



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



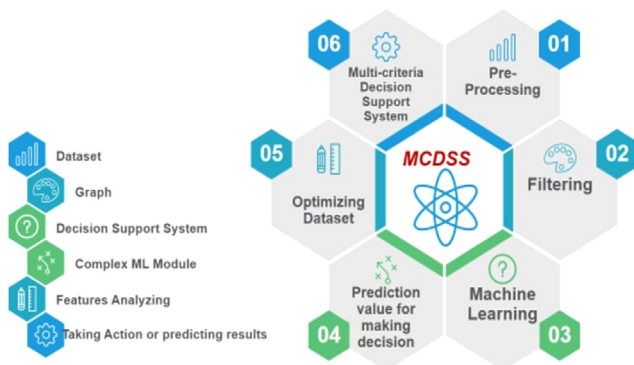
Multi-criterion Intelligent Decision Support system for COVID-19

Lakshita Aggarwal^a, Puneet Goswami^{a,*}, Shelly Sachdeva^b

^a Department of Computer Science & Engineering, SRM University Delhi-NCR Sonapat, Haryana, India

^b Department of Computer Science & Engineering, NIT Delhi, India

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 1 August 2020

Received in revised form 24 November 2020

Accepted 18 December 2020

Available online 29 December 2020

Keywords:

Multi-criterion Intelligent Decision Support system

Learning method

Epidemiology

Covid-19

Machine learning

ABSTRACT

COVID-19 is a buzz word nowadays. The deadly virus that started in China has spread worldwide. The fundamental principle is “if the disease can travel faster information has to travel even faster”. The sequence of events reveals the upheaval need to strengthen the ability of the early warning system, risk reduction, and management of national and global risks. Digital contact tracing apps like Aarogya setu (India) and Pan-European privacy preserving proximity tracing (German) has somehow helped but they are more effective in the initial stage and less relevant in the community spread phase. Thus, there is a need to devise a Decision Support System (DSS) based on machine learning algorithms. In this paper, we have attempted to propose an Additive Utility Assumption Approach for Criterion Comparison in Multi-criterion Intelligent Decision Support system for COVID-19. The dataset of Covid-19 has been taken from government link for validating the results. In this paper, an additive utility assumption-based approach for multi-criterion decision support system (MCDSS) with an accurate prediction of identified risk factors on certain well-defined input parameters is proposed and validated empirically using the standard SEIR model approach (Susceptible, Exposed, Infected and Recovered). The results include comparative analysis in tabular form with already existing approaches to illustrate the potential of the proposed approach including the parameters such as Precision, Recall and F-Score. Other advanced parameters such as, MCC (Matthews Correlation Coefficient), ROC (Receiver Operating Characteristics) and PRC (Precision Recall) have also been considered for validation and the graphs are illustrated using Jupyter notebook. The statistical analysis of the most affected top eight states of India is undertaken effectively using the Weka software tool and IBM Cognos software to correctly predict the outbreak of pandemic situation due to Covid-19. Finally, the article has immense potential to contribute to the COVID-19 situation and may prove to be instrumental in propelling the research interest of researchers and providing some useful insights for the current pandemic situation.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

The present COVID-19 (‘CO’ for Corona ‘VI’ for virus and ‘D’ for disease) approaches works on the “Best Efforts” method (i.e., tried, failed & sorry) and does not guarantee transfer of

* Corresponding author.

E-mail addresses: Lakshitaaggarwal31@gmail.com (L. Aggarwal), goswamipuneet@gmail.com (P. Goswami), sachdevashelly1@gmail.com (S. Sachdeva).

processes relevant information. The prediction accuracy of the data obtained from various authentic sources based on specific simulation tools like Weka is also not guaranteed. Additive Utility Assumption Approach for Criterion Comparison consideration in multi-criterion decision support system (DSS) is not taken into consideration so far to achieve a fruitful outcome with relevant inputs/parameters in the present scenario by most of the researchers. A Decision Support System is (DSS) is a computer-based system that combines data and decision logic as a tool for assisting a human decision maker. A decision support system does not make a decision. Instead, it facilitates the human decision-maker by analyzing the data and presenting the processed information in a form that is friendly to the decision-maker.

In today's pandemic situation of COVID-19, lots of data is being accumulated from the different sources, and the need is to analyze those data sets which pertains to useful information to us that can help in getting ubiquitous results for humankind. After the accumulation of large data sets, turning that data into useful or required information is a great deal. As in we have seen that for achieving the solution for COVID19, multiple dimensions needs to be looked upon. Various data related to the particular patient is required to be analyzed like age factor, nutritional intake, vaccination history, diabetes level, Blood pressure maintaining system, and many more.

So, we decided to prompt a solution which could look effectively into the matter and can contribute to making the best decision possible. A multi-criterion decision support system will look upon all the necessary factors and causes for a particular person and then will provide the results for the person. It will reduce human intervention or human mistakes sometimes caused and may lead to devastation. We can control such a tragedy with the help of a multi-criterion decision support system. A after all the data has been collected, we will analyze those data sets and will turn the raw data into information. Then, the information turns its phase to proper knowledge applying proper meanings to it. Now, the data is transformed into wisdom wherein appropriate tactful insights are applied to it so that it can take adequate and required decision using past experiences or past data sets and produce most appropriate decision possible taking in cause all necessary purpose and parameters so as to produce the perfect decision.

In this for analyzing, we took the sets that pertain to the eight most affected states of India Maharashtra, Tamil Nadu, Delhi, Gujarat, Uttar Pradesh, Rajasthan, West Bengal, and Madhya Pradesh. The data is retrieved from a repository accessible through the government link <https://news.google.com/covid19/map?hl=en-IN&mid=/m/03rk0&gl=IN&ceid=IN:en>. The database is openly available, showing data about different states of India and the index of people confirmed, recovered, or died.

The people in this extreme hardship are even surviving for the basic necessities of life, such as: - food, water, and shelter. Many migrant labourers and workers moved from their workplace to their native places, even bare-footed and with loads of families-striving children and aged parents due to immediate emergency throughout the country in the form of lockdown. As the disease is a highly contagious disease that spreads from a sick person to a healthy person even via a touch, a handshake, or any kind of physical contact. So, WHO issued guidelines for all on how to mitigate from this deadly virus and take preventive and corrective measures upon the successful alternatives to keep oneself safe from this fatal disease. Immediate corrective and preventive measures are necessary for this situation of a pandemic in the whole country. So, the need is to predict over Multiple Criterion Decision Making (MCDM) process. The international MCDM society defines it as "the study of methods and procedures by which

the concerns about multiple, usually conflicting, criterion can be formally incorporated into the management planning process" to fight from the disease.

Information at a point of care enables more realistic decisions providing health care with ubiquitous benefits from big data when it is structured, relevant, and accessible to all. Multi-criterion decision making has significantly beyond the conventional and formal methodologies and now also enters in the cognitive and informal processes to help humankind. In this paper, we discuss all the factors of multi-criterion approaches to decide on factors on how to flatten the growing curves of COVID-19 by multi-criterion decision making progress.

2. Literature review

The literature of the current research works related to the on-going buzz of COVID-19 is presented in this section.

Hu Yun et al. [1] worked on laboratory data analysis of COVID-19 screening 2510 patients. They collected nucleic acid and haematology data from 2510 patients for identifying COVID-19 infection for retrospective analysis. They tried to calculate the results of the influenza virus. In their paper, they decided to examine the effects of faecal matter. They proposed medical and clinical treatment by just introspecting 2510 patients, but the data could vary from one location to another. So the immunity and many other factors inside the body also differ from one region to the other. This paper does not takes into account number of factors such as: - people of different countries consuming varied nutritional food may show diverse nature against the virus and moreover they did not included the preventive measures in the paper.

Lixiang et al. [2] worked on the infectious disease modeling concept where they studied the transmission process of the coronavirus. They also worked on the backward and forward predictions before some tragedy happens. This paper uses the Gaussian distribution theory to construct a new model of coronavirus transmission from one to the other. In this paper, they tried to include necessary factors due to which the coronavirus might spread like reproduction number, incubation period, and the daily infection numbers. They also talked about the nutritive value that can help or the vaccine is taken from the infant stage. But, this model was unable to make some clinical treatment effective, which could efficiently help society.

Faizora et al. [3] stated that with coronavirus is still an unclear infectious disease, i.e., we can obtain an accurate SEIR (Susceptible, Exposed, Infectious, and Removed or Recovered) predictions only when the epidemic ends. In this, they made a corona tracker, which will identify based on the prediction model. But, other factors such as medical treatment analysis and preventive measures [4] were not discussed that may help in flattening the curve.

Palash et al. [5] worked on growth models where they worked on the Daily Infection Rate (DIR). They included that the data may vary according to the state or the area we are analyzing as the number of factors in the particular region also vary from one to another. They worked on state-wise analysis showing different DIR at different intervals of time, and varied results are cumulatively shown for each state. They also discussed some preventive measures where we can try making DIR to zero so that we can flatten the curve and may one day the war ends.

Samrat et al. [6] analyzed exploratory visual data through which they explained how early detection could help in better response in fighting the SARS virus. This paper investigated visual data where they studied three different categories: - confirmed cases, recovered cases and the number of people died in different countries. This paper also stressed upon the need of vaccines and other treatment for viral diseases since birth. But their observations lacked in areas where the need is to stress upon medical

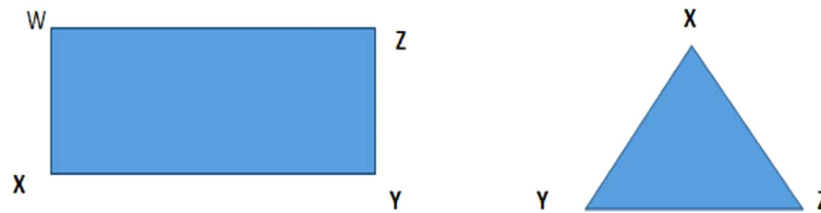


Fig. 1. Three and four length arrays.

Table 1

Comparative analysis.

S.No.	Parameters	Hu et al.	Lixiang et al.	Fairoza et al.	Palash et al.	Samrat et al.	Vinay et al.	Proposed model
1	Study Domain	Retro Analysis	Infectious disease modeling	Prediction model	Growth Models	Visually	Literture	Multi-criterion decision support system
2	Medicinal treatment	✓	×	×	×	×	✓	✓
3	Country wise Analysis	×	×	✓	✓	✓	×	✓
4	Nutritive value undertaken	×	✓	×	×	×	×	✓
5	Vaccines	×	✓	×	×	✓	✓	✓
6	Preventive Measures	×	×	×	✓	×	✓	✓

treatment or the preventive measure to be undertaken in the cause. Below, the comparative analysis clearly shows the idea.

Vinay et al. [7] have studied clinical features and transmission mechanisms of past pandemics like: - Spanish flu, Asian flu, Hong Kong flu, Swine flu. The paper also discussed various economic impacts and other risk factors on relevant statistics related to COVID-19. The authors have also taken into account signs and symptoms and patients to assess their cases carefully. In their work, they summarized their studies providing a brief overview on diagnosis, treatment and their prevention. The paper also discussed the pathway for transmission of disease of coronavirus from nasal passage till the lungs. The paper also discussed factors of AI, machine learning, & block-chain to flatten the curve (see Table 1).

So, to the best of our knowledge, this is the first attempt on COVID-19, which focusses on the multi-criterion decision support system where the data is trained from past experiences. Then it is transformed in a better way where it becomes ready to make smart decisions without erroneous human conclusions and, considering all the alternatives in one. There is an urgent need of an hour when we talk about infectious diseases, so urgency is to make some efficient ways through which less amount of human involvement is done. So, a multi-criterion decision support system is the one that will help in making the best decision considering all the alternatives available and producing the best viable choice before the situation worsens.

3. Data for decision making

A way to solve a particular set of problems that every person or group of people have when they try to figure out-“what’s the best option available-” for a problem that requires a solution. Fuzzy data is used to represent real expressions in maths that cannot be made quickly. It allows us to represents values and expressions, i.e., Wrong, Right, High, and Low. The way we use this type of values and expressions in maths is fuzzy data (see Fig. 1).

The above expressions are three or four-length arrays to every real expressions and values. A (W, X, Y, Z) & B(X, Y, Z) are the coordinates of a rectangle and triangle of three or four-length arrays to every real value and expressions. These two types of

different arrays shows the different ways we can analyze the data to understand it in great depths and calculate more effective results in different number of steps.

3.1. Decision making

In an optimal engineering design environment as such, solving the multi-criterion decision-making (MCDM) [8,9] problem is considered as a combined task of optimization and decision-making process to fight from the disease in a manner to stop spreading its effect on a large number of people.

Before, making final decision some adequate actions needs to be performed to choose the best out of so many alternatives present for solving the problem some criterion’s needs to be looked upon. This criterion in an organization comes from history, norms objectives, and first step. It all comes from data mining. So, to apply this theory, we have to set the following components to make it happen, i.e., we need:-

- Decision-makers/Deciders
- Criterion comparison
- Alternatives (possible solution to our problems)/ classify options based on criterion

There could be number of deciders in the group like: - input deciders, normal deciders, and the important deciders in the group. The action to be taken for the particular group or society as the need or severity of the disease can be different in different areas [10] on how to save the mankind from this on-going pandemic situation.

Criterion comparison of every decider must know which one is the most important to take the rest of the decision. And; finally, we need to evaluate every decision to be an alternative based on every criterion. Criterion impacts the selection of other options. Criterion [11] include the following:

- Completeness – All the types of parameters should be included in it. It acts as a closure property of mathematics.
- Redundancy – All the duplicates should have been removed at early stages only so as redundant data can be avoided efficiently.
- Operational – Each alternative should be compared against every option in decision-making criterion.

Table 2
Problem solving techniques.

Method	Advantages	Disadvantages	Areas of application
Simple Additive Weighting (SAW)	Able to compensate among criterion's and it is ready to take decisions intuitively.	Estimates do not reflect the real situations result.	Financial management, water management and business.
Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS)	Simple process is involved using, and programming and the number of steps remains the same irrespective of its attributes.	Difficult to weight and use of Euclidean distance relation judgments.	Engineering, manufacturing, business and marketing.
Analytical Hierarchy Process (AHP)	Easy to use, not much focus on data.	Problem due to interdependence between alternatives and criterion's; which may lead to inconsistencies to the final results.	Performance type problems, public policy, strategy and planning related tasks.
Simple Multi-Attribute Rating Technique (SMART)	Simple allowing for weight techniques; less effort by decision-makers.	Not a convenient procedure, not considering the framework.	Transportation and logistics, military, manufacturing and assembly problems.
Fuzzy Set Theory	Takes insufficient information, allowing for imprecise input.	Difficult to develop as it requires numerous simulations before use.	Engineering, Economics and management techniques, Biological and health care estimation of data.
Elimination at Choice Translating Reality (ELECTRE)	Considers all uncertain data vagueness into account.	Difficult to explain, causing strengths and weakness of alternatives not identified.	Energy, economics management and transportation problems.

4. Methodology used

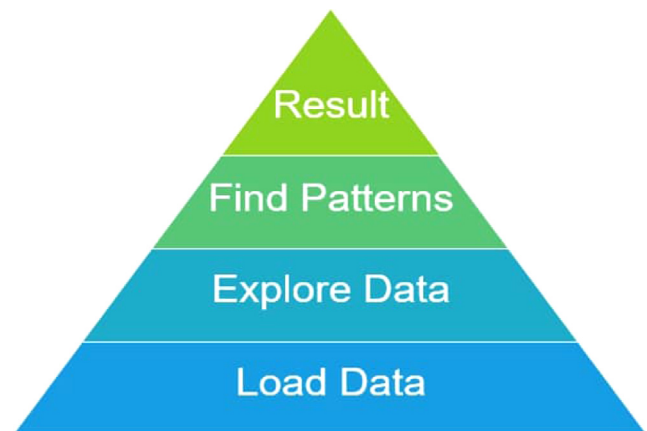
4.1. Discussions

The amount of data produced and communicated over the Internet is significantly increasing, thereby creating challenges for the organizations that would like to reap the benefits of analyzing this massive influx of big data. It is crucial to examine [12] which data sets can be converted from raw data to complete information that may be proved ubiquitous for humankind to grow fast and flatten the growing curve of cases among people.

- Loading info Establish the whole right information set to the well execute of the method. Load all the information into the data sets and implement the method well.
- Calculate results use above loaded data to calculate the results and to get the best solution from the problem.

Refer Fig. 2, in the methodology section, when there is a massive influx of raw data and we are interested in calculating the best results. So, the loading info, loads all the data in raw form or unstructured form then exploration of data takes place by implementing various methods and patterns to it. Then, that raw data is transformed from unstructured form to structured form and then the results are interpreted.

The Table 2 shows various methods with their advantages, disadvantages and their implications in different areas. Every method has its own advantages and disadvantages. Different methods are used in different fields to produce results in real simulations for taking the decision taking in cause various methods as in Simple Additive Weighting (SAW) method the main area of application is financial management where this method is used. Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS) method is generally used in areas where programming language is used to take decision so this method is majorly used in areas of engineering and marketing where diverse nature needs to be looked upon. Analytical Hierarchy Process (AHP) method is easy to use and data is not focussed much. This is entirely used in planning related tasks. Simple Multi Attribute Rating Technique (SMART) is simple technique where the decision maker's makes less effort entirely used in transportation

**Fig. 2.** The method used for calculating results.

and logistics related problems. The fuzzy set theory is used in areas where the input is the major information to be analyzed upon and fuzzy sets are basically used in the fields of electronic health related data sets. The Elimination at Choice Translating Reality (ELECTRE) considers all vagueness and uncertain data and makes it more certain and efficient to be used in energy and transportation related problems.

4.2. Steps to manage and process data

Big Data represents a new era in data exploration and utilization. While most data is stored in hard copy form, the current trend is toward the rapid digitization [13] of these massive amounts of data.

- Establish the number of deciders and alternatives.
- Establishing criterion
- Deciders choose their weights by consensus.
- Every decider differentiates every settled criterion with the others.
- Getting the consistent weights of criterion's.

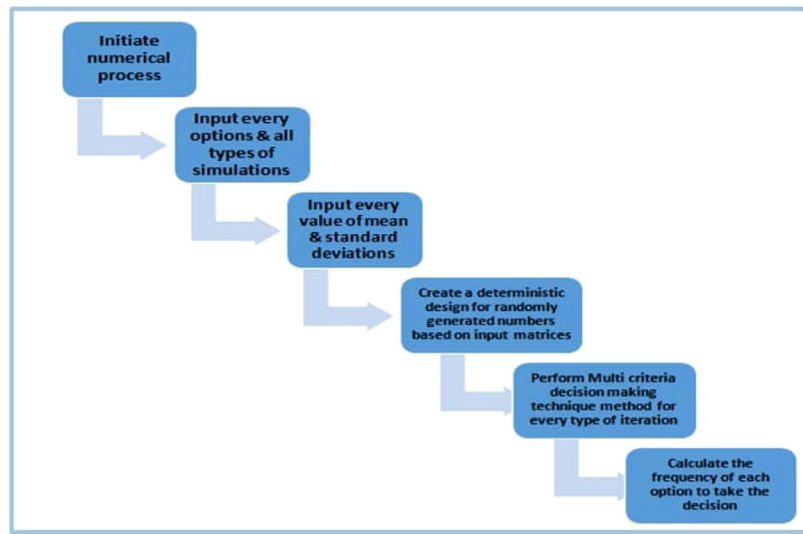


Fig. 3. Block diagram of multi-criterion decision support system.

- Every decider evaluates every alternative based on the criterion.
- Weighting of the alternatives with the consistent weights of each criterion.
- Group aggregations.
- Weighting of the resulting values on step 7 with consistent weights of deciders.
- Establishing max & min values.
- Distance between the final values with the max & min value.
- Calculate the final coefficient of the alternatives.

Firstly, start from formulating options, and decision-makers [11] must identify the decision goal, record the observations of all criterion's in a table by analysis process. The decision maker's then brainstorming, considering all the requirements to attain the best decision or the solution for the current problem. Then, it is to determine the decision alternatives, followed by the criterion. The decision set-up is entirely based upon the numerical values and also takes into account the representation of all the processes. The step before final decision to be taken depends upon the decision structuring process. Then, the final decision is taken. After; all the criterion weights or scores are computed, the decision-makers pairwise compare every alternative with respect to each of the criterion. Quantitative and qualitative measures are appropriately taken for each sets of criterion using the comparative approach and the numerical analysis of each value. Each analysis tool follows the different methods to calculate the final results of the multi-criterion decision-making approach. Many different techniques have emerged, which yields different results applied to an identical problem and helps in taking mitigate measures before the situation worsens (see Fig. 3).

4.3. Components of multi-criterion decision analysis

Multi-criterion decision analysis is a valuable asset that we apply for making complex decisions as needed in the case of COVID-19. It is mostly responsible and used in solving problems that have a choice among alternatives. It is a useful decision support tool, i.e., logical, consistent, and easy to use.

Multi-criterion decision analysis problems are composed of five components:-

- Goal
- Decision-makers
- Decision Alternative



Fig. 4. Flowchart demonstrating multi-criterion decision-making technique.

- Evaluation Criterion
- Outcomes or results

4.3.1. Algorithm for integrated approach to analyze multi-criterion spatial and non-spatial data

Start

Initialize all co-ordinates;

Input decision variables;

Calculate mean

$$\text{Mean} = \frac{1}{n} \sum_{i=0}^i x$$

Calculate standard deviation

$$\text{Variance} = \frac{1}{n-1} \sum_{i=0}^i (x - \text{Mean})^2$$

standard deviation = sqrt (variance);

Apply knowledge based tools;

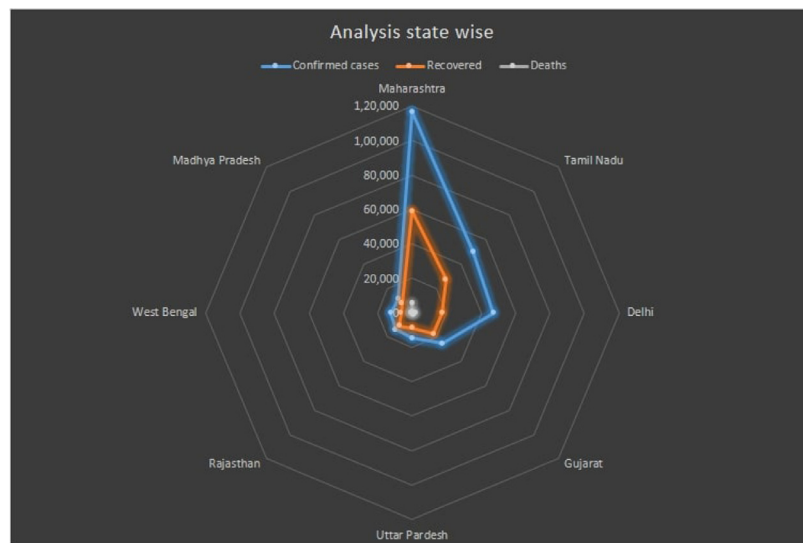


Fig. 5. Analysis state wise.

Perform MCDSS criterion using various ML techniques;
Calculate decision [14] using prediction approach;
End;

4.3.2. Flowchart demonstrating MCDSS

- To understand the approach multi-criterion system [14] uses to take best decision starts from defining the problem that includes the problem statement to be analyzed to make appropriate decision. Then, brainstorming different objectives using different techniques like questioning different set of people, using specialized tools and techniques to take decision. Set an alternatives or deciders to understand the multi-criterion decision support system (MCDSS) on provided parameters and criterions (see Fig. 4).
- After setting appropriate criterion using deciders update the filtered data on predicted data. Then, create new decision using knowledge based tools of complex Machine learning modules. Optimize the calculated datasets to take the best appropriate decision taking in cause the past approaches.
- Finally, the decision for analyzed algorithm is taken that takes into account multiple criterion for taking the best decision possible.

4.4. Factors

The factors needed to be analyzed are:-

- BCG (Bacillus Calmette Guerin) = in some countries like India BCG vaccine is a necessary vaccine given in childhood to infants, which is used for treating TB (Tuberculosis). People vaccinated with BCG [15] are proved to be least prone to the vaccine as it provides more immunity [7] to the body and helps the body to fight viral infection by making antibodies around it; if it may enter. It is observed that countries having policies of BCG vaccination [15,16] found a lower number of confirmed cases, and a lower death toll is observed. Although; data varies from country to country, and it also varies among different age groups. The vaccine founds to be ineffective in adulthood.
- Nutritional value – it is often said that citrus fruits or the food having nutrients like vitamin C, vitamin D, and magnesium ions in it proves to be ubiquitously resourceful for providing immunity against virus. So, health ministers and doctors strongly encourage people to boost immunity against

the virus so that the body prepares itself for the defense mechanism by making antibodies and fight the disease well. Moreover, results from the sensitivity analysis corroborate that people with adequate immunity are less prone to virus and are seen even to recover if caught in an infection at a mild stage only.

- Ministry of AYUSH – COVID-19 is SARS (Severe Acute Respiratory Syndrome) viral infectious disease which causes – “Corona virus”. At mild stage, ministry of AYUSH says it could be easily controlled with just home-made remedies. Medicines [17–20] made from home ingredients, natural herbs like Tulsi, Ashva gandha etc. helps in boosting immunity that helps to fight the virus. Yoga and breathing control exercises are found to be the effective measures in curing the viral disease accompanying Ayurveda, naturopathy, unani, siddha and homeopathy as well.
- Recently, Glenmark pharmaceuticals [21] on June, 20 have received approval from the drug controller general of India to manufacture and market Favipiravir, an antiviral drug to treat Covid-19 patients. Favipiravir is an antiviral drug that is approved in Japan for treating influenza and is also under the processing in many clinical trials; some trials have proved positive results while still are under the human based trial.

4.5. Results

The above discussed methodology about multi criterion decision making (MCDM) is a process of evaluating real-world situations, based on various qualitative or quantitative criterion in uncertain & risky environment to suggest an action or strategy or choosing the best option out of many available options. So, now to understand about eight worst-hit states due to COVID-19 we will be taking in cause the MCDM approach (see Fig. 5).

The data is analyzed on the Weka tool to see the number of people deceased from the total population. The data of the top 8 states in which the effects of COVID-19 is seen on a large scale, the data for the total population is taken. Then out of the total population number of confirmed cases, cases recovered, and the number of people died, statistics have been [3,22] observed to see the percentage of people affected out of the total population in the particular state.

So, Table 3 is experimental on the Weka [11] tool showing results as out of a total population in all the eight states, 0.537%

```

=== Run information ===

Evaluator:      weka.attributeSelection.CfsSubsetEval -P 1 -E 1
Search:         weka.attributeSelection.BestFirst -D 1 -N 5
Relation:       state
Instances:      8
Attributes:     5
               states
               Population
               confirmed
               recovered
               deaths
Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
  Best first..
  Start set: no attributes
  Search direction: forward
  Stale search after 5 node expansions
  Total number of subsets evaluated: 11
  Merit of best subset found: 0.537

Attribute Subset Evaluator (supervised, Class (numeric): 5 deaths):
  CFS Subset Evaluator
  Including locally predictive attributes

Selected attributes: 1,2,4 : 3
                   states
                   Population
                   recovered

```

Fig. 6. Weka analysis.

Table 3
Eight worst hit states due to COVID-19.

States	Population (Cr)	Confirmed	Recovered (k)	Deaths
Maharashtra	11.42	1.2L	59	5K
TamilNadu	6.79	50K	27	576
Delhi	1.9	47K	17	2K
Gujarat	6.27	25K	17	2K
UttarPradesh	20.42	15K	9	435
Rajasthan	6.89	14K	10	313
WestBengal	9.03	12K	7	506
MadhyaPradesh	7.33	11K	8	482

of people died, and rest are recovered even after being affected. Through the above depicted data, we conclude that India encountered least infectious mutant form because of which the recovery rate is quite high. Many people are seen to recover [23] from the virus in its mild stage only. The reason could be the nutritional values Indians consume since birth. Also, the vaccine BCG (Bacillus Calmette Guerin) is one of the factor proven to one of the recovery measure against the virus. Infants from birth to ten weeks of age are injected with BCG vaccine [3] so that the measures of high fever or TB (Tuberculosis) can be prevented (see Fig. 6).

Then, the K-means cluster analysis [24] is also performed on the data sets showing the statistical analysis, i.e., mean and standard deviation analyzed in the training table. The above analysis works on the five attributes i.e., states, population, confirmed, recovered, and deaths. Evaluation model evaluates on all training data sets. The above analysis method uses “Best first” technique for search based method. It uses forward search direction. Attribute subset evaluator evaluates each attributes individually using supervised learning model. It uses CFS subset evaluator as the feature selection algorithm. After performing, CFS pre-processing the number of steps are significantly reduced and more precise build-in classifications are obtained. Data set is divided into attributes and evaluates attributes by the degree of consistency.

After, attribute selection statistical analysis [25] takes place where mean and standard deviation on all four parameters such as: - Total population, confirmed cases, recovered cases, and the number of people died due to the pandemic (see Table 7).

Table 4
Statistical analysis.

Parameters	Mean	StandardDeviation
Total population	8.7562	5.0718
Confirmed cases	21.9	16.5248
Recovered cases	19.25	16.2692
Deaths	290.125	232.9117

Here, using Weka [26] we also analyzed mean and standard deviation of all four parameters:-

$$\text{Mean} = \frac{1}{n} \sum_{i=0}^i x \quad (1)$$

$$\text{Variance} = \frac{1}{n-1} \sum_{i=0}^i (x - \text{Mean})^2 \quad (2)$$

$$\text{Standarddeviation} = \sqrt{(\text{Variance})} \quad (3)$$

Using the above formulas, results are obtained (see Fig. 7).

Here, for analyzing clustered k-means analysis methods is chosen with 100% precision working on all eight instances of Table 4. We also calculate the correlation coefficient, which is a numerical measure of some types of correlation in the total population, confirmed cases encountered, recovered cases, and the number of people who died. Then, the mean absolute error, root mean absolute error, relative absolute error and total relative squared error is analyzed on the all right instances or the top eight states where the COVID-19 is spreading in a massive manner for public health emergency for fast-spreading and endangering a large number of health of the people, and it requires immediate actions to prevent the disease at the community level. The patterns of public sentiment on analyzing related health information and assess influenza at the political and economic spread of the virus.

We obtain clustered instances on k-means clustering method on all eight instances and the value of log likelihood suggests that the variables are discrete. This means that analysis obtained on the type of graph is discrete type variables. So, the value of log likelihood is -20.43 (approximately) showing discrete natured variables in the file (see Fig. 8).

As, we calculated statistical analysis following results are obtained on all the eight instances. This shows disease is a fast-spreading disease that is spreading at the community level, and


```

Population
mean      8.7562
std. dev. 5.0718

confirmed
mean      21.9
std. dev. 16.5248

recovered
mean      19.25
std. dev. 16.2692

deaths
mean      290.125
std. dev. 232.9117

Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      8 (100%)

Log likelihood: -20.42369

```

Fig. 7. Statistical analysis.

```

Correlation coefficient      0
Mean absolute error         215.3438
Root mean squared error     232.9117
Relative absolute error      100 %
Root relative squared error  100 %
Total Number of Instances    8

```

Fig. 8. Relative instances.

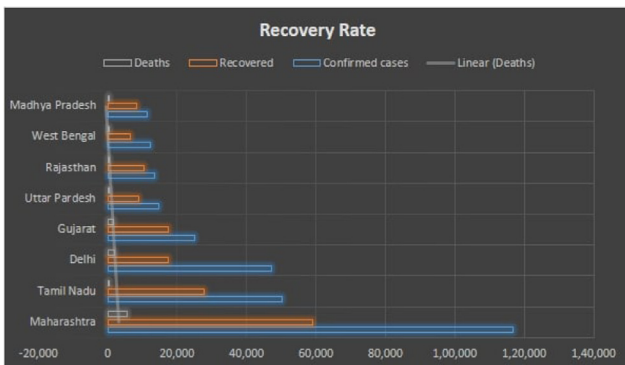


Fig. 9. Chart showing recovery rate.



Fig. 10. Analysis of all states on three different parameters.

immediate steps have to be taken to prevent the spreading [27, 28] of the disease at the community level. The disease is a health threat to humankind WHO [29] declared it a pandemic disease researchers from the different institutions are trying to make an Antidote for the disease so that large spread could be saved. The chart shows out of the total population number of confirmed cases found, cases recovered, and the number of people died linearly. We can see from the graph that linear regression in deaths has been shown, which even can be controlled if we take immediate measures for the people to solve the extreme situation (see Fig. 9).

The above table predicts the results of the fatality rate of people state wise. It takes into account the number of people died in a particular state due to pandemic till now as this number is increasing to a great extent. We can see that Maharashtra is the worst hit state still we can predict that the number of people died are very less in comparison to positive cases. So, we can say that

we can recover from the virus with improved recovery rate with appropriate prevention before situation worsens (see Fig. 10).

We concluded from the above graph that, the confirmed cases of COVID-19 are high in few states. The rate of recovery is appreciable that proves the fact that Indians have developed immunity against virus. But, still we have few cases of fatality rate generally among those people who are either aged, or in infant age group, or they are already suffering from any ailment related to breathing (see Fig. 11).

We have also calculated complete accuracy by class, including the precision values and weighted average value. [26] The TP rate i.e., the true positives (are the instances classified as a given class) that shows the value 1.000 that means all the instances are examined properly. The FP rate i.e., rate of false positives (are the instances falsely classified as a given class).

Precision is the instance that are calculated as the truly positives of a class divided by the total instance of the class.

$$\text{Precision} = \frac{(\text{TruePositive})}{(\text{TruePositive} + \text{FalsePositive})} \quad (4)$$

The Recall factor is the proportion of instance calculated as the given class divided by the TP rate in that class (or the actual total in that class).

$$\text{Recall} = \frac{(\text{TruePositive})}{(\text{TruePositive} + \text{FalseNegative})} \quad (5)$$

F- Measure is a combined factor of precision and recall.

$$F - \text{Measure} = \frac{(2 * \text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (6)$$

```

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          1.000    1.000    0.660    1.000    0.795    ?    0.500    0.660    c0
          0.000    0.000    ?        0.000    ?        ?    0.500    0.340    c1
Weighted Avg.  0.660    0.660    ?        0.660    ?        ?    0.500    0.551

=== Confusion Matrix ===

  a  b  <-- classified as
66  0  | a = c0
34  0  | b = c1

```

Fig. 11. Detailed accuracy and matrix.

Table 5

Result comparison of different learning methods [26,30,31].

Parameters	TreesJ48	Logistics	Decision table	ZeroR	Proposed paper values
Precision	0.629	0.5	0.375	0.423	0.66
Recall	0.643	0.4	0.5	0.6	1.00
F-measure	0.632	0.44	0.429	0.5	0.795

Table 6

Result comparison of different parameters.

Parameters	Casel	Casell	Caselll
MCC	0	0	0
ROC	0.5	0.5	0.5
PRC	0.6	0.3	0.551

$$i.e. F - Measure = \frac{(2 * 0.660 * 1.000)}{(0.660 + 1.000)} = 0.795 \quad (7)$$

Table 5 presents results obtained on three parameters i.e., Precision, Recall and F-measure by comparing varied Learning methods. After analyzing the values on Weka we have calculated improved results on all three parameters i.e. Precision, Recall, and F-measure (as shown in Table 5).

We also tried to analyze another three parameters i.e. MCC, ROC, and PRC.

MCC (Matthews Correlation Coefficient) is generally regarded as a balanced measure. It is even used with classes that are of very different sizes.

ROC (Receiver Operating Characteristics) area measurement gives an idea of how the classifiers are performing in the real simulation of data.

PRC (Precision-Recall) area gives more information than the ROC plot area. PRC plot [29] is generally used when we evaluate binary classifiers on variable datasets. PRC is more informative than MCC and ROC. [26]

The values of MCC (Matthews Correlation Coefficient), ROC (Receiver Operating Characteristics) and PRC (Precision-Recall) are calculated number of times in different cases i.e. I, II and III to improve the efficiency. It is estimated from the above analysis (Table 6) that in India, we have the least mutant form of infection in which if proper care and treatment happen, we can save the lives of all around us. As per the MOHFW report [1], it is observed that people having diseases like hypertension and diabetes are more prone to the virus. Thus, special attention should be given to all those suffering from the ill-effects of it.

Confusion matrix is used to describe the performance of a classification model on a set of data sets assigned to understand the true values. It also helps in improving the performance of an existing data models by evaluating the data set.

So, in the above confusion matrix [25] obtained results are interpreted as:-



Fig. 12. SEIFR model.

Table 7

Different classifiers used to analyze results.

Classifier	Figure	Attributes	Comments
TreesJ48Graph	6	Default	Uses pruning
Logistics	7	Default	Function logistics
Decision table	8	Default	Rules decision table
Zero R	11	Default	Rules zero R

a = Tested negative (It means the target variables are functional)
b = Tested positive (It means the target variables are non-functional)

'a' and 'b' are the letters assigned to class values.

So, using these models improves the efficiency of class variables to a greater extent.

We can say that few factors needed to be taken in the cause before coming to a decision; decision-makers in multi-criterion decision [14,30,31] making progress pose numerous alternatives before predicting a person may/may not possess a disease all the factors must be analyzed.

5. Epidemiology modeling of data

Epidemiology model [32] is a dynamic model that says it is a study of the pattern of the illness of the disease occurred in the human population by connecting it with the various vectors like: - air, water, environment, and individual droplet infection from one to the other. "Epidemiology" is the study "upon people". It applies to the human population that helps in identifying illness patterns with the environment, demography, etc. Now, in 21st epidemiology has become the fundamental priority to study trends or illness or the change in SARS (Severe Acute Respiratory Syndrome) pattern in human-related diseases. As COVID-19 is the seventh member of the SARS family, it has created a buzz of fear among the whole country.

When the biological and epidemiological knowledge is translated scientifically into the mathematical equations, this is known as mathematical epidemiology [1]. The purpose of mathematical epidemiology is to understand the patterns on the transmission of disease. It also takes into account how fast the disease will spread infecting number of people. Mathematical epidemiology [33–36] is also used to identify the number of risk factors of the disease.

Epidemiology started its phase from SIR taking into account the number of people who are Susceptible, Exposed, and Infected.

Then the shift took into a pace where another phase was also added, i.e., Exposed, making it Susceptible, Exposed, Infected and Recovered (SEIR).

5.1. Proposed methodology

Now, we propose the model that also takes into account the number of people those are died, making it another parameter fatality rate $F(t)$. $F(t)$ is the proportion of infected people who will die due to the pandemic. It is an important parameter to be taken in cause to understand the threshold amount of people able to tackle the effects of the virus and also to track the number of people not able to struggle with the virus. This parameter $F(t)$ will also help us to understand which level or what kind of historical background people are not immune to the virus or the special care to be taken upon such people as an immediate measure.

So, this SEIFR model will help the society to make better decisions in accurate time, which might not complicate efforts and pin down the number of people dying due to its ill effects. Through this number, we will be able to bridge the gap and help humankind in a better way.

$SIR \rightarrow SEIR \rightarrow SEIFR$

5.2. SIR

Susceptible–Infected–Removed model, developed by Ronald Ross¹, William Hamer. It was first used by Kermack and McKendrick in 1927. It was used for variety of infectious diseases such as: - Measles, airborne diseases.

5.3. SEIR

Susceptible–Exposed–Infected–Recovered model, developed by J.L. Aron and I.B. Schwartz. It is used to predict the infected numbers. It was used for variety of diseases like HIV etc (see Fig. 12).

S = Susceptible
E = Exposed
I = Infected
F = Fatal
R = Recovered

This SEIFR model is the novel approach where we will also take in count the fatality rate so that we can improve the rate by predicting the conditions before they go out of hand.

The complete flow of the proposed SEIFR model is as shown in Fig. 13. The existing SEIR algorithm has been modified so that it can be applied to the given healthcare domain i.e., to calculate infectious fatality rate to analyze the death rate and other features for fatality rate prediction.

Hence, the modifications to an existing SEIR are proposed.

Under this section, the experimental setup, input parameters and implementation of SEIFR are discussed.

5.4. Terminology

$S(t)$ is the number of individuals who are not affected and are even not immune to the disease at time t . If such people come in contact with another person who is the carrier of disease, they may also become exposed.

$E(t)$ is the number of individuals who are infected, i.e., exposed to the virus, but these individuals are asymptomatic, showing no sign or symptoms.

$I(t)$ is the infected individuals who are currently infected and can transmit the disease to others in near time t . These are

the infected individuals which show the sign of virus and also transmit the infection to healthy individuals.

$F(t)$ is the number of individuals who may fall under the category of death. Still, if some special precautions at some specific time t will be taken, then we can even save those individuals via decreasing the rate of fatality and increasing the number of recovered ones.

$R(t)$ are the individuals those have become immune to virus and are successfully recovered from the virus, and these persons might have also developed immunity against the virus.

β = infected individuals due to contact per day.

γ = time period of infectious disease.

$\frac{1}{\gamma}$ = average period of infection

μ = per capita death rate

μN = per capita birth rate

$N = S(t) + E(t) + I(t) + R(t)$ = total population size

$$\frac{dS}{dt} = \mu(N - S) - \beta SI \quad (8)$$

$$\frac{dE}{dt} = \beta SI \quad (9)$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I \quad (10)$$

$$F(t) = S(t) - R(t) \quad (11)$$

$$\frac{dR}{dt} = \gamma I - \mu R \quad (12)$$

Mathematical models are the important tools for analyzing the control and spread of infectious disease. The disease that have longer incubation period are divided into subsequent layers such as: - Susceptible, Exposed, Infected, Fatal and Recovered. Above equations, shows the way to calculate the modeling of data in SEIFR model taking initial parameters as β , γ , & μ as defined above. The unique differentiation of the above parameters yields the results of SEIFR model using the epidemiology of data. The model is run with ordinary differential equations. SEIFR Model tries to predict how a disease spreads or the number of infected people or the duration of epidemic or the onset of symptoms and the recovered stage (see Fig. 14).

In mathematical modeling [] of data, SEIFR layers interact i.e., Susceptible, Exposed, Infected, Fatal, Recovered. Infection starts from the susceptible layer that shows the beginning of the infectious state. The infection may spread to others in the time period of incubation i.e., onset of infection till the seven or nine days from the beginning denoted by $1/\gamma$. This layer is called exposed layer where the droplet infection spreads to other healthy individuals. Then, after two–three days of infection the symptoms arises in the body that becomes the state of onset of symptoms or the infected state. Symptoms such as fever, nausea, cough, breathing problem, sneezing etc. arises. If the person successfully, completes the incubation period of nine days after the onset of infection. Then he/she may enter into recovered state or the person may enter into fatal state. Fatal state is the end of life that shows the immunity of a person is not strong enough that does not went into recovered state. If person successfully completes, incubation period time then recovered state arises from where the person will recover completely. Then, enters into healthy state. The model shows that the recovery rate is good in India. As Indians are showing great recovery against virus and we can even heal ourselves from the pandemic situation only if we take some immediate measures. Protective measures can only help in fighting from the pandemic such as: - wearing masks, maintaining social distance, washing hands with soaps to kill germs, avoid touching face, stay home, avoiding gatherings, and many more to keep oneself safe and healthy. As COVID-19 has led the world into an unprecedented state because there is no such

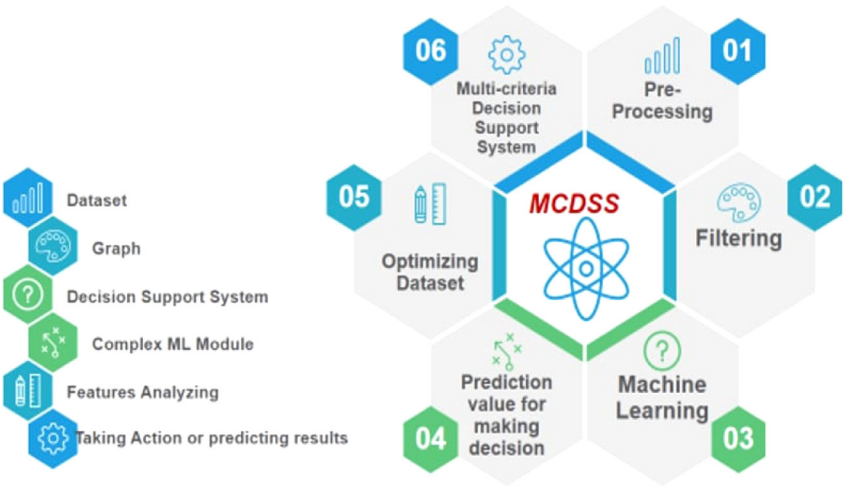


Fig. 13. Flow diagram of the proposed SEIFR model.



Fig. 14. Sequential phase.

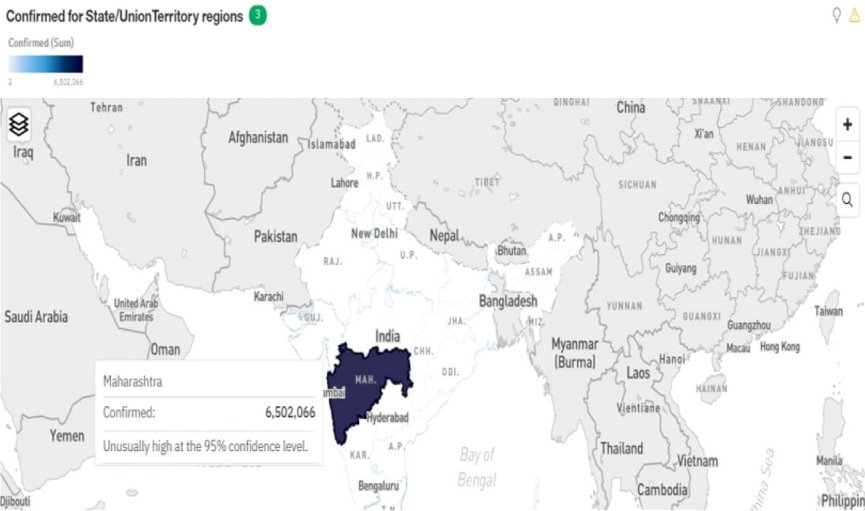


Fig. 15. Map view of worst hit state.

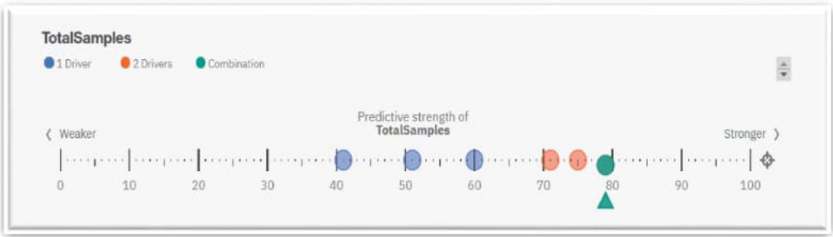


Fig. 16. Predictive driver analysis.

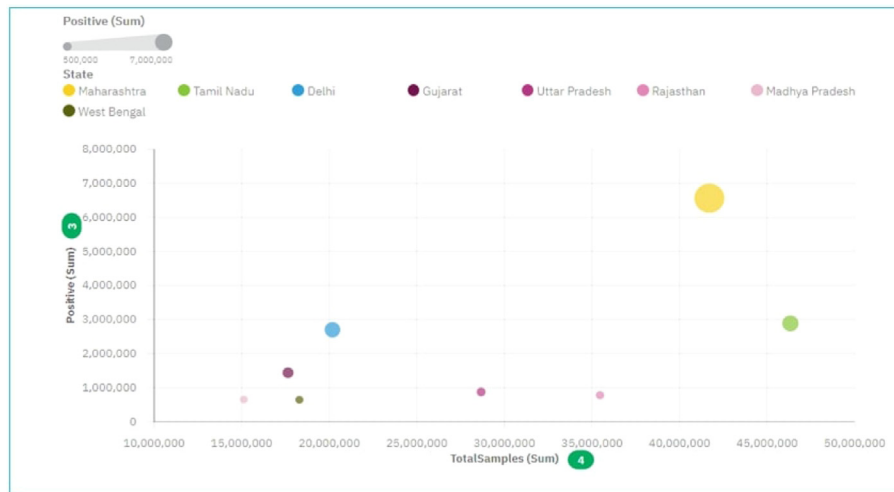


Fig. 17. Prediction analysis of worst eight hit states.

```
In [43]: import scipy.integrate
import numpy
import matplotlib.pyplot as plt
```

Fig. 18. Libraries used for EDA.

```
In [44]: def SEIFR_model(y,t,beta,gamma,mue):
    S,E, I, R = y
    dS_dt = mue*(N-S) - beta*S*I
    dE_dt = beta*S*I
    dI_dt = beta*S*I - gamma*I - mue*I
    dR_dt = gamma*I - mue*R
    return([dS_dt, dE_dt, dI_dt, dR_dt])
```

Fig. 19. Definition of SEIFR model.

symptoms. There are no specific symptoms [30,31] exhibited by the patient. Every patient undergoes treatment according to the symptoms arises. Even some are recovering on their own having mild symptoms with home quarantine (see Fig. 15).

Scenic map view of worst eight hit states through which we predict the confirmed cases out of total population using predictive driver analysis with 95% confidence level. In the total population given we can predict the number of confirmed cases out of total population using prediction driver analysis. If the same situation continues the Cognos IBM software predicts the rate at which the positive or confirmed cases can be predicted from total population.

Predictive driver analysis predicts the cases which can be controlled at great extent if the person knows the proper preventive steps or the measures so that he/she may save from being trap against the virus (see Fig. 16).

Predictive driver analysis out of total samples is performed that shows the probability of number of confirmed or positive cases out of total population. Different predictive strengths are obtained on different parameters. Different parameters matching's are analyzed to calculate the best possible predictive strength for analyzing which was chosen to be 79%.

5.5. Predictive strength obtained on different parameters are

- State is a predictor of total samples with the predictive strength of 41%.
- Number of negative cases is a predictor of total samples with the predictive strength of 51%.
- Number of positive cases is a predictor of total samples with the predictive strength of 60%.
- The combination of positive cases and negative cases is a predictor of total samples with the predictive strength of 71%.
- The combination of negative cases and states is a predictor of total samples with the predictive strength of 75%.
- The combination of states, positive cases and negative cases is a predictor of total samples with the predictive strength of 79%.

Table 8 shows number of people will be found positive or under confirmed cases out of Total population if analyzed using the same predictive strength of 79% (see Fig. 17).

Bubble graph analyzed using predictive driver analysis on IBM Cognos software yields results of top worst eight hit states with their predictive strengths taking in cause the total population,


```
In [45]: S0 = 0.9
E0 = 0.1
I0 = 0.1
R0 = 0.0
N = S0 + E0 + I0 + R0
beta = 0.35
gamma = 0.1
mue = 0.1
t = numpy.linspace(0,50,10000)

solution = scipy.integrate.odeint(SEIFR_model, [S0,E0, I0, R0], t, args=(beta,gamma, mue))
solution = numpy.array(solution)
```

Fig. 20. Initial parameter.

Table 8

Predictive analysis of number of patients out of total population.

States	Confirmed cases (total population)
Maharashtra	65,02,066
TamilNadu	28,89,463
Delhi	27,10,367
Gujarat	14,37,808
UttarPradesh	8,72,302
Rajasthan	7,81,856
WestBengal	6,47,832
MadhyaPradesh	6,54,873

```
In [51]: plt.figure(figsize=[6,4])
plt.plot(t, solution[:,0], label="S(t)")
plt.plot(t, solution[:,1], label="E(t)")
plt.plot(t, solution[:,2], label="I(t)")
plt.plot(t, solution[:,3], label="R(t)")
plt.grid()
plt.legend()
plt.xlabel("Number of people")
plt.ylabel("Range")
plt.show()
```

Fig. 21. Plotting graph of SEIFR model.

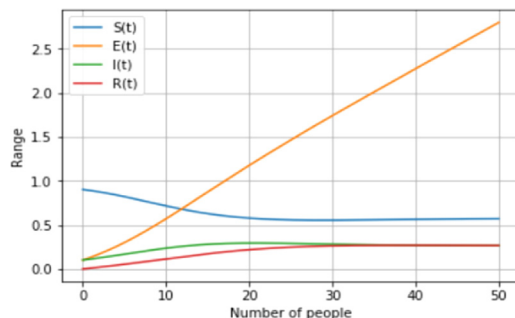


Fig. 22. Graph depicting the relationship between Susceptible, Exposed, Infected & Recovered rate.

positive cases and negative cases in particular state. Based on these factors, predictive strength is calculated.

5.6. Exploratory data analysis (EDA)

To summarize main characteristics from our dataset we used several python libraries (see Figs. 18 and 20–22).

Here, through this graph, we can estimate the amount of susceptible, exposed, infectious, and recovered among the population affected due to COVID-19 the graph (Fig. 19) depicts the number of people exposed in various ways as we can see

some people show susceptible changes. Exposed changes are those changes when people show remarks which even sometimes do not show symptoms. So, the number of exposed cases are high. People showing some infectious remarks generally tend to be recovered, which indicates the number of people recovered signifying the reasonable recovery rate if appropriate measures are taken. The fatality rate to be calculated depends upon the total population size subtracting the recovered ones showing that if nutritional value and other factors like vaccination and other needs since birth to be taken in the cause. We also predict that the Fatality rate is very less, which says that if appropriate measures are taken, we can achieve excellent results in favour of humankind. Thus, this model will contribute a lot in helping the society to have a life free of the pandemic situation by taking appropriate measures at the correct time.

6. Conclusion

While the world continues to fight with the growing impact of COVID-19 on the entire world. Human-kind was fighting with the pandemic situation, complementary efforts of various emerging technologies such as machine learning, IoT, etc. are endeavoring to alleviate its impact. Keeping it as the base foundation of this work we offer some of the mitigation steps and strategies that may help in the fighting pandemic. We begin this paper with a comprehensive review of the COVID-19 itself, in which we explore the multiple criterion's for better decision making In this paper, after analyzing the data of the pandemic situation, the multi-criterion decision making progress is considered to make the best decisions under the situation. It has tried to introduce some basic methodologies (SEIFR) for taking the appropriate decisions under varied alternatives and also some problem-solving techniques to solve the problem. It has analyzed the data of states that are most prone to COVID-19, and also proved the factors using empirical relationship like mean, standard deviation. Results are analyzed for curing the viral disease and providing immunity to the people against the infection. We have also studied the variation in confirmed cases, recovered cases, and the people who died in various states comparing it with the total population, which has projected that generally older people or the children are more prone. The ones with a weak immune system tend to be affected by the virus. Thus, the paper suggests to include some measures like BCG vaccine, nutritional improvement techniques, and guidelines of WHO & AYUSH for guiding the people to immune themselves and save the country from the epidemic and flatten the growing curves. Finally, we hope that this article will make some contributions to the world's response to the situation and will also help in providing some references for future research as well.

7. Future research

In the near future, we will try to develop a prototype that is based on machine learning supporting multi-criterion decision support system(MCDSS) [4,37] so that it can be commercially deployed in fields like healthcare, education and many more areas where diverse nature is to be included for a single decision to be taken. The purpose of building the application will be basically to help the humankind from the pandemic situations and control the situation before they worsen. We will be using a machine learning model that will be free of human intervention. We will train the pre-existing data with supervised learning algorithm where the data is fed with appropriate knowledge with all set of input output possible to take the decision. Then, analysis is possible, which is a totally automated process where we will not consider manual automation. We will make use of the criterion and deciders all in one proving it ubiquitously helpful for serving the society in a better way possible.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Hu Yun, Zhuoran Sun, Jun Wu, Aiguo Tang, Min Hu, Zhongyuan Xiang, Laboratory data analysis of novel coronavirus (COVID-19) screening in 2510 Patients, *Clin. Chim. Acta* 507 (2020) (2019) 94–97.
- [2] Lixiang Li, Zihang Yang, Zhongkai Dang, Cui Meng, Jingze Huang, Haotian Meng, Deyu Wang, Guanhua Chen, Jiaxuan Zhang, Haipeng Peng, Yiming Shao, Propagation analysis and prediction of the COVID-19, *Infect. Dis. Modell.* 5 (2020) 282e292.
- [3] F. Jiang, L. Deng, L. Zhang, Y. Cai, C.W. Cheung, Z. Xia, Review of the clinical characteristics of coronavirus disease 2019 (COVID-19), *J. Gen. Intern. Med.* 35 (2020) 1545–1549, <http://dx.doi.org/10.1007/s11606-020-05762-w>.
- [4] Mark A. Musen, Blackford Middleton, Robert A. Greenes, Clinical decision-support systems, *Biomed. Inform.* (2014) 643–674.
- [5] Palash Ghosh, Rik Ghosh, Bibhas Chakraborty, COVID-19 in India: State-wise analysis and prediction, 2020, <http://dx.doi.org/10.1101/2020.04.24.20077792>, medRxiv preprint.
- [6] Samrat K. Dey, Md. Mahbubur Rahman, Umme R. Siddiqi, Arpita Howlader, Analyzing the epidemiological outbreak of COVID-19: A visual exploratory data analysis approach, *J. Med. Virol.* Wiley (2020).
- [7] V. Surveillances, The epidemiological characteristics of an of 2019 novel coronavirus diseases (COVID-19) China, 2020, *China CDC Wkly.* 2 (8) (2020) 113–122.
- [8] Y. He, X. Wang, J.Z. Huang, Recent advances in multiple criterion decision making techniques, *Clin. Chim. Acta* 507 (2020) 94–97.
- [9] G. Kou, Y. Peng, A bibliography analysis of multi-criterion decision making in computer science (1989–2009), *Cut. Edge Res. Top. Mult. Criterion. Decis. Mak. Commun. Comput. Inf. Sci.* 35 (2009) 68–71.
- [10] Kuldeep Singh, Aryan Agarwal, Impact of weather indicators on the COVID-19 outbreak: A multi-state study in India. CC-BY-NC-ND 4.0 International license, 2020.
- [11] Sudhir B. Jagtap, et al., Census data mining and data analysis using WEKA, in: ICETSTM – 2013, International Conference in Emerging Trends in Science, Technology and Management-2013, Singapore.
- [12] T. Wright, Blockchain app used to track COVID-19 cases in Latin America, 2020, [Online]. Available: <https://cointelegraph.com/news/blockchain-app-used-to-track-covid-19-cases-in-latin-america>.
- [13] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus infected pneumonia in Wuhan, China, *J. Amer. Med. Assoc.* 323 (11) (2020) 1061.
- [14] Albert A. Angehrn, et al., Case-base decision support.
- [15] A. Miller, M.J. Reandelar, K. Fasciglione, et al., Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study. medRxiv. Preprint posted March 28, 2020, <http://dx.doi.org/10.1101/2020.03.24.20042937>.
- [16] Uri Harniel, MD Eran Kozler, MD, Ilan Youngster, MD, MMSc, SARS-CoV-2-rates in BCG-vaccinated and unvaccinated young adults in JAMA May 13, 2020, <http://dx.doi.org/10.1001/jama.2020.8189>.
- [17] 1918 Pandemic (H1N1 Virus), Centers for Disease Control Prevention (CDC), 2020, [Online]. Available: https://www.cdc.gov/_u/pandemic-resources/1918-pandemic-h1n1.html. (Accessed 20 April 2020).
- [18] E.D. Kilbourne, In-uenza pandemics of the 20th century, *Emerg. Infect. Dis.* 12 (1) (2006) 9–14.
- [19] 1957–1958 Pandemic (H2N2 Virus), Centers for Disease Control Prevention (CDC), 2020, [Online]. Available: https://www.cdc.gov/_u/pandemic-resources/1957-1958-pandemic.html. (Accessed 20 April 2020).
- [20] 2009 H1N1 Pandemic (H1N1pdm09 Virus), Centers for Disease Control Prevention (CDC), 2020, [Online]. Available: https://www.cdc.gov/_u/pandemicresources/2009-h1n1-pandemic.html. (Accessed 20 April 2020).
- [21] H. Chen, J. Guo, C. Wang, F. Luo, X. Yu, W. Zhang, J. Li, D. Zhao, D. Xu, Q. Gong, J. Liao, H. Yang, W. Hou, Y. Zhang, Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: A retrospective review of medical records, *Lancet* 395 (10226) (2020) 809–815.
- [22] S. Salehi, A. Abedi, S. Balakrishnan, A. Gholamrezaezhad, Coronavirus disease 2019 (COVID-19): A systematic review of imaging findings in 919 patients, *Amer. J. Roentgenol.* (2020) 1–7, <http://dx.doi.org/10.2214/AJR.20.23034>.
- [23] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J. Xia, T. Yu, X. Zhang, L. Zhang, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study, *Lancet* 395 (10223) (2020) 507–513.
- [24] Suman Madan, Puneet Goswami, K-DDD measure and map reduce based anonymity model for secured privacy-preserving big data publishing, *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* 27 (02) (2019) 177–199, <http://dx.doi.org/10.1142/S0218488519500089>, published in a SCI(E) and Scopus indexed journal with ISSN (print): 0218–4885 | ISSN (online): 1793–6411.
- [25] Layla Safwat Jamil, Data analysis based on data mining algorithms using Weka workbench, Thomson Reuters EndNote 5 (8) (2016).
- [26] Aaditya Desai, Sunil Rai, Analysis of machine learning algorithms using weka in IEEE, 2013.
- [27] WHO, Coronavirus disease 2019 (COVID-19) situation report 87, 2020, [Online]. Available: <https://www.who.int/docs/defaultsource/coronaviruse/situationreports>.
- [28] WHO, Coronavirus disease 2019 (COVID-19) situation report 79, 2020, [Online]. Available: <https://www.who.int/docs/defaultsource/coronaviruse/situationreports>.
- [29] F. Shi, et al., Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19, *IEEE Rev. Biomed. Eng.* (2020) <http://dx.doi.org/10.1109/RBME.2020.2987975>.
- [30] Waleed, et al., Multi-criterion framework for emergency department in Irish hospital, in: C. Laroque, J. Himmelsbach, R. Pasupathy, O. Rose, A.M. Uhrmacher (Eds.), Proceedings of the 2012 Winter Simulation Conference, IEEE, 2012, 978-1-4673-4780-8/12/\$31.00 ©.
- [31] Nengjun, et al., A decision support system with intelligent recommendation for multi – disciplinary medical treatment, *ACM Trans. Multimedia Comput. Commun. Appl.* 16 (1) (2020) <http://dx.doi.org/10.1145/3352573>, Article 33, Publication date.
- [32] Charlie H. Zhang, Gary G. Schwartz, Spatial disparities in coronavirus incidence and mortality in the United States: An ecological analysis as of May 2020, *J. Rural Health*.
- [33] M. Frackowiak, J. Darlak, A. Kobylinska, T. Lund, Factbox: Europe Begins Easing Coronavirus Lockdowns, Reuters, 2020, [Online]. Available: <https://www.reuters.com/article/us-health-coronaviruseurope-factbox>.
- [34] M. Godin, These European Countries are Slowly Lifting Coronavirus Lockdowns. Here's What that Looks Like, Time, 2020, [Online]. Available: <https://time.com/5822470/countries-lifting-coronavirusrestrictions->.
- [35] R. Martin, What countries are still in lockdown and how many weeks has it been for them? Metro, 2020, Available: <https://metro.co.uk/2020/04/17/countries-still-lockdown-many-weeks-1257>.
- [36] R. Staff, Here's What Lockdown Looks Like Around the World, World Economic Forum, 2020, Available: <https://www.weforum.org/agenda/2020/04/coronavirus-lockdowns-global/>.
- [37] V. Chauhan, S. Galwankar, B. Arquilla, M. Garg, S. Di Somma, A. El-Menyar, V. Krishnan, J. Gerber, R. Holland, S.P. Stawicki, Novel coronavirus (COVID-19): Leveraging telemedicine to optimize care while minimizing exposures and viral transmission, *J. Emerg. Trauma Shock* 13 (1) (2020) 20.

Further reading

- [1] A. Mosavi, A multi-criterion decision making environment for engineering design and production decision-making, *Int. J. Comput. Appl.* 69 (1) (2013).
- [2] G. Wang, A. Gunasekaran, E.W. Ngai, T. Papadopoulos, Big data analytics in logistics and supply chain management: Certain investigations for research and applications, *Int. J. Prod. Econ.* 176 (2016) 98–110.
- [3] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, *Health Inf. Sci. Syst.* 2 (1) (2014) 3, <https://link.springer.com/article/10.1186/2047-2501-2-3>.

- [4] Michael M. Trauer, Frcem, PgCert, Ashley Matthies, Nick Mani, Cian Mcdermott, Robert Jarman, Utility of lung ultrasound in COVID-19: A systematic scoping review, 2020, medRxiv.
- [5] Novel coronavirus (COVID-19) cases data, 2020, <https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases>. (Accessed: 06 June 2020).
- [6] V. Chamola, V. Hassija, V. Gupta, M. Guizani, A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact, IEEE Access 8 (2020) 90225–90265, <http://dx.doi.org/10.1109/ACCESS.2020.2992341>.
- [7] World Health Organization, Digital technology for Covid-19 response, 2020, [Online]. Available: <https://www.who.int/newsroom/detail/03-04-2020-digital-technology-for-%covid-19-response>.