and Zeng (2020), proposed a new deep learning framework to classify child respiratory disease image by combining residual thought and dilated convolution. To overcome the over-fitting and the degradation problems of the depth model and the problem of loss of feature space information caused by the increment in depth of the model, the proposed method used the residual structure and dilated convolution respectively.

Jaiswal et al., (2019) proposed a deep learning-based method to identify and to localize the pneumonia in Chest X-Rays images. The identification model is based on Mask-RCNN that can incorporate global and local features for pixel-wise segmentation. Good performances evaluated on chest radiograph dataset showed the effectiveness and robustness of the model.

The investigation of post-stroke pneumonia prediction models using advanced machine learning algorithms, specifically deep learning approaches has been presented in Ge et al work in (2019).

Indeed, authors have used the classical methods (logistic regression, support vector machines, extreme gradient boosting) and have implemented methods based on multiple layer perceptron neural networks and recurrent neural network to make use of the temporal sequence information in electronic health record systems. The obtained results showed that the deep learning-based predictive model achieves the optimal performance compared to many classic machine learning methods.

Sirazitdinov et al., (2019) proposed an automated detection and localization method of acute respiratory disease on chest x-ray images using machine learning solutions. They presented two convolutional neural networks (RetinaNet and Mask R-CNN). The proposed method was validated on a dataset of 26,684 images from Kaggle Pneumonia Detection Challenge and the obtained results were satisfactory.

Khalid et al. (2020) proposed a Deep Learning framework for examining lung pneumonia and cancer. Indeed, they proposed two different deep learning techniques: the first one was Modified AlexNet. It was intended to classify chest X-Ray images into normal and pneumonia class using Support Vector Machine and its performance were validated with other pretrained deep learning (AlexNet, VGG16, VGG19 and ResNet50).

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

Whereas the second one implemented a fusion of handcrafted and learned features in the MAN for improving classification accuracy during lung cancer assessment. The work presented in

Behzadi-khormouji et al. in (2020) presented a method that can detect consolidations in chest x-ray radiographs using Deep Learning especially Convolutional Neural Networks assisting radiologist for better diagnosis. Authors have used a Deep Convolutional Neural Network pre-trained with ImageNet data to improve the accuracy of the models. Then, to enhance the generalization of the models, they proposed three-step pre-processing approach.

When we use DL and image processing, especially in medical imaging, a natural question arises: how we use this DL, and which is the best architecture of DL will be used? Following the context of XRay images classification, we present a comparison of recent Deep Convolutional Neural Network (DCNN) architectures for automatic binary classification based fined tuned versions of (CNN, VGG16, VGG19, DenseNet201, Inception_ResNet_V2, Inception_V3, Xception, Resnet50, and MobileNet_V2.

2.5 SUMMARY OF REVIEW

S/N	AUTHOR	YEAR	RESEARCH WORK	CONTRIBUTION(S)	LIMITATION(S)
1	James, et al.	2016	Fuzzy cluster means algorithm for the diagnosis of confusable disease	The system gives a measure of each disease within a set of confusable disease. The proposed system had a classification accuracy of 60%.	Their work did not use a well define data to ascertain the classes of disease in the diseases syndromes
2	Faith-Michael et al	2017	A framework for early differential diagnosis of tropical confusable diseases using the fuzzy cognitive map engine.	Their system also takes cognizance of some risk factors that could pre-dispose an individual to infection. In addition, it recognizes the sematic causative relationships among symptoms, and can account for comorbidity in the seven diseases under consideration.	This system did not used a competent soft computing algorithm that guarantees effective diagnosis.
3	Fatima and Pasha	2017	A survey of Machine Learning Algorithms for Disease Diagnostic.	Their work provides the comparative analysis of different machine learning algorithms for diagnosis of different diseases such as heart disease, diabetes disease, liver disease, dengue disease and hepatitis disease. It brings attention towards the suite of machine learning algorithms and tools that are used for the analysis of diseases and decision-making process accordingly.	There was no new technological innovation in diagnosis of confusable disease, since the work was just a review of previous technology used

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 12, December 2024
DOI: 10.17148/IJARCCE.2024.131240

4	Obot and Inyang	2015	An intelligent clustering based methodology for confusable diseases diagnosis and monitoring.	An intelligent system driven by fuzzy clustering algorithm and Adaptive Neuro-Fuzzy Inference System for the investigation, diagnosis and management of similar and confusing symptoms of confusable diseases was developed.	The system was not able to compares favorably with diagnosis arrived at by experienced physicians and also provides patients" level of severity in each confusable disease and the degree of confusability of any two or more confusable diseases.
5	Aniobi et al.	2020	Development of an expert system framework for the diagnosis of confusable diseases using neutrosophicahp technique.	Ability to designed and implemented a framework for the diagnosis of confusing diseases using Neutrosophic-Analytic Hierarchy Process (AHP) technique.	Diagnosis with this solution was a difficult procedure due to the influence of therapeutic uncertainties that arise from confusion in disease symptoms that occur between two or more diseases.
6	Toğaçar et al.	2019	A robust hybrid deep convolutional neural network FOR covid-19 disease identification FROM chest x-RAY images	They employed convolutional neural network as feature extractor and used some existing convolutional neural network models like AlexNet, VGG-16 and VGG-19 to realize a specific task.	Their work did not diagnize ARDSs to know their actual signs but rather, Covid-19 disease
7	Liang and Zeng	2020	Automated Methods for Detection and Classification Pneumonia based on X-Ray Images Using Deep Learning	They proposed a new deep learning framework to classify child respiratory disease image by combining residual thought and dilated convolution. To overcome the over-fitting and the degradation problems of the depth model and the problem of loss of feature space information caused by the increment in depth of the model, the proposed method used the residual structure and dilated convolution respectively.	Their work did not use a well define data to ascertain the classes of disease in the diseases syndromes but based all research on dictating and classifying Pneumonia, a single type of ARDS
8	Jaiswal et al.	2019	Identifying pneumonia in chest X-rays: A deep learning approach	They proposed a deep learning-based method to identify and to localize the pneumonia in Chest X-Rays images. The identification model is based on Mask-RCNN that can incorporate global and local features for pixel-wise segmentation. Good	Their work did not use a well define data to ascertain the classes of disease in the diseases syndromes but based all research on identifying

DOI: 10.17148/IJARCCE.2024.131240

				performances evaluated on chest radiograph dataset showed the effectiveness and robustness of the model.	Pneumonia in chest, a single type of ARDS
9	Ge et al.,	2019	Predicting post-stroke pneumonia using deep neural network approaches	They investigated the post-stroke pneumonia prediction models using advanced machine learning algorithms, specifically deep learning approaches is captured	Their work did not use a well define data to ascertain the classes of disease in the diseases syndromes but based all research on predicting post-stroke pneumonia, a single type of ARDS
10	Sirazitdinov et al.	2019	Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database	They proposed an automated detection and localization method of acute respiratory disease on chest x-ray images using machine learning solutions. They presented two convolutional neural networks (RetinaNet and Mask R-CNN). The proposed method was validated on a dataset of 26,684 images from Kaggle Pneumonia Detection Challenge and the obtained results were satisfactory.	Their work did not use a well define data to ascertain the classes of disease in the diseases syndromes but based all research on localizing Pneumonia from a large-scale chest x-ray.
11	Khalid, et al.	2020	Using X-ray images and deep learning for automated detection of coronavirus disease	They proposed a Deep Learning framework for examining lung pneumonia and cancer. Indeed, they proposed two different deep learning techniques: the first one was Modified AlexNet. It was intended to classify chest X-Ray images into normal and pneumonia class using Support Vector Machine and its performance were validated with other pretrained deep learning (AlexNet, VGG16, VGG19 and ResNet50). Whereas the second one implemented a fusion of handcrafted and learned features in the MAN for improving classification accuracy during lung cancer assessment.	Their work did not use a well define data to ascertain the classes of disease in the diseases syndromes but based all research on using X-ray images and deep learning for automated detection of coronavirus disease

Table 1: Summary Evaluation of Literature Review.

DOI: 10.17148/IJARCCE.2024.131240

Moreover, looking at the review of related works done by other reserchers above, we've evaluated the current "state of the art" for the body of knowledge reviewed, pointing out major areas covered, inconsistencies in theory and findings, and we then notice that there are areas or issues pertinent to future study which includes developing a soft computing model for early detection of acute respiratory diseases and that is the research topic we area working on.

2.6 Literature Gapes

In our research, reviewed numerous works done by researchers on acute respiratory diseases and we noticed that there are areas yet untouched or under-explored, theses areas include: records on the levels of increase or decrease on the death tolls of acute respiratory disease patience, the gender that are mostly affected and the age bracket of people who are mostly affected. They did not also capture the location with the highest prevalence of the disease. So, we will capture all this in our research and draw an inference on them.

III. SYSTEM ANALYSIS AND DESIGN

A research methodology is a systematic programming approach of well-defined procedure that should be followed in carrying out a thorough research project. An adequate and suitable methodology would ensure a very detailed research work and a higher degree of accuracy. It will helps to know if there should be a total overhauling of the existing system or for us to only carry out an improvements only.

3.1 Research Methodology

After due consideration of the reasons in the introduction aabove, Analytical and Empirical Research methodology (AERM) will be adopted in our research. This methodology is important for we need to carry out a thorough, adequate and comprehensive evaluation of the existing system with a view to identifying its strengths and weakness. For a sensitive topic of this sort which concerns human health, we need such methodology to avoid wrong medication prescription.

Reasearch Methods

Data Collection:

- Gather historical patient data, including symptoms, age, duration, and diagnosis labels.
- Samples of features: gender, age, duration of illness, and binary symptoms like cough, fever, etc.

Data Pre-processing:

- **Scaling**: Standardize input data using Standard Scaler.
- **Encoding**: Convert categorical variables (e.g., gender) to numeric values.
- **Feature Selection**: Choose features relevant to respiratory conditions.

Model Development:

- **Fuzzy C-Means (FCM)**: Use for clustering symptoms to identify distinct patterns (e.g., clusters for mild, moderate, and severe cases).
- **Machine Learning Model**: Train a classifier (e.g., Decision Tree, Random Forest) to predict conditions based on input data.

System Integration:

- Integrate models into a user-friendly application using Streamlit for real-time prediction and clustering.

Sources of Data

We shall visit and dialog with the Virologists and Medical Lab Technologist in various Hospitals to track their expert knowledge about the Acute respiratory Disease (ARS) and ask for records of cases of acute respiratory diseases in their hospital.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 12, December 2024
DOI: 10.17148/IJARCCE.2024.131240

Method of Data Collection

1. We will develop research questionnaire to collect data from other medical personals who are inclined or have abilities relating to this research area.
2. We shall consult journals, textbooks, and online materials, which will enable us to track the records of other researchers on this research area.

Method of Data Analysis

1. A MatLab Software Tool will be used to train or analyze the data and develop a Model of the intended/new system.
2. We then implemented our findings by Model development using Python Programming Language.

Presentation of Data

We shall present the data collected in excel format capturing the various diseases on the rolls and their signs on the columns and a definition of the data captured will be giving for clearer understanding of the calssees of the diseases.

3.2 Analysis of the Existing System

In the case of acute respiratory diseases, it is common for people to die due to incorrect diagnosis or incorrect administration of medication. When people suddenly fall ill with these acute respiratory diseases, they quickly assume a certain type of disease based on the signs they feel at that moment, which leads to self-medication or treatment. On the other hand, the researchers have adopted a suggestive/predictive model that classifies these diseases, their system can recognise the causal semantic relationships between the symptoms and can reflect the compatibility in the seven diseases they considered, and they have also developed a model that carries out performance evaluation on the chest X-ray dataset, which shows the effectiveness and robustness of the model. They use machine learning techniques to propose an automated method for diagnosis and localisation of acute respiratory diseases using chest X-ray images. Therefore, we are stirred to develop a model that can group these different diseases and categorise them into their different categories to facilitate access and prompt diagnosis of these respiratory diseases.

Weakness of the Existing System

One of the drawbacks of the exixting system or method is that it is mostly based on assumptions or suggestions. They were not able develop a system suitable to diagnose the acute respiratory diseases types and determine their primary causes or identify the main feature that distinguishes the diseases, as they are all similar. Our system should therefore be able to fill these gaps. It should be able to summarise these many diseases and their signs, and then classify the signs according to their pathological syndrome.

The weakness of the existing system is also captured as below:

- **Reliance on Expertise**: Diagnoses often depend on human expertise, which is prone to variability and error.
- **Time-Consuming**: Manual examination and testing delay diagnosis.
- **Inaccessible Technology**: Advanced diagnostic tools may not be available in low-resource settings.
- **Limited Data Utilization**: Most systems lack robust algorithms for symptom analysis and clustering.

3.3 Analysis of the Proposed System

1. Our model will hinge on experiential data obtained from physicians. 60% of the data would be used as a ‘training set’ in model development, while 40% would be utilized as a ‘test set’.
2. The model, which will be based on the FCM procedure will attempt to do the following:

a) Utilize some brute-force algorithmic procedure to mimic the mental algorithm employed by physicians in the diagnosis process.
b) Minimize the number of steps in the initial investigation process in order to arrive at a hypothesis.
c) Increase the efficiency and effectiveness of diagnosis by considering a broader spectrum of symptoms and affecters, and implementing meta-hypothesis that increases physician’s confidence in the diagnosis process.
d) Groups symptoms based on similarity, allowing early-stage identification of potential conditions.
e) Example: Symptoms like "cough, fever, shortness of breath" may belong to a "Severe Acute" cluster.

FCM has been chosen for this study for the following reasons:

▪ It is a dynamic modeling tool that helps in representing knowledge in an environment of uncertainty and imprecision. John and Innocent (2015) considered the aetiology of diseases and demonstrated that FCMs can encode fuzzy causal structures in order to support symptom information elicitation and diagnostic reasoning.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

- FCMs can exhibit enhanced usefulness through their ability to evolve dynamically through recurrent feedback and the facility to combine knowledge bases through the union of a number of FCMs. Khan et al. (2017), Khan and Quaddus (2014), Papageorgiou et al. (2019)
- FCM is one of the few decision support systems that are characterized by fast computation, which provides real-time results, and provides fast and crisp modeling methodology for creating and simulating qualitative dynamic systems with feedback. Tabar (2017), Carvalho and Tome, (2019).

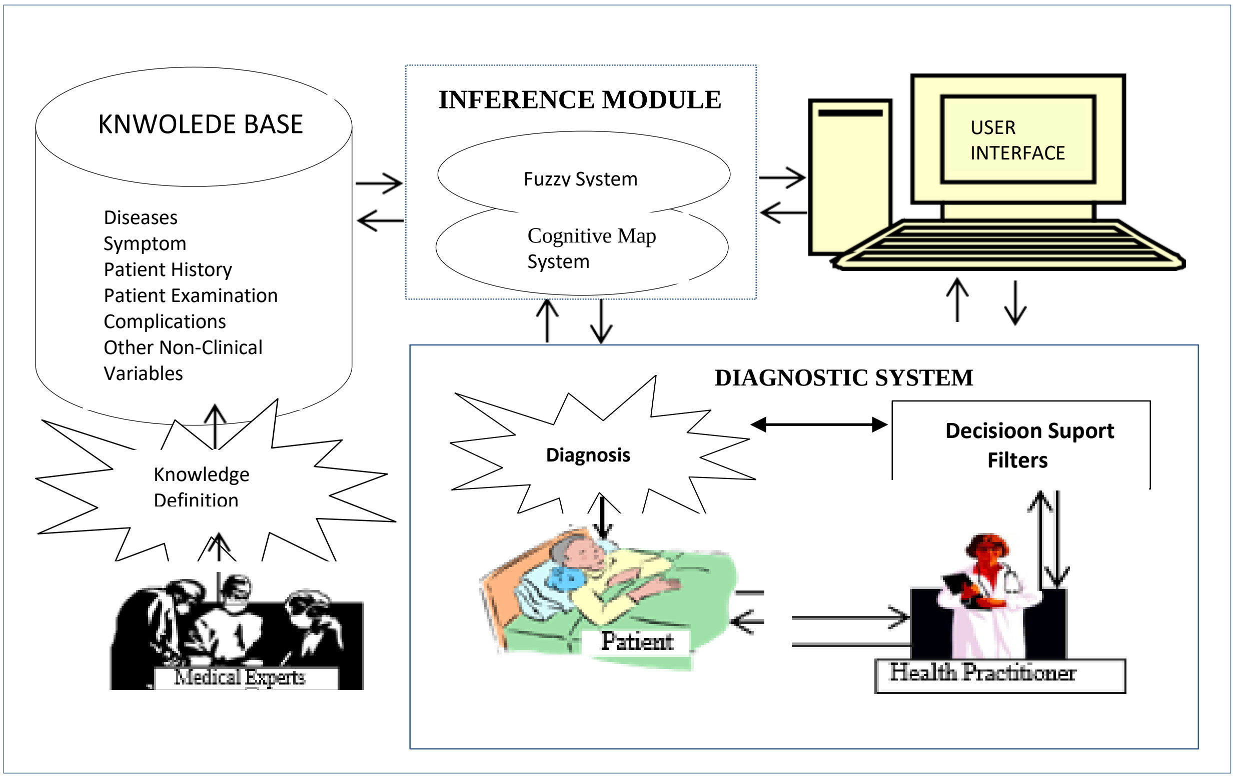
The architecture of our proposed system in Figure 1 below is adapted from the hybridization model in Uzoka and Barker (2017) and the FCM model with learning and feedback mechanism presented by Papageorgiou et al. (2018). It consists of the knowledge base, inference module; diagnostic system; and user Interface.

3. Machine Learning Predictions will also be adopted to**:**

- Predict specific conditions using a trained model.
- Example: Predict "Bronchopneumonia" or "Asthma" based on patient inputs.

4. There will be Ease of Use which will**:**

- Streamlit app enables healthcare professionals to input symptoms and receive predictions/clusters instantly.



3.4 Design of the Proposed System

The method employed for this work is the Object-Oriented Analysis and Design (OOAD). This is the procedure of identifying software engineering requirements and developing software specifications in terms of a software system's object model which comprises of interacting objects. An object-oriented software development methodology is a software engineering process that passes through the generic development phases of analysis, design, implementation and testing, based on analyzing existing methodologies and techniques, identifying their strengths and weaknesses, and producing a set of requirements defining the characteristics of the target methodology. The method can then be developed through making utmost use of existing techniques in such a way as to satisfy the requirements.

Phases In Object-Oriented Software Development

The major phases of software development using object–oriented methodology are object-oriented analysis, object-oriented design, and object-oriented implementation.

DOI: 10.17148/IJARCCE.2024.131240

Object–Oriented Analysis

In this stage, the problem is formulated, user requirements are identified, and then a model is built based upon real–world objects. The analysis produces models on how the desired system should function and how it must be developed. The models do not include any implementation details so that it can be understood and examined by any non–technical application expert.

Object–Oriented Design

Object-oriented design includes two main stages, namely, system design and object design.

System design

In this stage, the complete architecture of the desired system is designed. The system is conceived as a set of interacting subsystems that in turn is composed of a hierarchy of interacting objects, grouped into classes. System design is done according to both the system analysis model and the proposed system architecture. Here, the emphasis is on the objects comprising the system rather than the processes in the system.

Object design

In this phase, a design model is developed based on both the models developed in the system analysis phase and the architecture designed in the system design phase. All the classes required are identified. The designer decides whether:

- New classes are to be created from scratch,
- Any existing classes can be used in their original form, or
- New classes should be inherited from the existing classes.

The associations between the identified classes are established and the hierarchies of classes are identified. Besides, the developer designs the internal details of the classes and their associations, i.e., the data structure for each attribute and the algorithms for the operations.

Object–oriented implementation and testing

In this stage, the design model developed in the object design is translated into code in an appropriate programming language or software tool. The databases are created and the specific hardware requirements are ascertained. Once the code is in shape, it is tested using specialized techniques to identify and remove the errors in the code.

Below are other important part of the proposed system design and their definitions:

a) **System Architecture**:
- **Input Module**: Collects user data (age, gender, symptoms, etc.).
- **Pre-processing Module**: Standardizes and prepares input data.
- **Prediction Module**: Predicts conditions using a trained ML model.
- **Clustering Module**: Assigns patients to clusters using FCM.
- **Output Module**: Displays results as predictions and membership scores.

b) **Data Flow Diagram**:
- Show how data flows from input to preprocessing, prediction, clustering, and output.

c) **UML Diagrams**:
- **Class Diagram**: Represent classes such as Patient, Cluster Model, and Predictor.
- **Sequence Diagram**: Illustrate interactions between the user, input interface, and backend logic.

3.5 Model Framework

Three AI models will be used to denote the probability of having a defined defined type of ARDS. The models included two major modules: Judgment of ARDS based on clinical data and Extreme Definition & Meaning (eXtreme) of the clinical data as a support method

3.6 Modeling of the Proposed System

The Figure below shows the **use case diagram** for the proposed system.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

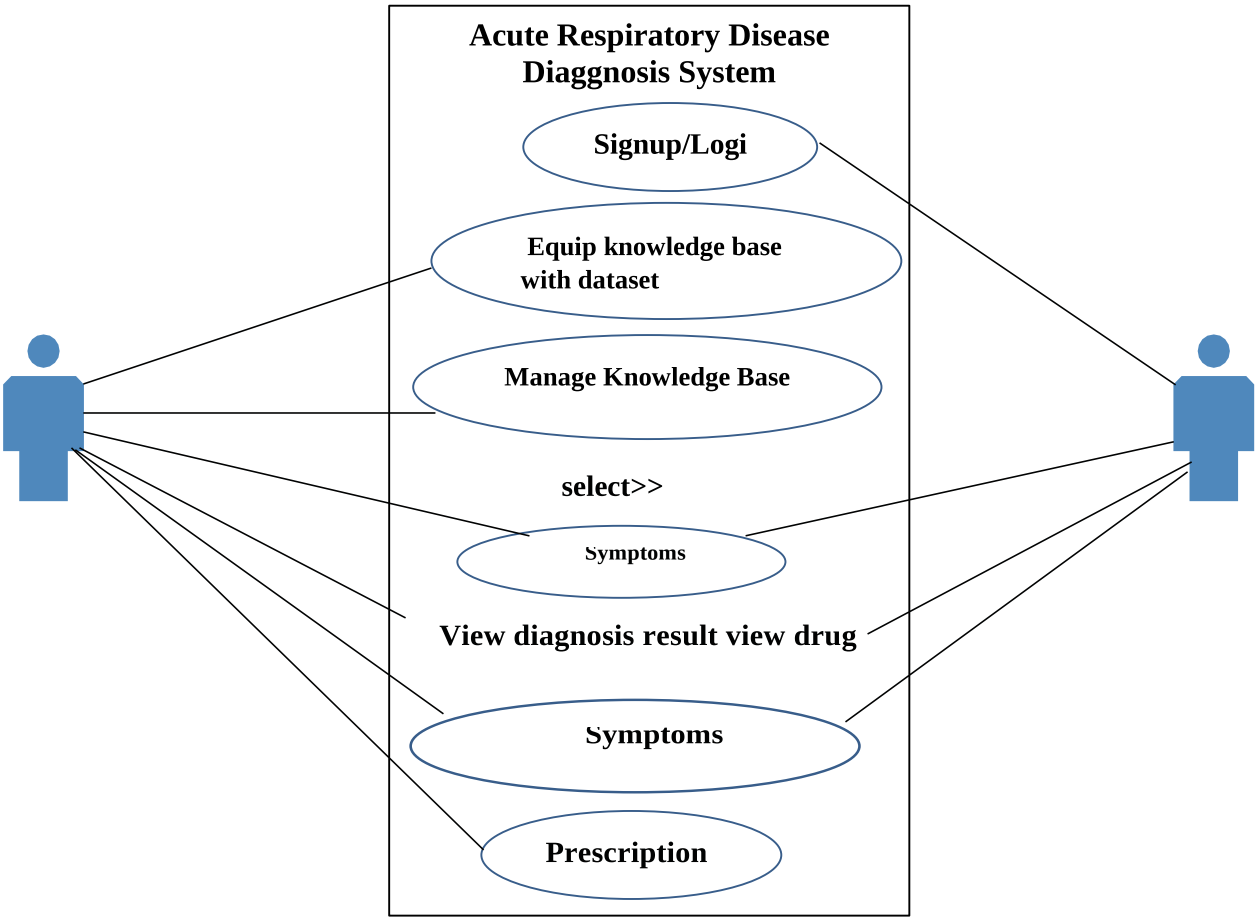


Figure2: Proposed Use Case Diagram (UCD):

3.7 Data Collection (Using Record Data, DataSet)

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes. Questionnaires and other Survey Methods are to be used in doing that. This research work used interviews method by visiting various Specialist hospitals across the country to make findings about the classes of diseases in the classes of severe respiratory track diseases. We shall also visit the various isolation centers to observe the patients, note down the symptoms of the diseases thaey are suffering from, note their attributes, as well as gather data from those centers. A readily collected data shall be sorted for to enable us have a face hand data to train.

The enrolled patients to be considered are those who had been reviewed as with or without ARDS by the working group from October 2019 to December 2023. They also met the following criteria:

(1) Age ≥ 20 years old.

(2) Stayed in an ICU for more than 48h.

(3) Patients readmitted to an ICU and whose CXR images were missing will be excluded.

Expected total of 1577 patients will be included in the final data analysis. The patient selection flowchart will also be presented.

IV. SYSTEM IMPLEMENTATION

System implementation is the planned and orderly conversion from an existing system to a new one. Implementation also enable programmers to work modules in parallel and periodic testing and checking on performance of the whole system

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 13, Issue 12, December 2024
DOI: 10.17148/IJARCCE.2024.131240

in order to allow manageable growth in complexity without introducing untraceable bugs. It involves the training of users in the operation of the new system.

Therefore implementation is the total organization activities involved in the adoption, management and implementation of innovation. Basically, this is concerned with the orderly schedule of events and list of new system analysis and design implementation involving all the practical method of putting into work all the theoretical design and getting the system into operation.

On the other hand, in our effort to Develop a Soft Computing Model for Early Detection of Acute Respiratory Diseases, implementation is the act of brining in a new system to coordinate the internal procedures of information communication and data processing as it regard classifying acute respiratory diseases according to their various signs. At this stage of implementation, the model application developed is test-runned, and setting up sub-task like training users of the new system are carried out.

Finally, the implementation of the Soft Computing Model Developed for Early Detection of Acute Respiratory Diseases involves the following task lists:

1. Specification of hardware and software requirement
2. Evaluation of system test-run
3. Under-taking system maintenance and
4. System change over.

So, This chapter provides an in-depth explanation of the implementation process, highlighting the tools, methods, and testing strategies.

4.1 Choice of Development Environment

The new system will be developed in one of the high level programming languages known as Python Programming Language, that will help us implemented our findings by Model development. The Python programming language will equally be used for: Pre-processing data (Pandas, NumPy), Implementing FCM (fcmeans library), and Building the user interface (Streamlit).
The Libraries/Packages that will be employed include**:**

- Fcmeans for fuzzy data clustering.
- Scikit-learn for preprocessing and model training.
- Matplotlib for visualizations.
- Pickle for save and load trained models.

The Development Tools will be**:**

- Jupyter Notebook for model training and testing.
- Streamlit for web-based application.

Justification of Programming Language Used
The language is chosen because of their flexibility and ability to generate a well-defined user graphical interface. The languages inglude: SQL which is used for database query, MysQl Server (PHPmyAdmin) used as the data backend (database), and Random Forest Algooriithm which was used for the training of the dataset collected from the hospiitals.

4.2 Implementation Architecture

This is a piece of software that acts as a bridge between a particular architectural style and a set of implementation technologies. It is prototype implementation based on open-source tools using Python which support rapid development and fast prototyping.

It consists of three layers: Data Layer, the Process Layer, and the Output Layer consists of three layers: Data Layer, the Process Layer, and the Output Layer.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

The Main Menu

The main menu in the prediction app provides two options:

a) **Prediction Application**:
- Inputs: Age, gender, duration, and binary symptoms.
- Function: Predicts the likely respiratory condition using the ML model.

Figure3: The Prediction Software Menu

Fuzzy C-Means Cluster:
Inputs: Same as in Prediction Application
Function: Assigns the user to a symptom cluster using the FCM model.
The clustering is patterned in a way that colours are used to denote the primary, the associate, and the secondary signs which predicts a certain disease. The primary sign is that sign that uniquely defined the particular disease and it is not found as a sign/symptom of any other disease. The secondary and asociate signs are those signs whice can be found with some other disease.

The fuzzy clustering algorithm equally captured the duration or how long the patient have been surfering with the disease and the ages of the concerned patient during the clustering. Kindly see the clustering result below and the implementation program.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

IMPLEMENTATION ALGORITHM

```
# Visualize the clusters (optional - replace 'Feature1', 'Feature2' with actual features for visualization)
plt.figure(figsize=(10, 7))
for i in range(n_clusters):
    plt.scatter(data[data['Cluster'] == i][features[0]],
                data[data['Cluster'] == i][features[1]],
                label=f'Cluster {i}')
plt.xlabel(features[0])
plt.ylabel(features[1])
plt.title('Fuzzy C-Means Clustering Results')
plt.legend()
plt.show()
```

The Submenus

The submenu in the predition app provides two options:

- Prediction Inputs for feature fields which includes: gender, age, duration, and symptoms.

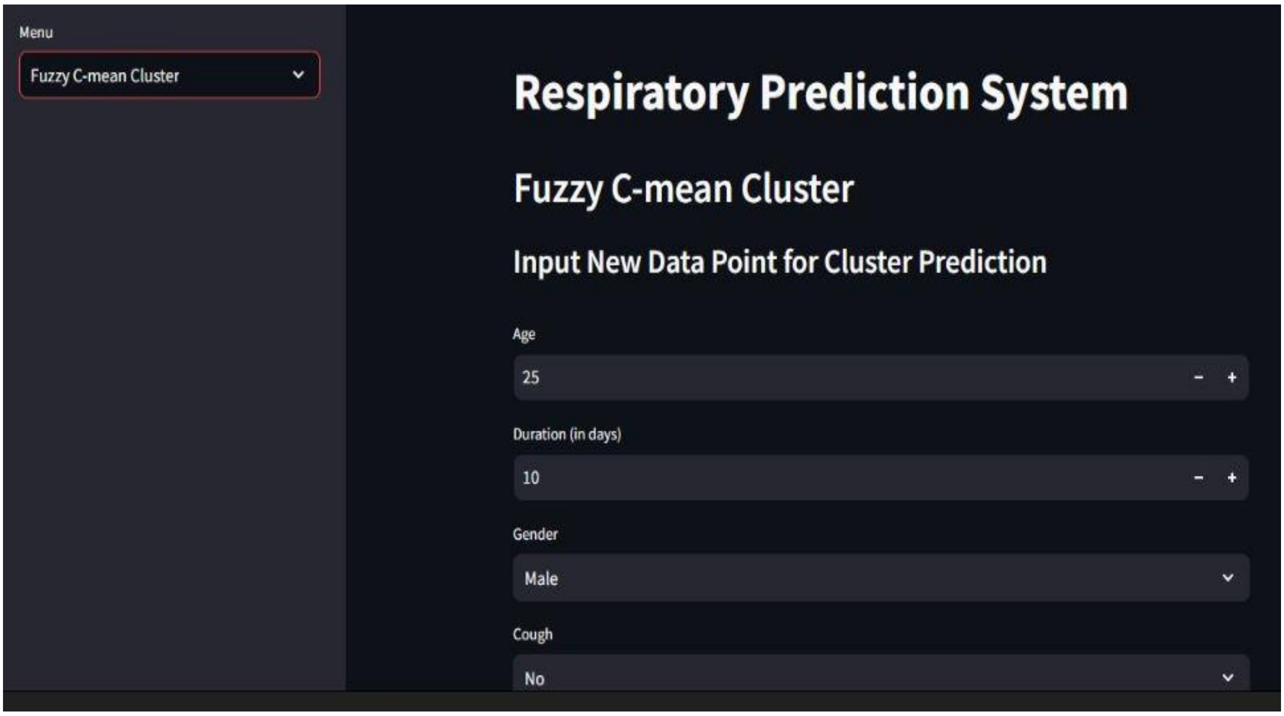


Figure5: The Prediction Software Submenu

- This is a case of clustering outputs for membership scores (the diseases) and assigned clusters. Colours are equally used to show the clusters and it is reduced to two which are the primary sign/symptom and the associated symptoms that clearly defines the disease. Please see the image and the implementation algorithm below:

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 12, December 2024
DOI: 10.17148/IJARCCE.2024.131240

IMPLEMENTATION ALGORITHM

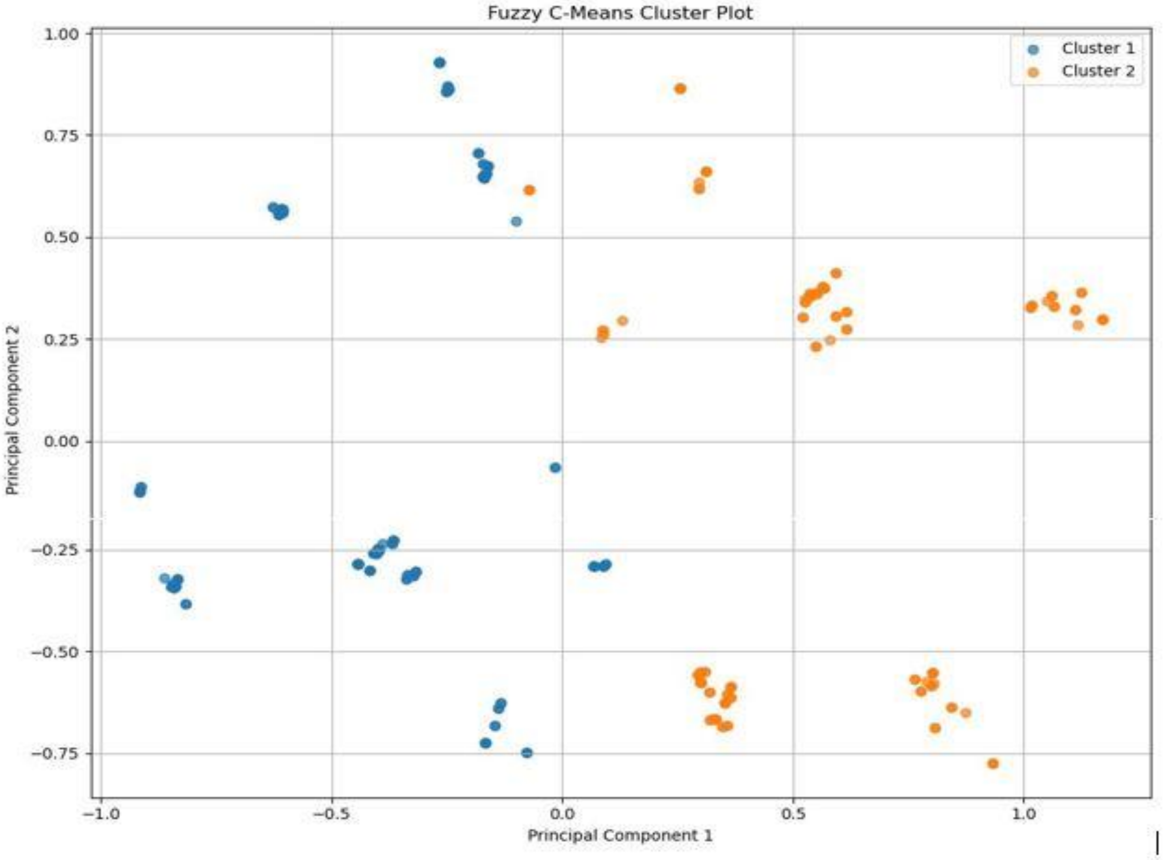
```
## plt.figure(figsize=(10, 7))
for i in range(n_clusters):
    plt.scatter(df[df['Cluster'] == i][features[0]],
                df[df['Cluster'] == i][features[1]],
                label=f'Cluster {i}')
plt.xlabel(features[1])
plt.ylabel(features[2])
plt.title('Fuzzy C-Means Clustering Results')
plt.legend()
plt.show()
```

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240



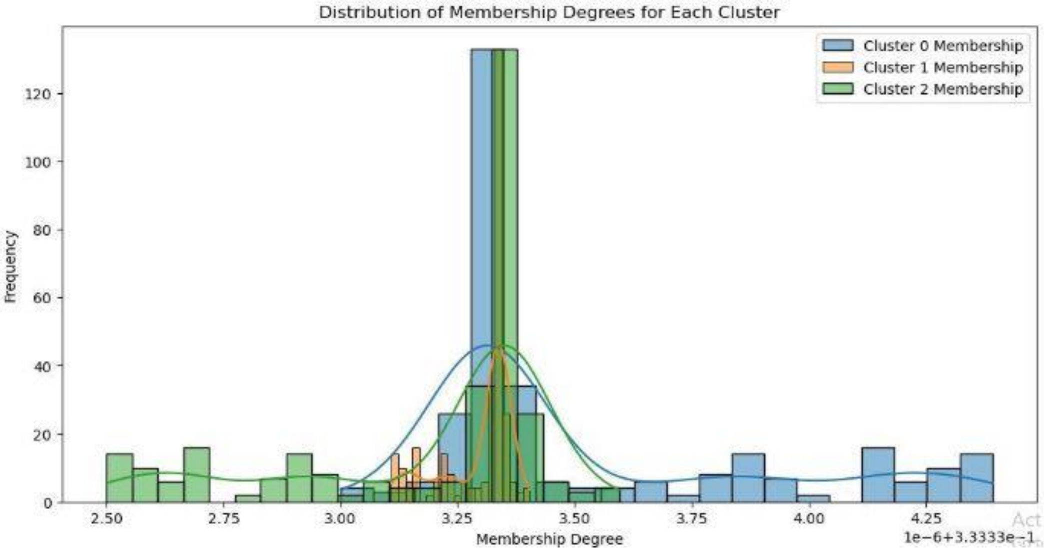
Implementation Algorithm

```
# Cluster Visualization using PCA
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X_normalized)

plt.figure(figsize=(10, 8))
for cluster in range(n_clusters):
    cluster_points = X_pca[df['Cluster'] == cluster]
    plt.scatter(
        cluster_points[:, 0],
        cluster_points[:, 1],
        label=f'Cluster {cluster + 1}',
        alpha=0.7,
    )

plt.title('Fuzzy C-Means Cluster Plot')
plt.xlabel('Principal Component 1')
plt.ylabel('Principal Component 2')
plt.legend()
```

The blue colours denotes the actual or the primary symptom/sign of the disease the patient is suffering from, while the orange colour shows the associate and the secondary signs of the disease which the patient is suffering from.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 12, December 2024
DOI: 10.17148/IJARCCE.2024.131240

Plotted Histogram of the Membership Values for Each Cluster



Implementation Algorithm

```
import seaborn as sns

# Plot histogram of membership values for each cluster
plt.figure(figsize=(12, 6))
for i in range(n_clusters):
    sns.histplot(u[i], kde=True, bins=20, label=f'Cluster {i} Membership')
plt.xlabel('Membership Degree')
plt.ylabel('Frequency')
plt.title('Distribution of Membership Degrees for Each Cluster')
plt.legend()
plt.show()
```

4.3 Software Testing

This is the process of assessing the functionality of a software program. The process checks for errors and gaps and whether the outcome of the application matches desired expectations before the software is installed and goes live. It is of the following types:

Unit Testing:
Used to test individual components like predict cluster () and model. Predict () functions.

Integration Testing:
Used to verify data flow between pre-processing, prediction, and clustering modules.

System Testing:
Used to test the overall performance of the app, including edge cases like missing or invalid inputs.

User Testing:
Used to gather feedback from healthcare professionals on usability and accuracy.

4.4 Documentation

This is the information that describes the product to the people who develop, deploy and use it. It includes the technical manuals and online material, such as online versions of manuals and help capabilities.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 12, December 2024
DOI: 10.17148/IJARCCE.2024.131240

User Manual

1. **Navigating the Interface**:
   - Select the "Prediction App" or "Fuzzy C-Mean Cluster" from the main menu.
   - Enter the patient's details in the fields provided.
2. **Understanding Outputs**:
   - **Prediction App**: Displays a condition diagnosis.
   - **Fuzzy C-Means Cluster**: Shows the assigned cluster and membership scores.

4.5 Source Code Listing (An Appendix)

Provide the full source code as an appendix. Include:

- **Prediction Code**: Implements condition prediction based on user inputs.
- **Clustering Code**: FCM implementation for symptom clustering.
- **Interface Code**: Streamlit app structure, menu creation, and user input handling.

V. PERFORMANCE EVELUATION AND DISCUSSION OF RESULT

This is a formal assessment in which the user of a software evaluates a software developer's deliverables performance to understand its strengths and weaknesses. Its goal is to provide feedback, identify areas for improvement, and recognize accomplishments on the side of the users of the softwares.

Ricardo M. C. et al (2019) states that evaluating a software product is an ongoing effort to define quantitative Performance Evaluation (PE) of the software, written to clarify the understanding of some concepts on terminology thus avoiding ambiguity in daily efforts when investigating systems characteristics. The intention of the work is to discuss dependability and select important notions applicable to Performance Evaluation and present means and approaches suitable for analysis and decision making.

On the other hand, result discussion section is where you delve into the meaning, importance, and relevance of your results. It focuses on explaining and evaluating what you found, showing how it relates to your literature review and paper or dissertation topic, and making an argument in support of your overall conclusion.

5.1 System Performance Evaluation Matrices and Tools Used

The system performance eveluation metrics provide insights into various aspects of system performance, including user experience, efficiency, and reliability. The key metrices in the prediction software we developped which we are to consider will include: Response Time, Throughput, Latency, Scalability, Resource Utilization, and Error Rates.

Response Time:
The time it takes for a system or person to react to a request or input. It can be measured in units of time, like milliseconds or seconds.

Throughput:
This is a measure of how many units of data a system can process in a given amount of time. It is applied broadly to systems ranging from various aspects of computer and network systems to organizations.

Latency:
This is an expression of how much time it takes for a data packet to travel from one designated point to another. It is in simple term, a synonym for delay.
In other words, it is the delay that happens between when a user takes an action on a network or web application and when it reaches its destination, which is measured in milliseconds.

Scalability:
This is the measure of a system's ability to increase or decrease in performance and cost in response to changes in application and system processing demands.

Resource Utilization:
This is a KPI (or metric) of resource planning used to help project managers and leaders understand performance and effort over a specific amount of time. It measures your team's productivity and can help you understand if your organization is over or underutilizing resources.

DOI: 10.17148/IJARCCE.2024.131240

Error Rates:
This is a measure of the degree of prediction error of a model made with respect to the true model. It can be calculated by add up all process-related errors in a reporting period and divide them by the total number of processes completed within the same reporting period.

Tools used on the other hand talks about the programming languages we use in the development of the prediction software and they include: SQL which we employed for the database query and MysQl Server (PHPmyAdmin) which we used as the data backend (database).

5.2 Performance Experiment and Measurement

This is a software engineering practice of managing the developmental processes of software which can be achieved by means of continuous planning, measurement, monitoring and assessment of process performance indicators. One can equally say that it is a software testing process characterized by numerous, well-known performance indicators, based on which it is possible to plan and measure its performance. It ensures that software meets specific performance requirements. For example, it can verify whether an application can handle thousands of users logging in or performing various actions simultaneously. It is an organization's resolve to achieve its goals.

Key Employee Performance Metrics to Track
Quantity: This talks about the amount of work produced.

Quality: Talks about the standard or excellence of work compared to others.

Efficiency: Talks about the ratio of input (time and money) to output (products or services)

Effectiveness: Talks about the impact of the work produced.

Steps Follows in Measuring the Performance of a Software:
- Employees Per Product Produced Eveluation.
- Customer/Employee Retention Rate Eveluation.
- Staff Morale Eveluation.
- Customer Satisfaction Eveluation.
- Overhead Ratio Eveluation.

Employees Per Product Produced Eveluation.
A performance measurement that calculates employee productivity by the number of products delivered within a timeframe. This method is simple, saves time, and works best for small businesses.

Step-By-Step Guide to Effectively Evaluating Employees
I. Set Performance Standards.
II. Set Specific Goals.
III. Take Notes Throughout the Year.
IV. Be Prepared.
V. Be Honest and Specific with Criticism.
VI. Don't Compare Employees.
VII. Evaluate the Performance, Not the Personality.
VIII. Have a Conversation.

Criteria for Employee Evaluation
i. Quality of work which is determined by its accuracy, thoroughness, and competence.
ii. Quantity of work which is determined by its productivity level, time management, and ability to meet deadlines.
iii. Job knowledge which is determined by skills and understanding of the work.

Customer/Employee Retention Rate Eveluation
An Employee retention rate is a metric that measures the percentage of employees who remain employed over a specific period. This vital measurement allows organizations to gauge the success of their human resource practices. It is measured in percentage.
To calculate the employee retention rate, we divide the total number of employees at the end of a period by the total number of employees at the beginning of that same period, and multiple the result by one hundred.

DOI: 10.17148/IJARCCE.2024.131240

NB: The higher your retention rate, the better.

Staff Morale Eveluation
This is a method adopted to assess employees' overall mood, satisfaction, and engagement levels within an organization. These evaluation aim to identify factors contributing to high morale and pinpoint areas causing dissatisfaction or disengagement.

Customer Satisfaction Eveluation
This determines how satisfied customers are with the services, goods, business, or customer service team. It helps to determine customers level of happiness regarding one or more aspects of the business. The answers taken from a CSAT are expressed as a percentage, ranging from 0 to 100%.

To evaluate or measure client satisfaction, we do a surveys. Customer satisfaction surveys can gauge the thoughts, feelings, and opinions of customers on ones company overall or with specific aspects of his/her product or service.

Overhead Ratio Eveluation
This is a measurement of the operating costs of doing business compared to the company's income.
This is evaluated by dividing the indirect costs by the direct costs and multiplying by 100. That means, if your overhead rate is 40%, it implies the enterprise spends 40% of its revenue on making a good or providing a service.
NB: A low overhead ratio indicates that a company is minimizing business expenses that are not directly related to production.

5.3 Performance Measurement of the New System
This is the process used to assess the efficiency and effectiveness of a newly developed software projects, programs and initiatives. It is a systematic approach to collecting, analyzing and evaluating how "on track" a new software project/program is to achieve its desired outcomes, goals and objectives.

Performance measurement of a newly developed software refers to the process of collecting, analyzing, and reporting information about the performance of the developed software. In the context of coding and software engineering, it involves evaluating various aspects of code efficiency, system performance, and development processes.

The Importance of Performance Measurement
Performance measurement is essential in software development for the following reasons:

Identifying Bottlenecks: It helps in identifying performance bottlenecks in code or systems, allowing developers to optimize and improve efficiency.

Ensuring Quality: By measuring performance, developers can ensure that their code meets quality standards and performs as expected.

Guiding Optimization: Performance metrics guide optimization efforts, helping developers focus on areas that will yield the most significant improvements.

Tracking Progress: For learners using platforms like AlgoCademy, performance measurement provides a way to track progress and identify areas for improvement.

Meeting Requirements: In professional settings, performance measurement ensures that software meets specified performance requirements and service level agreements (SLAs).

Key Performance Metrics in Software Development
Several key metrics are commonly used to measure performance in software development:

Code-Level Metrics
- **Time Complexity:** Measures how the execution time of an algorithm increases with the input size, often expressed using Big O notation.
- **Space Complexity:** Assesses the amount of memory an algorithm uses relative to the input size.
- **Cyclomatic Complexity:** Measures the number of linearly independent paths through a program's source code.
- **Code Coverage:** Indicates the percentage of code that is executed during testing.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⪾ Peer-reviewed & Refereed journal ⪾ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

System-Level Metrics

- **Response Time:** Measures the time taken for a system to respond to a user request.
- **Throughput:** Indicates the number of tasks or operations a system can handle in a given time period.
- **Latency:** Measures the delay between a user action and the system's response.
- **Resource Utilization:** Tracks the usage of system resources like CPU, memory, and disk I/O.

Process Metrics

- **Velocity:** Measures the amount of work completed in a given time period, often used in Agile methodologies.
- **Defect Density:** Calculates the number of defects per unit of code.
- **Lead Time:** Measures the time from the start of development to delivery.
- **Cycle Time:** Tracks the time from when work begins on a task to when it's completed.

5.4 Model Training And Evaluation

The data we used were randomly split five times into training datasets (80%) and validation datasets (20%). Each data point was only in one of the training and validation datasets. For the precdition model evaluation that we developed, the results of the classification performance are presented in terms of accuracy, sensitivity, specificity, and AUC. The calculations of accuracy, sensitivity, and specificity were as follows:

$$\text{Accuracy} = \frac{tp + tn}{t_p + t_n + f_p + f_n}$$

$t_p + t_n$

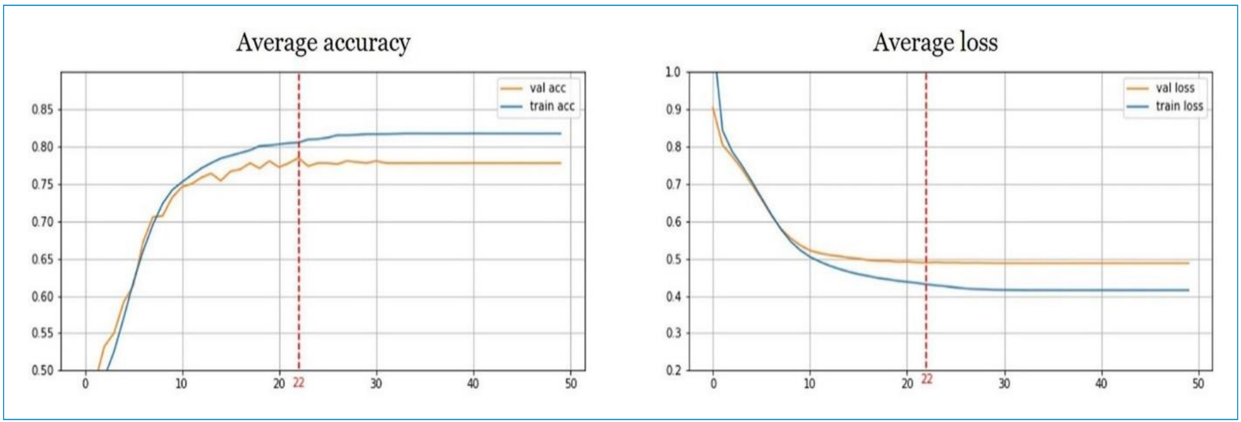
Sensitivity = $t_p + f_n$

$$\text{Specificity} = \frac{t_n}{t_n + f_p}$$

where true positive (t_p) is an ARDS patient classified as an ARDS patient, and false positive (f_p) is a non-ARDS patient classified as an ARDS patient. True negative (t_n) is a non-ARDS patient classified as non-ARDS, and false negative (f_n) is an ARDS patient classified as a non-ARDS patient. 645% were male. We also found out that 24.2% (383/1577) of the patients had ARDS. They had a higher average and stayed longer in the Intensive Care Unit (ICU).

5.5 Model performance

The prediction model we developed were evaluated using Random Forest Algorithm based on its accuracy**.** Values obtained after the prediction model for training accuracy, validating accuracy, training loss, and validation loss are shown in the figure below. It indicates the mean training accuracy of over 80% and the mean of training loss becomes stable after 22nd times training. The proposed model is stopped early in the 22th time to avoid overfitting.



Prediction model for training accuracy, validation accuracy, training loss, validation loss

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 13, Issue 12, December 2024

DOI: 10.17148/IJARCCE.2024.131240

Random Forest Algorithm

This is a commonly-used machine learning algorithm, trademarked by Leo Breiman and Adele Cutler. It combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems.

It is a supervised machine learning algorithm made up of decision trees which is used for both classification and regression. It has the following advantages over other tpyes of training algorithm:

a. It handle missing data without imputation, which is particularly beneficial in real-world datasets with common missing values.

b. The algorithm can maintain accuracy by estimating missing values based on the data's proximities, making it robust in scenarios where data quality might be compromised.

Its example is, classifying whether an email is “spam” or “not spam”.

Other algorithms that was examined before our resolve to use Random Forest are: Dynamic Programming Algorithm, Greedy Algorithm, Backtracking Algorithm, Brute Force Algorithm, Divide And Conquer Algorithm, Sorting Algorithms, Recursive Algorithm, Search Algorithms, Graph Algorithms.

5.6 Comparison of the New and the Existing System

A total of 1577 dataset gathered from different hopspitals were enrolled. Their mean ages of the affected patients was 66 ± 15 years and they are Scented in table below. Comparing the machine learning models constructed using clinical data, all classifiers had a good Area Under Cover (AUC) value and high model performance, with accuracy, sensitivity, specificity.

Moreover, the ensemble of three classifiers using average probability performed the highest specificity at 0.934 but the lowest sensitivity, at 0.676. The ensemble of three classifiers using maximum probability had the highest sensitivity of 0.849 and AUC of 0.912.

An Abridged form of ARDS Classification Results Based on Clinical Data.

```
sex,age,duration,cough,night_fever,catarrh,cold,tenderness,headache,black_stool,jerking_movement_of_the_hands,unable_to_sit_unsupported,throat_fe
M,13,7,1,1,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,MALARIA
M,13,7,1,1,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,URTI
F,4,14,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,HEAMORRHAGIC DISEASE
F,4,14,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,PERINATAL  ASPHYXIA
F,5,28,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,FERNATAL ASPYHXIA
F,5,28,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,NNS
M,5,7,1,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,MACROSOMU BABY
M,5,7,1,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,RVD-RETRO VIRUS DISEA
M,10,14,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,GET SEPSIS
F,12,21,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,SEVERE MALARIA
F,12,21,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,HYPERPYREXIA
M,17,14,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,MALARIA
M,17,14,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,URTI
M,21,14,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,ACUTE URTI
M,21,14,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,PNEUMONIA
F,12,49,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,MALARIA
F,12,49,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,BPN
F,12,21,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,DIARRHEA
F,12,21,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,DEHYDRATION
F,12,21,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,MALARIA
M,16,35,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,URTI
M,16,35,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,MALARIA
M,11,14,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,URTI
```

We captured the Sex of the patient, the Age, the duration, the signs that defines the ARD, and the ARDS itself

5.7 Discussion of Results

ARDS is a common disorder among critically ill patients requiring mechanical ventilation support. Sometimes, it is not recognized immediately and treated appropriately. We have successfully developed a machine learning prediction model

DOI: 10.17148/IJARCCE.2024.131240

to classify patients with ARDS. We used clinical data captured in a prediction models to enhance sensitivity and specificity. The important features relevant to clinical viewpoints were also explored. Our novel model design can be integrated into clinical practice in the future and remind clinicians to implement the optimal bundle of actions for ARDS at the right time.

ARDS is a highly prevalent clinical disorder with high morbidity and mortality rates in critically ill patients. The application of machine learning algorithms for the diagnosis and management of ARDS has emerged in previous years. ARDS is clinically and biologically heterogeneous, and phenotyping by experimental biomarkers potentially offers insights for prediction and treatment.[34] The application of machine learning models by using readily available clinical data can help classify ARDS phenotypes with high accuracy.[35] This may enable rapid phenotype identification at the bedside for prediction of diagnosis and possibly decisions about intervention.

VI. SUMMARY AND CONCLUSION

We have successfully developed a novel machine learning model to classify patients with or without ARDS based on the combined features of clinical data and CXRs. The model has been designed for a real-life scenario. We have started to integrate the model into our clinical practice and are conducting a study to investigate its impact on the outcomes of ARDS patients.

VII. CONCLUSION

In this work, we shall develop a model for diagnosing Acute Respiratory Diseases (ARD) driven by Fuzzy Cluster Means Algorithm. The work will be divided into two major parts; the first shall involve making use of the idea of SARS to be able to acquire the relevant knowledge from a medical practitioner from the stand point where it is uncertain about making decision from laboratory procedures in diagnosing a patient but have overlapping symptoms that creates doubts and indecision for the doctor in deciding the exact diseases based on presented symptoms. This knowledge shall be presented in the sets of membership function format capturing the certainty (truthfulness) of the doctor that a disease is true, the uncertainty and falsehood of the disease being true. The second phase shall be the passing of the knowledge base data set to the inference engine with the help of the application of Multi-Criteria Decision Making (MCDM) method to provide the exact disease.

We shall conduct due analysis of the existing systems, and spot out some limitations for considerations. Finally, the proposed framework shall provide an interface where the patient’s symptom will be captured by the system, the confusability measure will be calculated and in consultation with the knowledge base, the inference mechanism shall makes its diagnosis to the user and in turn make appropriate detection of the exact disease in the disease set.

RECOMMENDATION

We also used limited clinical features and limited data to avoid model complexity and overfitting and would be unlikely to leak brands. We found better results even with a limited set of data, I therefore request a more complex work to be done to improve on our effort.

Second, we collected only a small sample size data from various hospitals and conducted our study at a single center. External validation of multiple centers should be conducted in the future and visit should be made to more hospitals.

REFERENCES

[1]. Amit, K. J. et al, (2019) Identifying pneumonia in chest X-rays: A deep learning approach, Elsevier, Measurement 145, 2019, pp 511-518, www.elsevier.com/locate/measurement

[2]. Aniobi, D. E., et al, (2020) Development of an Expert System Framework for the Diagnosis of Confusable Diseases Using Neutrosophic-Ahp Technique, FULafia Journal of Science & Technology, ISSN: 24490954, Vol. 6 No.1 March 2020.

[3]. Carvalho J.P. andTomé J.A. (2019); Rule Based Fuzzy Cognitive Maps and Fuzzy Cognitive Maps - A Comparative Study, Proceedings of the 18th

[4]. Faith-Michael, E. U. et al (2016), A framework for Early Differential Diagnosis of Tropical Confusable Diseases using the Fuzzy Cognitive Map Engine, Conference Paper · February 2016. https://www.researchgate.net/publication/310141600

[5]. James, G.G., Okpako, A.E., and Ndunagu J.N., (2017) Fuzzy cluster means algorithm for the diagnosis of confusable disease, Journal of Computer Science and its application, ISSN: 2006-5523, Vol. 23 No. 1 (2017). https://www.ajol.info/index.php/jcsia/article/view/153911

[6]. John, R. I., and Innocent, P. R. (2005). Modeling uncertainty in clinical diagnosis using fuzzy logic. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 35(6), 1340-1350.

[7]. Khalid, E. A., et al, (2020) Automated Methods for Detection and Classification Pneumonia based on X-Ray Images Using Deep Learning, Conference Paper, 2020 https://www.researchgate.net/publication/340331870

[8]. Khan, M. S., and Quaddus, M. (2014). Group decision support using fuzzy cognitive maps for causal reasoning. Group Decision and Negotiation, 13(5), 463-480.

[9]. Khan, M. S., Chong, A., and Quaddus, M. (2017). Fuzzy cognitive maps and intelligent decision support–a review. School of Information Technology, Murdoch University, Graduate School of Business, urtin University of Technology, GPO Box U. Downloaded on December 10, 2013 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.7815&rep=rep1&type=pdf

[10]. Meherwar, F. and Maruf, P., (2017) Survey of Machine Learning Algorithms for Disease Diagnostic, Journal of Intelligent Learning Systems and Applications, ISSN: 2150-8410, pp 1-16, 2017. http://www.scirp.org/journal/jilsa

[11]. Obot, O. U., Inyang, U. G., **(2015)** An intelligent clustering based methodology for confusable diseases diagnosis and monitoring, Journal of Computer Science and Its Application, ISSN: 2006-5523, Vol. 22 No. 1 (2015). https://www.ajol.info/index.php/jcsia/article/view/150075

[12]. Okure U. Obot and Udoinyang G. Inyang (2015), an Intelligent Clustering Based Methodology for Confusable Diseases Diagnosis and Monitoring, Journal of the Nigerian Association of Mathematical Physics, Volume 30, (May, 2015), pp 271 – 280

[13]. Papageorgiou, E. I., Papandrianos, N. I., Apostolopoulos, D. J., and assilakos, P. J. (2018, June). Fuzzy cognitive map based decision support system for thyroid diagnosis management. In Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference on (pp. 1204-1211). IEEE.

[14]. Taber, R. (2017). Knowledge processing with fuzzy cognitive maps. Expert Systems with Applications, 2(1), 83-87.

[15]. Theodora, S. et al, (2014) A Robust Hybrid Deep Convolutional Neural Network for

[16]. COVID-19 Disease Identification from Chest X-ray Images, International Journal of Molecular Sciences, 2023, 14, pp 294-310 https://www.mdpi.com/journal/information

[17]. Yanqiu G., et al, (2019) Predicting post-stroke pneumonia using deep neural network approaches, International Journal of Medical Informatics, ISSN: 1386-5056, 2019. www.elsevier.com/locate/ijmedinf

[18]. Uzoka, F.M.E., and Barker , K. (2017). Expert Systems and Uncertainty in Medical Diagnosis: A Proposal for Fuzzy-ANP Hybridization. International Journal of Medical Engineering and Informatics , 2(2), 329-342.