Original Article

Fuzzy based classification and COVID-19 Patient's Data analysis

Manisha Shinde-Pawar[1], Alok Shah[2], Madhumita Nayak[3], S. A. Jadhav[4], Akhilesh Jadhav[5], Dhanashree Sahasrabuddhe[6], Shabana Memon[7], Anjali Gaikwad[8]

[1,3,5,6]Bharati Vidyapeeth (Deemed to be University), Pune, Institute of Management and Rural Development Administration, Sangli

[2]Bharati Vidyapeeth (Deemed to be University), Pune, Department of Management Studies, Navi Mumbai

[4,7]Bharati Vidyapeeth (Deemed to be University), Pune, Institute of Management, Kolhapur

Email: mjs.imrda@gmail.com

Abstract

COVID-19 respiratory viral disease has changed the entire research communities' focus towards Diseases, Health Care, Treatment and related resources decision making and services. Fuzzy approach gives simplification and improvement in processing logic and speed applied to Data analytics and testing of algorithm for different business purposes may be future research domain. Significant COVID-19 analytics can help to classify and prioritize the resources the future consequences to take enhanced self-ruled decisions, to recognise and design pattern in data spread to design strategic policies for medical, health care units and stakeholders.

Keywords: COVID-19, decision making, Fuzzy Logic, pattern, medical stakeholders.

1. Introduction

1.1 Background of study and Significance of Study

COVID-19 is a Respirational disease-causing contamination which grief specifically lungs, nose, and throat. Spanish flu (1918) which infected around one third populations of planet and now the COVID-19 also infected to large population of the world. Peoples infected experienced unimportant to reasonable symptoms and also so many recovered without any special treatment, but on the other hand so many patients lost their lives because of not exact, timely and proper treatments or existing medical problems and other threatening parameters in related field. The COVID-19 has banged the limitations of decision making in Respiratory viral infections treatments and decision making at different level of severity of such diseases. Technical solution can address different problems in manual decision making by analyzing voluminous data on few clicks can be combined with expertise skills and knowledge of medical professionals, pharmacists, pharma companies, pathologies and also diagnostic centres for treatment and decision making in such respiratory viral infections.

1.2 Symptoms Classification

COVID-19 affects different people in different ways. Most infected people will develop mild to moderate illness and recover without hospitalization.

Address for correspondence: Manisha Shinde-Pawar,
Institute of Management and Rural Development Administration, Sangli
Email: mjs.imrda@gmail.com

Access this article online

Quick Response Code:

Website: www.pnrjournal.com

DOI: 10.47750/pnr.2022.13.03.054

How to cite this article: Manisha Shinde-Pawar, Alok Shah, Madhumita Nayak, S. A. Jadhav, Akhilesh Jadhav, Dhanashree Sahasrabuddhe, Shabana Memon, Anjali Gaikwad, Fuzzy based classification and COVID-19 Patient's Data analysis, J PHARM NEGATIVE RESULTS 2022;13:339-354.

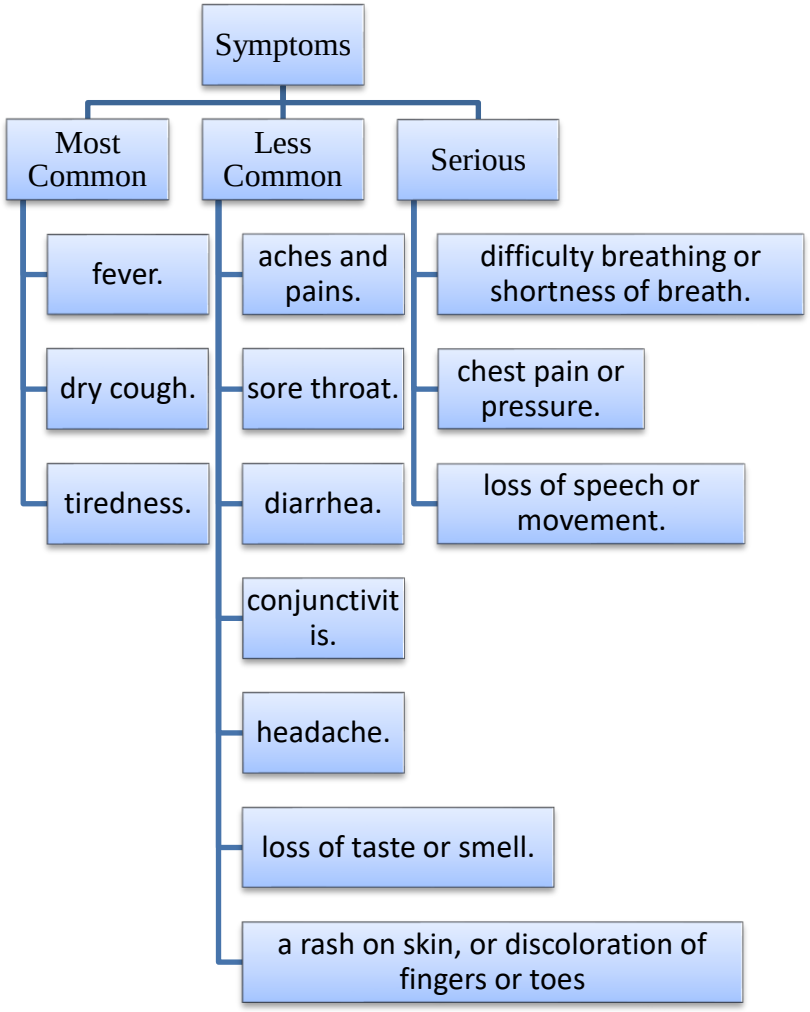


Figure 1

As shown in Figure 1 Symptoms can be observed in three categories:

Most common symptoms:

Less common symptoms:

Serious symptoms:

Till today there is no certain medicine or treatment is available, hence it is essential to create awareness among the stakeholders, experts, medical professionals, pharma's and related people about covid-19 symptoms, its treatment and how it impacts to various body organs after the treatment and what is the probability of recurrence of disease. This will support to these different professional with technical support system for decision making.

1.3 Digital Support / Solution to get the track of treatment

An existing research study helped to learn techniques and framework for fuzzy classification, Prioritization and implementation and machine learning techniques applied in different areas.

Fuzzy model is based on technique which has strong mathematical foundation. To deal with vague data fuzzy approach helps to reach to approximation of certainty from imperfection and uncertainty of data. COVID-19 data is also uncertain, vague and imperfect data. To capture similarity from certain classes of different degrees and to get dominance in data, it's always helpful to use fuzzy model.

1.4 Fuzzy Logic

Fuzzy Logic (FL) is a method of reasoning that resembles human reasoning. This approach is similar to how humans perform decision making. And it involves all intermediate possibilities between YES and NO. The Fuzzy logic works on the levels of possibilities of input to achieve a definite output.

One legacy artificial and machine learning technology is fuzzy logic. Traditional and classical logic typically categorize information into binary patterns such as: yes/no, true/false, or day/night. Fuzzy logic instead focuses on characterizing the space between these black-or-white scenarios.

1.5 Issues & Challenges

- Data categorization :

i. Infected and non-infected classification is general – COVID- 19Needs to deal with cases not exactly lying in two classes.

- Major Challenge for researcher:

i. Extracting and analyzing real time or vague data.

ii. Non Textual contents graphics, Audio, video, images etc.

iii. Imprecision, Incompleteness, variety, privacy and Security, Scale and Timeliness are confronted while managing COVID-19 data.

According to Stephen Kaisleret. al. the data stored with machine plays very important role in decision making and knowledge discovery. A major challenge for IT researchers and practitioners is that growth rate is fast exceeding our ability to both:

(1) Design appropriate systems to handle the data effectively

(2) Analyse it to take out important gist for choice construction.

1.6 Need of Analysis

COVID-19 cases outbreaks with voluminous and variety of data, discovering knowledge by analyzing this data helps in medical diagnosis, decision making and investigation process. But due to unstructured form of COVID-19 data there are so many challenges in analysis of data. Proper analytic solution provides in depth understanding of data which maximizes benefits to organizations and also maximizes value relationship with its stakeholders. The Computerized diagnosis, treatment decision making and prediction produces thought practice and impacts domain information.

1.7 Important Parameter of Analysis

- Infected and non-infected Categorization

The main terms identified as normal form of data after pre-processing unstructured data from COVID-19 and then further is compared with different Dictionary based or feature based data libraries to get the real attitude behind the text is Positive or Negative.

- Similarity and Pattern Detection

Dataset is generally evaluated to identify similarity in data so as consequent impact/decision can handled in solution form. So, Similarity detection techniques will be very much applicable to COVID-19 data behaviour analysis.

- Certainty

Covid -19 suddenly changed the way and parameters to deal with resources available for respiratory infectious disease particularly vast number of patients, comparatively short of medicines, hospitals, oxygen provisions to treat and care patients. Decision making in the complexity seems major challenge for medical stakeholders, governments to frame certain policies, to set priorities while facing the pandemic.

So, researcher deals with problems in circumstances of disease that are observed and captured as representations of any various feature or some part of reality of COVID-19 patient's admitted to Bharati Hospital, Sangli. When refining and smearing model to set of conditions to provide solution to real-world problem, the researcher will emphasis and effort to exploit worth of available data with aspect of different relations among uniqueness of trouble, integrity and ambiguity. Though not complete, this incomplete and vague data can be used with its implausibility as analytical and rigid way to maximize worth in decision making.

Application of Machine Learning with Fuzzy classification fuzzy cluster means, fuzzy k-means and Artificial neural network can be used for classification, prediction, grouping purpose to analyze to uncertain real-world problems like COVID-19 Pandemic.

Fuzzy helps to get the result from uncertainty to Certainty with different defuzzification methods and degrees of membership with interval ranges.

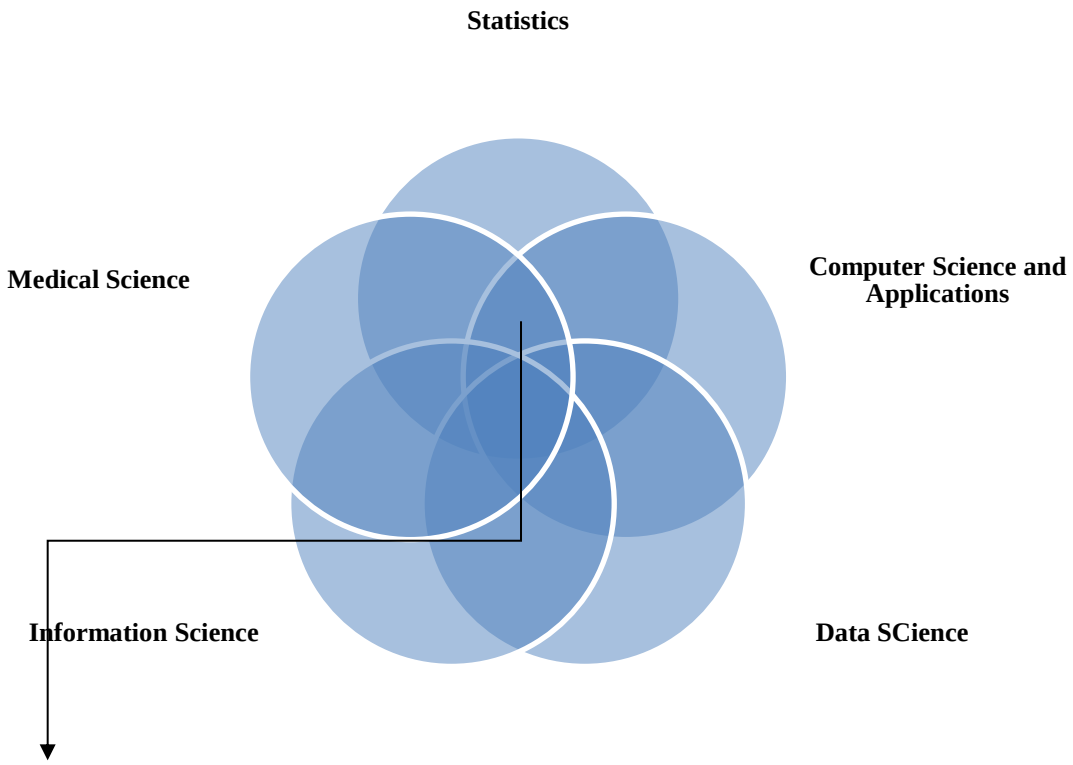
Mostly unstructured and varied form of COVID-19 data is uncertain and imprecise need to bring to certainty measures.

1.8 Ways of Analysis

COVID-19 data, symptoms, diagnosis and treatment related data is very recent and problem is also very recent. This data generated is being used by different medical practitioners and stakeholders to take decisions, to frame strategies and to deal with cases to beat the pandemic.

The figure No. 2, Data Analysis is proper mix of computer application of medical science, statistics, medical expert skills, data science and information science, ANN Techniques and associated data analysis policies. It needs theories and knowledge applications from various fields mainly computer science and machine learning application associated with statistics, medical science and information science.

Figure 2 Integration of technological advancement with various fields



Source: Compiled by Researcher

1.9 Applications of Programming languages for Data Analysis

Amarendra Babu[R-31] explained that, with aspect of powerful statistical base and computational efficiency for general purpose to specific purpose of data analysis, there are few top languages. Every language has different feature with respect of its usage.

Researcher has studied and traced some highly popular and recommended programming languages in recent data analysis and computational trends of data analysis. R is one of the most popular Programming languages in recent computational trends of Data Analysis. R is Open Source, Statistical Computing and Graphics, Machine Learning and Data Science, Microsoft, RStudio support to R Based Computing. R is widely used for Statistical Models and applications, Data Mining, Best choice of by Google and Facebook Companies.

1.10 COVID-19 Specific Functions

Certain systems and models designed provides specific functions (W-26) as listed below

- COVID-19 Case-based surveillance
- Contact registration & follow-up program
- Ports of Entry screening & follow-up program
- COVID-19 Surveillance Event Program
- COVID-19 Aggregate Surveillance

II. Fuzzy Approach for COVID-19 Analysis

Fuzzy Logic Process

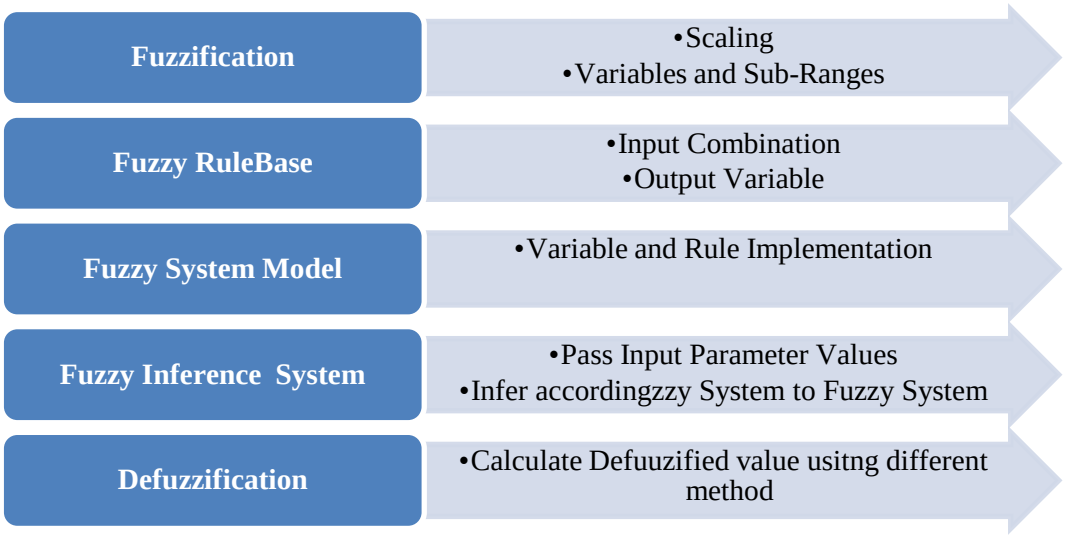
Fuzzy logic involves five main steps as explained below

- Fuzzification
- RuleBase
- Fuzzy Model for Rule Base
- Fuzzy Inference System
- Defuzzification

Fuzzification is process of creating fuzzy variables with numeric sub-ranges values in between continuous values of with binary or Boolean logic. Then rule base is created for different vague combinations or possibilities of mixture of input sub-ranges. Fuzzy System Model is designed for combinations of multiple vague inputs to generate particular output class sub-range by matching pattern of combination as

depicted in Figure No.3. Input is passed to Fuzzy inference system which evaluates to FIS value of Membership and it is defuzzied using different methods of defuzzification.

Figure 3 Fuzzy Logic



Source: Compiled by Researcher

III. Fuzzy Based COVID-19 Classification and Pattern Analysis

By applying Fuzzy Logic, Multi-criteria or multiple parameters are considered for decision making, so as fuzzy approach considers multiple inputs with different levels of value for precision of sub ranges integrated with relevant studied parameters can give better result than old methods.

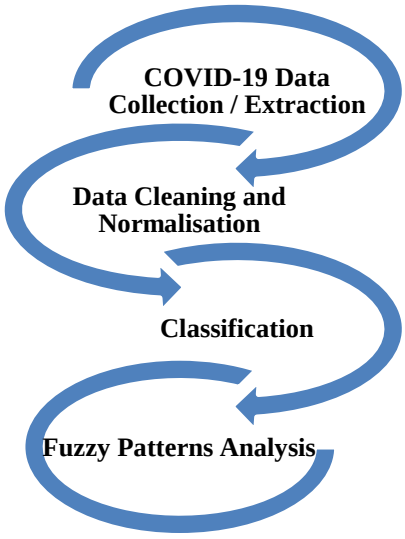
Fuzzy takes different input parameter and then identical to input for fitness. As data is rising every day in form of digital records and communication channels, the value mining process to trace out new discoveries from enormous data for better decision making is need of time for hospitals, medical stakeholders and so tempting magnetism for scholars so Medical Data Analysis is need of time now days to overcome the uncertainty of the pandemic.

This process involves phase-wise extraction, normalization, classification and analysis as depicted in figure number 4.

Researchers (2021) [39] with the context of Fuzzy Rule based systems and Deep Neural Network hybrid approach shows significant feature ranking and features connectivity graph.

Different Soft Computing technique as Fuzzy Logic, Neural Network and Genetic Algorithm [3], every technique has its own unique strength and its limitation also. So combined strategies can help with such limitations.

Figure 4: COVID-19 Data Analysis Process



Source: Compiled by Researcher

IV. R Tool

R Tools with R programming framework is used in design and implementation experiments with different software components

Table 3 RStudio Team R packages

Sr. No.	Package	Used For
1	Tidyverse	data science, including ggplot2, dplyr, tidyr, and purr
2	Shiny	web technology
3	RMarkdown	Insert R code into markdown documents
4	knitr	Dynamic reports combining R, TeX, Markdown & HTML
5	Packrat	Package dependency tool
6	devtools	Package development tool

Source: Compiled by Researcher

Table 4 R Packages for data processing and fuzzy logic

Sr. No.	Package	Used For
1	Dplyr	Data manipulation, Transformation and summarizes data frames
2	Readr	Quick and easy data retrieval like .csv, .tsv data
3	Tm	Provides framework and methods for Text mining
4	ggplot2	Mapping and Visualization of data and variables using Graphics
5	Tidyverse	Philosophy, grammar, data structure and Opinionated collection of R for Data Science
6	Set	Data structures and basic operations for ordinary sets, generalizations such as fuzzy sets, multisets, and fuzzy multisets, customizable sets, and intervals.

Source: Compiled by Researcher

Table 5 R Packages for Statistical Analysis Used

No.	kage	d For
	rics	d for statistical analysis
	MTools	l for statistical analysis

Source: Compiled by Researcher

As researcher planned to work on introduced social problem, researcher worked on COVID-19 data to gain the classes, predictions and dominant data glimpse using fuzzy approach and principal component analysis, the researcher studied concept of COVID-19 outbreak and its data for required input parameters and how to apply classification methods. Further fuzzy logic, Rulebase, fuzzy inference system and defuzzification methods are studied to learn fuzzy control process. And R programming and R packages learnt for experimentation and implementation are studied.

Research Motivation:

After experiencing situation of COVID-19, one can judge the future impact of such viral infections. Main challenge of Covid-19 is that the symptoms are changing from patient to patient as days passes so it's very difficult to set line of treatments. And also new trends are found in European countries.

So, there should be an ICT based model that will take patient data as an Input and after processing that data it should generate some results which could be useful for the doctors for deciding future line of treatment for patient based on the patients' health condition or severity of Infection.

So, The Researcher has identified some technical challenges for analysis of data through preliminary investigation and observation. To explore unfamiliar, exploration of new patterns and to scrutinize exceptions, incidents, cases by using strategic, structured and organized value mining to reach at sound decision is motivating potency for research in COVID 19.

To deal with vague data, fuzzy logic has played vital role in different domains. So, by considering the need of COVID research and treatment different analysis tools and algorithms for COVID can be focused to design in an action on existing data.

V. Literature Review:

To understand the concept in depth, literature review is performed. Further the reviews are classified on the basis of its application and thus the following six categories were obtained.

V.1 COVID 19

V.2 Measures of Interestingness

V.3 Fuzzy Based Analysis Algorithm COVID 19 treatment

V.4 Analysis using R

V.1 COVID-19

Authors have discussed about Chinese experience of dealing with COVID-19 cases for COVID-19 prevention and Management [R-6]. Enforcement of drastic quarantine measures were proactive measures to control the infection rate and delays resulted in death tolls. Researchers [R-1] explored the impact of vital signs, chronic disease, preliminary clinical data, and demographic features to predict mortality and survival. Grid search technique was applied for parameter optimization. Researchers suggested model needs improvement for multiple datasets. Authors suggested that, both population-wide and custom-made approaches should be developed as infectious diseases have diverse infectivity, incubation time, transmission mode, and will lead to different

signs dependent on the host [R-11].

For future work authors suggested [R-22], if data sets can be gathered by researchers or WHO organization or based on personal efforts to include symptoms and other information of suspects of COVID-19, researcher can diagnose that new Coronavirus. Secondly, this classification can be further improved if other symptoms like loss of taste and smell, a rash on the skin, etc., are considered.

V.2 Measures of Interestingness

This experiment [R-26] shows application of measures of interestingness and association rules for cluster analysis. Authors explained that, pruning of rules needs proper rules ranking, selection and association analysis for identifying representative rules as it is very complex process and can affect result. Author suggests pruning and efficiency classification for future work.

Author has studied [R-15] different 53 objective measures for associative rule classification. Number of rules may lead in repetitive result conditions so reducing rules can lead in improved result. The observation of experiment shows that no standard single measure can be identified for classification and few set of measures have same impact and effectiveness, so one representative measure from resultant clustering of measures improves results. Author indicates that relevant feature or data set may direct right measure to apply.

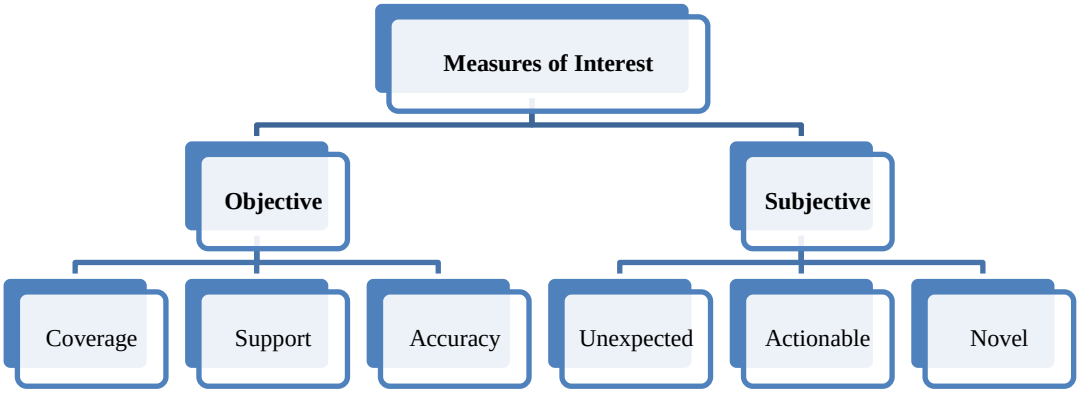
In this paper[R-6] authors have proposed a methodology for evaluating interestingness measures for closed itemset selection. The study gives comparative analysis of different measures of interestingness and evaluates to selection of suitable measure with ranking and classification techniques. Authors have mainly studied 3 measures as leverage, stability and difference measure. The study was performed to examine up to what extent measure evaluation matches to expert's evaluation for different datasets. Finally, author concluded with stability as better measure over other two leverage and difference.

According to author [R-37], there are different measures of interestingness for data mining real world applications, so part of author's work also focuses effectiveness of measures on real world application with web usage mining based application for information retrieval and management over internet, intranet and extranet. Author also analysed effectiveness of measures of interestingness with association rules, relationships among measures and rules and pruning of rules.

Ken defined about Pattern as: [R-24]

"Given a set of facts (data) F, a Language L, and some measure of certainty C, a pattern S is a statement S in L that describes relationships among a subset Fs of F with certainty C, such that S is simpler (in some sense) than the enumeration of all facts in Fs"

Figure 5: Taxonomy of Measures of Interest



Source: Compiled by Ken Mcgarry (2005) [R-23]

"The interestingness (I) of a rule is defined by its support (S), confidence (C) and knowledge of common sense rules (κ) and is given by I = f(S, C, κ). The relative entropy measure given by D(p(x) || q(x)) computes the differences in probability distributions when q(x) is p(x)."

According to author separation of subjective and objective measures can overcome knowledge acquisition problems with different aspects of evaluation. And if prior idea about patterns association is there then the prior knowledge about event can lead in better judgments and evaluation measures. Goal driven subjective measures approach only focuses to set assumption and ignores to unset and unforeseen patterns and relations. So the statistical measure can be combined with correlation measure to get improvement.

V.3 Fuzzy Based Analysis Algorithm COVID 19 treatment

Authors revealed that [R-19] several studies are being conducted to overcome this lethal disease caused by Coronaviruses. These clinical specimens were subjected to genetic, serological, and biochemical testing. For further processing, several characteristics are collected from clinical samples, and also the CT-lung screening speeds up the process and aids in identifying people who have been infected with the coronavirus. Based on this, the COVID test is performed, the genuine positive, negative, and false positive, negative are separated, and the accuracy, precision, recall,

and F1-factor are calculated.

Authors explained that, FL is a viable classifier that can be used for COVID-19 diagnosis because of the following reasons; (i) FL is powerful, simple, flexible, fast, and appropriate to the real world applications, (ii) FL can handle problems with imprecise and incomplete data, hence, it can make accurate predictions even with small amount of training data, (iii) FL is less sensitive to missing data, it is also resistive resistance to noisy data which avoids over-fitting the dataset, and (iv) when new data or rules are added to the system, there is no need to re-train the system, mainly just adding new rules (besides rule conflict check). Authors also suggested future work; the element of uncertainty is one of the biggest challenges in this field. Studying the input uncertainty simulation with input distributions that are based on real-world data.

[33] Fuzzy classification applied by authors and identified more refined multi-valued classes.

V.4 Analysis using R

The impetration of COVID-19 data research using R mentioned limitation of large corpus volume, which can be handled by using machine learning methods available in R[R-21].

V.5 Literature Summary:

The Literature review presented in this chapter embodies mainly the framework of COVID-19 Data Analysis and application of soft computing techniques for pre-processing, standardisation process and methods that is to bring data to normal or regular form data contents such as symptoms, O2 level, age, lab results, severity levels and to analyse this data from the text obtained to support medical stakeholders with computerised decision-making abilities. COVID-19 Impact analysis process is studied with different categories of COVID-19 Impact as depicted in table no. 2.2. Different advanced techniques like machine learning based, cluster based and fuzzy based different approaches are discussed exploring strength and weakness of each method for classification, prediction and grouping study. So the researcher carved out that pattern detection, learning based prediction, hybrid or/and optimised techniques integrated with statistical methods can lead in improved COVID-19 Impact analysis approach.

VI. RESEARCH GAP

After going through the preliminary research investigation, it seems as problem is very recent, very less work is addressed so more reliable and innovative Fuzzy based classification model can be developed to address the problem with immediate technical classification and prioritization solutions. Existing Research shows, two possible classes infected and non-infected, but multiple classes can be formed to plan precise patterns of treatments. Severity level and recurrence of COVID-19 infection is needed to be addressed as parameters or determinants of infection. Amongst so many parameters dominant parameters need to be traced to focus separately.

VII. Statement of the Problem

COVID-19 pandemic exposed human limitations in global world and medical resources incapability to serve in uncertainty like disastrous condition, to address the problem researcher proposed to study and provide improvement in technological solutions in form of classification and prioritization of COVID-19 treatment based on patients data by applying fuzzy approach to respiratory viral disease and focused to identify constraints of decisions in form of input parameter, fuzzy patterns, associations (relevance analysis).

VIII. Title of the research work

The researcher aimed to address the algorithmic approaches to COVID-19 data analysis by proposing study titled,

"Fuzzy based classification and Impact analysis of Respiratory Viral Infections diseases"

After detailed review and study of the related literature, the researcher traced out the framework and research gap about computational techniques lacking in Treatment and Decision making especially for Respiratory Viral infections, The Researcher framed title and objectives accordingly and designed, applied fuzzy approaches and planned Machine Learning based algorithm on COVID-19 data with three parallel approaches.

IX. Objectives of Research Work

The main purpose of current study is to apply fuzzy based pattern of COVID-19 media data using R tool that would help the expert, professionals, governments as well as students in the process of medical treatment and care.

The objectives of present research were to collect the COVID-19 patients' treatment and pre and Post COVID data and parameters required for the study, accordingly to collect from Bharati Hospital, Sangli and to apply fuzzy rule-based system to COVID-19 Data for right classification, prediction and grouping and supremacy analysis. For the said research work the researcher has set following objectives:

For the supposed research study the researcher has established the following objectives:

To identify pattern and association of covid-19 constraints and its impact analysis to support medical, pharmaceuticals, diagnostic and laboratory related decision making and treatment care.

In this research study various approaches of machine learning algorithm implemented. The results are tested for Bharati Hospital, COVID Section, and Sangli Data. The current study

is based on literature analysis depends on the resources listed in the references. In addition, this is significant mix of empirical and analytical study followed by using design and creation research strategy. Primary part of input parameter relevance analysis and identification performed after analysing literature and analysing dataset in light of relevance analysis. The entire machine learning application were carried out using R-tool, R-Studio for r programming, Google Forms and Excel were used for dataset generation.

X. Research Strategy

After Parameter identification part, the researcher followed design and Creation Research Strategy. Initially researcher studied the Parameter Identification and relevance analysis through Google Forms and Excel Visualization techniques and the strategy spotlights on application of machine learning algorithm for COVID-19 data representations to provide significant treatment and decision making model for respiratory infections diseases.

By setting the fuzzy weights to extract identical features and comparison for parameter passed as input. If sample data is collected, filtered and classified for datasets for analytics range partitions can help to form dataset in meaningful manner and then for every data set parameter membership for fuzzy specification can be evaluated to apply membership based rule to predict it's category of Covid Test Result is (positive, negative) more specifically. By applying machine learning approaches, fuzzy static and dynamic grouping mechanism and its different evolutionary combinations algorithm to COVID-19 datasets, stakeholders can identify and derive related test results, can identify similar and dominant groups to get insights of data.

XI. Rationale and Significance of the Study

The proposed study will be significant for machine learning and fuzzy based algorithm for impact classification and pattern detection to support COVID 19 treatment decision making.

- This study primarily determines various possible input parameters influencing COVID 19 treatment and care.
- The classification and patterns will help medical researchers and experts to identify tracks for treatment, priorities and segregation of patients. The study classifies impact into Serious, Average and deferred cases so accordingly research can focus to class specific symptoms and accordingly treatment can be planned.
- This will help for hospital and pharmaceutical resources management and similar situations.
- As incident based data captured will be transformed to contextual information will assist to design strategic policies with additional importance in all fields of COVID-19 research and impact.

XII. Scope of the Study

- Irrespective of varied usefulness of COVID-19 patient's data, researcher has planned to application of machine learning and related blended technological solutions to COVID-19 data for selected input parameters only.
- Initially the Researcher studied data for identified parameters using Google Form and Excel Analysis to understand relevance and importance of parameters in data collected from Bharati Hospital, Sangli.
- In form of secondary data analysis, researcher explored available research articles, news, publications to related context.
- COVID-19 have drawn attention of all different researchers, so investigator planned to address technological solution application of machine learning to respiratory diseases by applying machine learning to COVID-19 Data as the disease demands solutions in all aspects.
- The established tactic streamlines the challenges of data pre-processing correlation analysis and Fuzzy based approaches for classification, Prioritization in COVID-19 data analysis.
- The current research work gives simplification and improvement in processing logic and speed up Data analysis to support all different medical, pharmaceutical, pathology laboratory and diagnostic field stakeholders.

XIII. Materials and Sample Size

Primary Data:

Bharati Hospital, COVID Section, Sangli

Data Collected for the Period of Patients Admitted - July 2020 and August 2020 records.

(First Wave in Sangli District) through Google survey Data Entry.

For relevance analysis investigator also performed google survey data collection from 25 direct respondents observed randomly.

Secondary Data:

News, articles, Research articles, from Media like websites, News Papers, TV and other publications and media resources

Sample Size

Data Collected for the Period of Patients Admitted - July 2020 and August 2020 records.

(First Wave in Sangli District).

Table 5 Data Analysed by Researcher

proach	**a Set**	**e**
a Pre-processing	iddt.csv	Data records
		records ult stored to newt.csv
		lained pre-processing for 200 records, separation of symptoms and count of symptoms restored to dataset
zy Classification	rtable.csv	-processed, selected columns to newt dataset

Source: Compiled by Researcher

XIV. Software Used:

Software used for programming and implementation in windows-7 environment are as shown in Table No. 6

Table 6 Software Used

Sr. No.	**Software Used**	**Version**
1	R Tools Framework	3.5.2
2	R Studio	1.1.463
3	Google Forms	Version 103.0.5060.66 (Official Build) (32-bit)
4	Microsoft Excel	Office 2007

Source: Compiled by Researcher

XV. Fuzzy Based Algorithm:

Fuzzy approaches for COVID-19 patient's data to analyse it for improvement in treatment and care using option based fuzzy form to deal with different needs of medical stakeholders are recognized and medical stakeholders/experts/doctors can determine to any or multiple decision approaches as specified in Table no. 7:

Table 7 Algorithms approach

Sr. No.	**Algorithm**	**Approach**
1	Fuzzy Classification	To get Fuzzy patterns of Output class
		And to get certain outcome membership frame for impact analysis.

Source: Compiled by Researcher

Algorithm Fuzzy based COVID-19 Patients Data Classification Analysis:

Step 1: Collect data from Hospital records or using Patient Survey method

Step 2: Create Dataset, Apply Data Preprocessing created dataset, to deal with noisy data.

Step 3 Apply fuzzy part I static classification for identified parameters

Apply Rule based fuzzy logic to patterns as.

3.0 Verify identified parameters significance in fuzzy approach consideration by using correlation analysis to parameters dataset

3.1 Apply Fuzzification- Generate Fuzzy System with Variables, RuleBase with pre-defined Fuzzy sub ranges (Static Fuzzy Partitions) and Get Fuzzy Model for Membership Fitness to infer Output Variable.

3.2 Input: Extract parameters from pre-processed dataset in required Data frame.

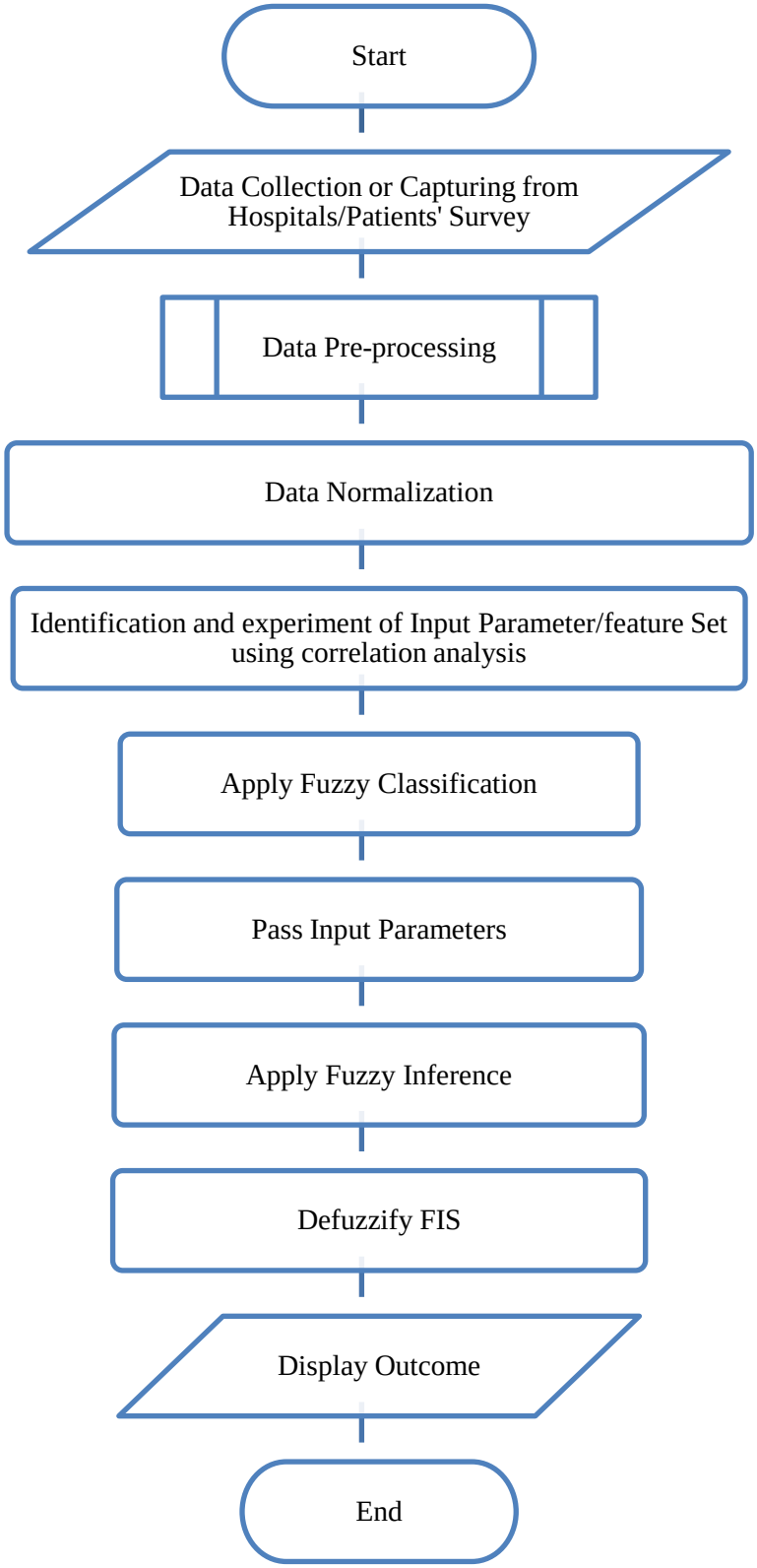
3.3Process: Apply fuzzy inference rule according to comparison of extracted member category. Implement fuzzy model to apply fuzzification with generated Rulebase to infer the data for fitness function.

3.4 Output: Achieve De-fuzzification on Fuzzy Inference System (FIS) values produced using model.

Step 4 Stop.

XVI. Flowchart:

Figure 6 Flowchart for fuzzy Based COVID-19 Patient's Analysis



XVII. Data Collection

Data is collected from Bharati Hospital COVID Section, Sangli using Google Forms for required input parameters and further Data is extracted in .csv format using Google Sheets.

The Researcher has experimented the two ways of data collection as

1) By Using Hospital's Data.

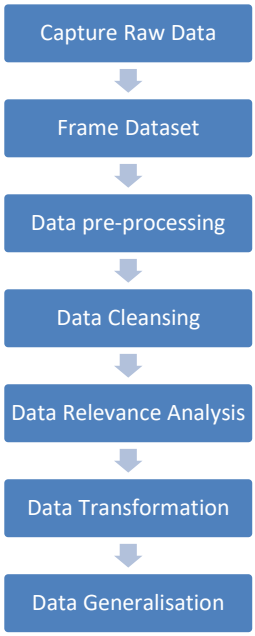
2) By using patient's Survey Method.

XVIII. Data Pre-processing

To handle missing data, or different forms of data and incomplete, inconsistent data, it needs data pre-processing which will avoid misleading outcomes. So, to bring data consistency applying pre-processing (as depicted in Figure No.7) to dataset can help.

By using text mining package tm symptoms concatenation is separated and counts of symptoms are represented to bring it to numeric form.

Figure 7 Data Pre-processing



Source: Compiled by Researcher

XVIII.1 Data Cleansing

Data cleansing or data cleaning involves removing the noise and treatment of missing values. The noise is removed by applying smoothing techniques and the problem of missing values is solved by replacing a missing value with most commonly occurring value for that attribute.

Data cleansing is the preliminary process for maximisation of accuracy in data analysis where in incorrect and invalid data is dealt logically. It involves deleting unfilled fields, duplicate fields, missing information systematically. This process is conducted to get clean and standard data for further analysis.

XVIII.2 Relevance Analysis (Feature Selection)

Relevance analysis is way of establishing and identifying interesting relationships among data. Sometimes datasets may have irrelevant data also, so isolating of relevant data from irrelevant data is necessary.

XVIII.3 Data Transformation: The data can be transformed by any of the following methods.

(1) Normalization/Regularisation: The data is transformed using normalization. Normalization involves scaling all values for given attribute in order to make them fall within a small specified range.

Normalization is used when in the learning step, the neural networks or the methods involving measurements are used.

(2) Generalization: Generalize the data to higher level concepts using concept hierarchies and or normalize data which involves scaling the values.

In this article, Research methodology is discussed along with plan of work, algorithm and flowchart is developed as general fuzzy framework for COVID-19 patient's data analysis of treatment and care strategic and decision making improvement. This chapter also explains title and objectives of research Model of data is represented. It also gives rational and significance, scope and sample.

Dependency and Vagueness among Parameters:

COVID-19 data as now a days being generated at different levels has different challenges of roughness and imperfections. To reduce them, researcher has emphasized on relative analysis of changes observed. Coefficient of co-relation is also applied to plug and get essential and independent set of input parameters. Symptoms, Oxygen level, CT Score, Diagnosis time all these parameters are observed uncertain and major determinants for treatment and care. So these parameters uncertainty is handled using fuzzy approach to get guesstimate.

Infected patients also have different levels of severity and needs different ways to treat and care. An occurrence showing mixed different values for these input reflects different gist so can't be exacted using binary classification as only infected and non-infected. Such classification causes wrong result and treatment and care decision which affects consequent strategies. Exact two way classification result possibly will damage the medical decision routine. The current discussion indicates the ambiguity in COVID-19, so machine learning based fuzzy approach will be applied.

XIX. Fuzzy Approach For Reducing Uncertainty Parameters:

This paper considers and predicts context of relative analysis with reference to graphical analysis. In the proposed research, the researcher has planned to focus on other dimensions or input parameters analysis by applying coefficient of correlation to identify independent and important multiple input parameters and further to classify with multivalve logic or in the form approximate degree of ranges.

Further the researcher would like to apply classification of COVID-19 patient's data for quantifying into the scale of impact as shown in table no.7.

Table 7 Fuzzy Approach-Degrees of Intervals

Sr. No.	COVID-19 IMPACT	Probability Range
1	Serious	0.6 to 0.8
2	Severe	0.4 to 0.6
3	Negative	0.2 to 0.4
4	Normal	0 to 0.2

Source: Compiled by Researcher

So based on resultant data researcher can evaluate similarity or patterns of degree membership to classify and predict COVID-19 impact category. The researcher has intended the fuzzy intervals partitions based on the representation in table no. 7.

Parameters Data Pre-Processing

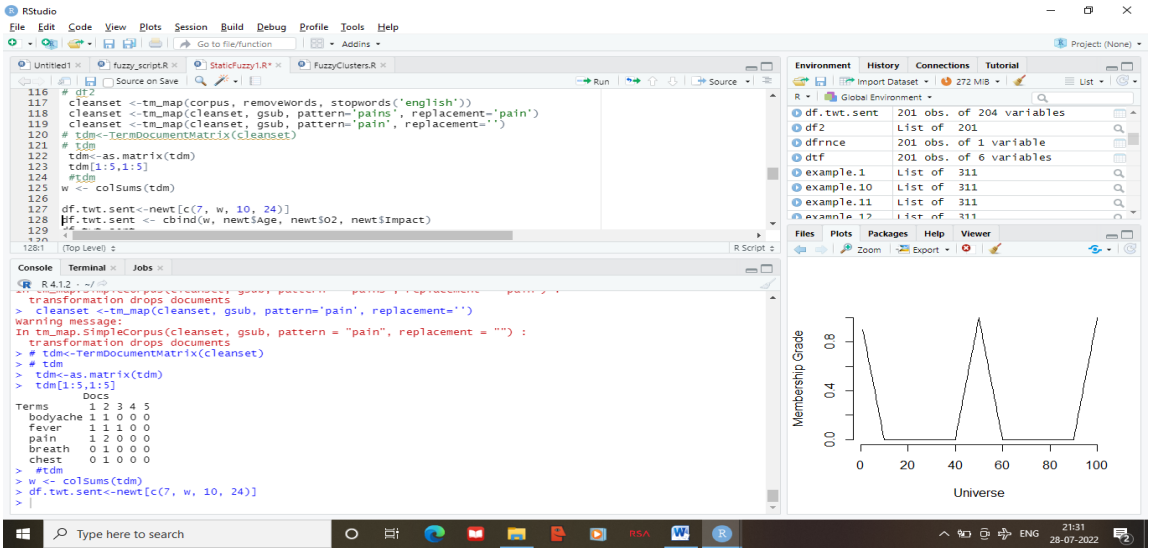


Figure 8 Data Pre-processing and Term Document Matrix

All Symptoms records collected are treated as corpus text by using text mining symptoms terms entered are separated and count of symptoms observed is determined with term document matrix which represents numeric value for symptoms to pass as key input parameter to study the impact pattern.

All other parameters NA or blank values are replaced with value zero.

XX. Conclusion

The potential benefits of a medical strategy and solution development in Sangli district are highlighted in this study

which is summarized as follows: Pattern in Symptoms occurred and infection level, Pre and Post COVID problems, major Impact of COVID-19 infection on body organs mechanism for improving medical strategies. A detail mapping of strategies will help in the future treatment and post COVID care of the patient and will support the medical treatment decisions; technically researcher has planned to design a fuzzy based model for COVID-19 patient treatment and decision making to address the problem, this stood the main purpose behind selection of the proposed research area.

The study will help all the stake holders in medical sector to craft management and decision making strategies for optimised treatment model. It can also help to create awareness amongst patients and common people about COVID impact. It will help to medical stakeholders and researchers in the area to identify major parameters of COVID-19 impact observed in patients. In the research work, the investigator has investigated and formed strategy for COVID-19 Data Normalization, fuzzy classification for pattern and certainty analysis used; graphical relative analysis, Coefficient of relationship which will help in analysis in form relevance analysis, classification, prioritization and grouping to support medical stakeholder's especially hospitals, pharmaceuticals and laboratories to manage their resources and in decision making in treatment and care for viral diseases.

Implementation at Micro level units for day to day or routine healthcare setup design.-

Diverse health settings in large population countries like India and China is also challenge to provide medical treatments and facilities, create awareness in different kind of society, so micro level units and hospitals, clinics or individual practitioners if would implement these automated solution at their service points or clinics, this will reduce the load on specialty hospitals and government hospitals.

Hospital and Pharmaceuticals Management-

As the proposed work focuses clinical decision making support it will provide management solutions and techno-medical- management strategic models can be designed in future work for Hospital Management.

XXI. Theoretical Contribution:

The investigator's key contribution is on studying COVID-19 infection related Fuzzy classification algorithm combined with data pre-processing for the classification, prioritization and certainty analysis.

XXII. Future Scope for Research:

In future research, deep learning can be applied to COVID-19 data set. Researcher would like to suggest developing AI based model for unstructured data also. The Researcher has planned to implement machine learning and evolutionary model for COVID-19 patient's data with different approaches as explained in table no. 8.

Table 8 Future Approach of implantation

Sr. No.	Algorithm	Approach
1	Fuzzy Classification	To get Fuzzy patterns of Output class.
2	Fuzzy-Genetic-Machine Learning	To predict values of RT-PCR testRersult for determined set of input parameters.
3	Fuzzy Clustering Means	To identify group correspondence based on data centres to form clusters.
4	Principal Component analysis	To analyse dominant components of COVID-19 patients data by applying principal Component analysis.

As the pandemic and related data is very recent, respiratory disease and solutions are attractions of recent research, so can be applied with different expert and recommendation systems. The emergent investigation and ultimatum in research scholar positions for COVID is also screening the interest of scientific community. Investigator would like to propose enhancement in the work in collaboration with medical research fellows to focus on innovative methods for computerised medical support systems. This work can be further enhanced fuzzification of images and video identifying as new parameters and by simplifying it to standard form and allocating weights for more exact set of perditions. One can also apply image processing to HRCT scan image reports to accept it as major, vital input constraint.

References

[1]. Aljameel Sunayh, Khan Irfan, Aslam Nida, Aljabri Malak, Alsulmi , "Machine Learning Based Model to predict the disease Severity and Outcome in COVID 19 Patients, "Scientific Programming, Vol 2021

[2] Assaf, D., Gutman, Y., Neuman, Y. et al. Utilization of machine-learning models to accurately predict the risk for critical COVID-19. Intern Emerg Med 15, 1435–1443 (2020). https://doi.org/10.1007/s11739-020-02475-0

[3]. Aydin N., Yurdakul G. Assessing countries' performances against COVID-19 via WSIDEA and machine learning algorithms. Appl. Soft Comput. 2020; 97:1–18. Doi: 10.1016/j.asoc.2020.10679. Elsevier.

[3]Cascella, M., Rajnik, M., Aleem, A., Dulebohn, S. C., & Di Napoli, R. (2022). Features, evaluation, and treatment of coronavirus (COVID-19). Statpearls [internet].

[4]. da Silva Neto SR, Tabosa Oliveira T, Teixeira IV, Aguiar de Oliveira SB, Souza Sampaio V, Lynn T, et al. (2022) Machine learning and deep learning techniques to support clinical diagnosis of arboviral diseases: A systematic review. PLoS Negl Trop Dis 16(1): e0010061.

[5]. Debnath, S., Barnaby, D.P., Coppa, K. et al. "Machine learning to assist clinical decision-making during the COVID-19 pandemic". Bioelectron Med 6, 14 (2020). https://doi.org/10.1186/s42234-020-00050-8

[6] Fujun Peng, Lei Tu, Yongshi Yang, Peng Hu, Runsheng Wang, Qinyong Hu, Feng Cao, Taijiao Jiang, Jinlyu Sun, Guogang Xu, Christopher Chang, "Management and Treatment of COVID-19: The Chinese Experience", Canadian Journal of Cardiology, Volume 36, Issue 6, 2020, Pages 915-930, ISSN 0828-282X

[7]. Khanday, A.M.U.D., Rabani, S.T., Khan, Q.R. et al. (2020) "Machine learning based approaches for detecting COVID-19 using clinical text data". Int. j. inf. tecnol. 12, 731–739 (2020). https://doi.org/10.1007/s41870-020-00495-9

[8]. LiqiangGeng, Howard Hamilton (September 2006), "Interestingness Measures for Data Mining: A Survey", ACM Computing Surveys, Vol. 38, No. 3, Article 9.

[9]. Mahmoudi , Mohmad and Heydari, Mohmmad Hpssein & Qasem, Sultan & Mosavi, Amir. ," Principal Component Analysis to study the relations between the spread rates of COVID 19 in High Risks Countries.", Alexandra Engineering Journal 2020

[10]. Mahdavi M, Choubdar H, Zabeh E, Rieder M, Safavi-Naeini S, Jobbagy Z, et al. (2021) A machine learning based exploration of COVID-19 mortality risk. PLoS ONE 16(7): e0252384. https://doi.org/10.1371/journal.pone.0252384

[11]. Marie Gousse, Pauline Penot, […] COCOREC study group." Clinical recurrences of COVID-19 symptoms after recovery: Viral relapse, reinfection or inflammatory rebound?" The Journal of Infection. Nov 2020, 81[5]

[12]. Mohammad-H. Tayarani N., Applications of artificial intelligence in battling against covid-19: A literature review, Chaos, Solitons & Fractals, Volume 142, 2021, 110338, ISSN 0960-0779, https://doi.org/10.1016/j.chaos.2020.110338.(https://www.sciencedirect.com/science/article/pii/S0960077920307335)

[13]. Mohammad Sadeq Mottaqi, Fatemeh Mohammadipanah, HediehSajedi (2021), "Contribution of machine learning approaches in response to SARS-CoV-2 infection,Informatics in Medicine Unlocked",Volume 23,2021,100526,ISSN 2352-9148, https://doi.org/10.1016/j.imu.2021.100526

[14]. Mojjada, R. K., Yadav, A., Prabhu, A. V., & Natarajan, Y. (2020). Machine Learning Models for covid-19 future forecasting. Materials today. Proceedings, 10.1016/j.matpr.2020.10.962. Advance online publication. https://doi.org/10.1016/j.matpr.2020.10.962

[15]. MojdehJalali-Heravi, Osmar R. Zaïane(2010), "A Study on Interestingness Measures for Associative Classifiers", University of Alberta, Canada.

[16]. Muhammad, L.J., Algehyne, E.A., Usman, S.S. et al.(2021) "Supervised Machine Learning Models for Prediction of COVID-19 Infection using Epidemiology Dataset". SN COMPUT. SCI. 2, 11 (2021). https://doi.org/10.1007/s42979-020-00394-7

[17]. N. Peiffer-Smadja, T.M. Rawson, R. Ahmad, A. Buchard, P. Georgiou, F.-X. Lescure, G. Birgand, A.H. Holmes(2020), "Machine learning for clinical decision support in infectious diseases: a narrative review of current applications,Clinical Microbiology and Infection",Volume 26, Issue 5,2020,Pages 584-595,ISSN 1198-743X, https://doi.org/10.1016/j.cmi.2019.09.009.(https://www.sciencedirect.com/science/article/pii/S1198743X1930494X

[18] Norah Alballa, Isra Al-Turaiki, "Machine learning approaches in COVID 19 diagnosis, mortality, and severity risk prediction: A review", Informatics in Medicine Unlocked, Volume 24, 2021. ISSN 2352-9148

[19]. Osaka University. "Sorting out viruses with machine learning." ScienceDaily, 11 November 2020.https://www.sciencedaily.com/releases/2020/11/201111102641.htm

[20]. Osama R. Shahin, Hamoud H. Alshammari, Ahmed I. Taloba, Rasha M. Abd El-Aziz, Machine Learning Approach for Autonomous Detection and Classification of COVID-19 Virus, Computers and Electrical Engineering, 2022, 108055, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2022.108055. (https://www.sciencedirect.com/science/article/pii/S0045790622003159)

[21]. Park, M.S., Jo, H., Lee, H. l. . , "Machine Learning-Based COVID-19 Patients Triage Algorithm Using Patient-Generated Health Data from Nationwide Multicenter Database", Infect Dis Ther (2022).

[22] Patricia Carracedo, Rosa Puertas, Luisa Marti, "Research lines on the impact of the COVID-19 pandemic on business. A text mining analysis", Journal of Business Research, Volume 132, 2021, Pages 586-593, ISSN 0148-2963, https://doi.org/10.1016/j.jbusres.2020.11.043.

[23]. Peltan ID, Beesley SJ, Webb BJ, Lopansri BK, Sinclair W, Jacobs JR, et al. (2021), "Evaluation of potential COVID-19 recurrence in patients with late repeat positive SARS-CoV-2 testing", PLoS ONE 16(5): e0251214.

[24]. Pijush Dutta, Shobhandeb Paul, Asok Kumar, Chapter 25 - Comparative analysis of various supervised machine learning techniques for diagnosis of COVID-19, Editor(s): Suman Lata Tripathi, Valentina E. Balas, S.K. Mohapatra, Kolla Bhanu Prakash, Janmenjoy Nayak, Electronic Devices, Circuits, and Systems for Biomedical Applications, Academic Press, 2021, Pages 521-540, ISBN 9780323851725, https://doi.org/10.1016/B978-0-323-85172-5.00020-4. (https://www.sciencedirect.com/science/article/pii/B9780323851725000204)

[25]. Prof. BhushanTrivedi(2017-2018), "Artificial Intelligence-I", Module 9, Module 36 by Prof. Bhushan Trivedi, Swayam Courses 15th Aug 2017 to 15th Jan 2018.

[26]. Quiroz-Juárez MA, Torres-Gómez A, Hoyo-Ulloa I, León-Montiel RJ, U'Ren AB. Identification of high-risk COVID-19 patients using machine learning. PLoS One. 2021 Sep 20;16(9):e0257234. doi: 10.1371/journal.pone.0257234. PMID: 34543294; PMCID: PMC8452016.

[26]. S. Kanan and S. Bhaskaran(2009), "Association Rule Pruning based on Interestingness Measures with Clustering", IJCSI International Journal of Computer Science Issues, Vol. 6, No. 1.

[26]. Said Agrebi, AnisLarbi(2020), "Use of Artificial Intelligence in infectious diseases.", Artificial Intelligence in Precision Health,Academic Press,2020,Pages 415-438,ISBN 9780128171332,https://doi.org/10.1016/B978-0-12-817133-2.00018-5. (https://www.sciencedirect.com/science/article/pii/B9780128171332000185)

[27]. Shaban, W. M., Rabie, A. H., Saleh, A. I., & Abo-Elsoud, M. A. (2021). Detecting COVID-19 patients based on fuzzy inference engine and Deep Neural Network. Applied soft computing, 99, 106906. https://doi.org/10.1016/j.asoc.2020.106906

[28].Toshiki Kuno, Yuki S, Shinpei K, Mai T, Masao I., Natalia N. E., "Prediction of in-hospital mortality with machine learning for COVID-19 patients treated with steroid and remdesivir", Journal of Medical Virology, Vol 94, Issue 3

[29]. Treskova-Schwarzbach, M., Haas, L., Reda, S. et al. Pre-existing health conditions and severe COVID-19 outcomes: an umbrella review approach and meta-analysis of global evidence. BMC Med 19, 212 (2021). https://doi.org/10.1186/s12916-021-02058-6

[30]. Uday Jain, "Effect of COVID-19 on the Organs",Cureus. 2020 Aug; 12(8): e9540.

[31]. Wang, C. J., Ng, C. Y., & Brook, R. H. (2020). Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. Jama, 323(14), 1341-1342.

[32]. Wang Y., Hu M., Li Q., Zhang X.-P., Zhai G., Yao N. "Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with covid-19 in an accurate and unobtrusive manner." 2020b. arXiv:2002.05534.

[33]. Warda M. Shaban, Asmaa H. Rabie,Ahmed I. Saleh, M.A. Abo-Elsoud[2021], "Detecting COVID-19 patients based on fuzzy inference engine and Deep Neural Network", Appl Soft Computing.

[34]. Weston C. Roda, Marie B. Varughese, Donglin Han, Michael Y. Li, Why is it difficult to accurately predict the COVID-19 epidemic?,Infectious Disease Modelling, Volume 5, 2020, Pages 271-281, ISSN 2468-0427, https://doi.org/10.1016/j.idm.2020.03.001. (https://www.sciencedirect.com/science/article/pii/S2468042720300075)

[35]. W.T. Li, J. Ma, N. Shende, G. Castaneda, J. Chakladar, J.C. Tsai, et al. Using machine learning of clinical data to diagnose covid-19 medRxiv (2020)

[36]. X. Wu, H. Hui, M. Niu, L. Li, L. Wang, B. He, et al. Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study, Eur J Radiol, 128 (2020), p. 109041

[37]. Xiangji Huang(March 2007), "Comparison of interestingness measures for Web usage mining: An empirical study", International Journal of Information Technology and Decision Making 06(01):15-41 .

[38] Y. Mohamadou, A. Halidou, P.T. Kapen ,[2020] A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of covid-19. Applied Intelligence (2020), pp. 1-13

[39] Ye Wenjing, Lu Weiwei, Tang Yanping, Chen Guoxi, Li Xiaopan, Ji Chen, Hou Min, Zeng Guangwang, Lan Xing, Wang Yaling, Deng Xiaoqin, Cai Yuyang, Huang Hai, Yang Ling, "Identification of COVID-19 Clinical Phenotypes by Principal Component Analysis-Based Cluster Analysis", Frontiers in Medicine, ISSN=2296-858X, Volume=7, Year=2020, URL=https://www.frontiersin.org/articles/10.3389/fmed.2020.570614, DOI=10.3389/fmed.2020.570614

[40] Z. Alyasseri, Al-Betar, M. A., Doush, I. A., Awadallah, M. A., Abasi, A. K., Makhadmeh, S. N., Alomari, O. A., Abdulkareem, K. H., Adam, A., Damasevicius, R., Mohammed, M. A., & Zitar, R. A. (2021). Review on COVID-19 diagnosis models based on machine learning and deep learning approaches. Expert systems, e12759. Advance online publication. https://doi.org/10.1111/exsy.12759

[41] Zoabi, Y., Deri-Rozov, S. & Shomron, N. "Machine learning-based prediction of COVID-19 diagnosis based on symptoms". npj Digit. Med. 4, 3 (2021)

[42]. https://www.datarevenue.com/en-blog/machine-learning-covid-19