J. Inst. Eng. India Ser. B (December 2021) 102(6):1249–1263
https://doi.org/10.1007/s40031-021-00547-z

ORIGINAL CONTRIBUTION

South Asian Countries are Less Fatal Concerning COVID-19: A Fact-finding Procedure Integrating Machine Learning & Multiple Criteria Decision-Making (MCDM) Technique

Soham Guhathakurata[1] · Sayak Saha[1] · Souvik Kundu[2] · Arpita Chakraborty[3] · Jyoti Sekhar Banerjee[3]

Received: 14 August 2020 / Accepted: 18 January 2021 / Published online: 6 March 2021

Abstract Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which caused an outbreak of pneumonia in December 2019 in Wuhan, China, has spread rapidly throughout the world. This ongoing pandemic has resulted over 55.6 million cases of COVID-19 leading to 1.34 million deaths in more than 188 countries. However, it has been observed that the death rate is significantly in the lower side for the SAARC countries compared to the First World Nations. In this paper, the possible factors have been represented that determine this uneven distribution of COVID-19 deaths. The significance of the factors has been presented in this paper after the data analysis of the factors from 165 different countries. Based on the correlation of the factors and their critical impact towards the concerned countries death toll, the risk index of each factor has been labeled using analytical hierarchy process (AHP)-based MCDM, i.e., multiple criteria decision-making technique. The risk index of all the factors has been used to generate the susceptibility of COVID-19 for each of the countries in study, specifically the SAARC Nations. Finally, the hierarchical clustering was applied to visualize the death toll of the countries corresponding to their susceptibility index.

Keywords COVID-19 · Risk factor · Susceptibility index · Hierarchical clustering · AHP · MCDM

✉ Jyoti Sekhar Banerjee
tojyoti2001@yahoo.co.in

1 Department of CSE, Bengal Institute of Technology, Kolkata, India 700150

2 Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA

3 Department of ECE, Bengal Institute of Technology, Kolkata, India 700150

Introduction

The World Health Organization declared the outbreak of COVID-19 [1] a Public Health Emergency of International Concern on January 30, 2020, and a pandemic on 11 March. The first confirmed death was in Wuhan on January 9, 2020. The first death outside China occurred on February 1, 2020, in Philippines and the first death outside Asia was in France on 14 February [2]. Since then, more than 188 countries and territories have had at least one case of COVID-19 so far. Containment measures such as quarantines and curfews have been imposed by many countries to restrict the spread of the virus. By late April, around 300 million people were under lockdown in many European countries, while 200 million people were under some form of lockdown in the USA. However, the death toll in USA has been constantly rising with over 2,39,853 deaths as of November 13, 2020. Similarly, the ratio of confirmed cases to deaths is significantly high for the First World Nations in Europe and Asia compared to the SAARC countries. South Asia accounts for approximately 1.5% of the total coronavirus cases worldwide and even a lower percentage of deaths nations.

The South Asian Association for Regional Cooperation (SAARC) includes India, Afghanistan, Bangladesh, Maldives, Nepal, Pakistan, Bhutan and Sri Lanka. One-fifth of the world's population resides in these nations. However, the death rate has been significantly low in spite of the high population and mediocre heath facilities. India has accounted for 83,64,086 confirmed cases and 1,24,354 deaths, which is the highest among the SAARC nations as of November 13, 2020, whereas the USA has 239,853 deaths as of November 13, 2020, and the death toll of France, Italy, Spain, the UK, and Russia is all high in comparison with their population.

Table 1 signifies the vast difference in the conversion rate of confirmed to death between the SAARC nations and the other developed countries. In spite of the fact that India's population is equivalent to 17.7% of the total world population and Dhaka, the capital of Bangladesh being the sixth most populated city in the world, the death toll is still low.

The paper focuses on the factors which are responsible for such a low death rate in SAARC nations. In spite of the fact that COVID-19 is highly transmissible, the transmission rate in these nations is low given that these countries house one-fifth of the world population. The factors [3] in the study include (1) The Average Temperature; the warmer and more humid weather in the South Asian region can temper the spread of the disease, (2) The Bacillus Calmette-Guerin (BCG) vaccine offered for the protection against tuberculosis in these nations creates a better immune response against the virus, (3) Critical Days, which signifies the minimum number of days taken by the government to take action after the first occurrence of COVID-19 case in that country, (4) Average age of the country; the younger population respond significantly better than the older population which plays a crucial for the death count of the countries, and (5) The Herd Immunity, which is an indirect protection against infectious diseases that occurs when a large percentage of a population has become immune to an infection. Observations and studies have unanimously indicated that these factors have played a significant role in the low death rate of the SAARC nations. We have assigned logical weights to these factors using analytical hierarchy process to calculate the susceptibility risk index of every country. Finally, the hierarchical clustering has been applied to visualize the distribution countries death rate corresponding to their risk index.

Related Studies

In the recent past, a lot of work in the field of market research, bioinformatics, data processing, image processing, etc., has been done by using the hierarchical clustering algorithm. A brief review is presented here.

Ying Zhao, George Karypis and Usama Fayyad in [4] applied hierarchical clustering [5] to cluster documents. In order to remove the early-stage error and improve the quality of clustering solutions the authors combined the features from both partitional and agglomerative approaches. They concluded that constrained agglomerative methods result in better solutions than agglomerative methods alone and observed that for significantly high cases outperform partitional methods, as well.

Feng Luo, Kun Tang, and L Khan in [6] proposed to obtain gene expression pattern by using a new hierarchical clustering which constructs a hierarchy from top to bottom and dynamically find the number of clusters at each level. The authors observed that the algorithm can recognize features in complex gene expression data by extracting patterns with different levels of abstraction.

Deng Cai et. al. [7] proposed a hierarchical method to cluster the web image search results into different semantic

Table 1 Statistical data of COVID-19 for (**a**) countries with highest deaths and (**b**) SAARC countries as of November 13, 2020

	Countries	Confirmed	Deaths
(a)	USA	9,802,374	239,853
	Russia	1,712,858	29,509
	Spain	1,356,798	38,118
	UK	1,099,059	47,742
	Brazil	5,590,941	161,170
	Germany	597,453	11,035
	Turkey	384,509	10,558
	France	168,693	38,674
(b)	India	8,364,086	124,354
	Pakistan	338,875	6893
	Bangladesh	416,006	6021
	Afghanistan	41,925	1554
	Sri Lanka	12,187	29
	Maldives	11,822	38
	Nepal	182,923	1052
	Bhutan	358	0

clusters. The hierarchical clustering method uses textual, visual and link analysis to obtain three kinds of representations. The representations comprise visual feature-based representation, textual feature-based representation and graph-based representation.

Seema Bandhyopadhyay and E.J. Coyle in [8] proposed a distributed, randomized clustering algorithm. The algorithm can organize the sensors in a wireless sensor network into cluster. The authors generated a hierarchy of cluster heads and observed an increase in energy savings with the increase in the level of hierarchy.

Richard Cheng and Glenn W. Milligan in [9] applied hierarchical clustering to map Influence Regions. They presented three-dimensional response surface plots for several hierarchical clustering methods and simulated core group data structures. The relative influence of the corresponding coordinate location is represented by the response surface in the bivariate data space on the clustering of the core groups. Substantial differences between clustering methods are revealed by the substantial plots.

Twinkle Tiwari and Nihar Ranjan Roy in [10] applied hierarchical clustering in heterogeneous wireless sensor networks. Wireless sensor network can sense the physical parameters like pressure, humidity, temperature, motion, etc. It is a network of small battery powered sensing node which collects information from its environment of deployment and finally report it to a central node called base station. These nodes fulfill their task by proper collaboration. The energy source should be used properly since it is constrained in WSNs. Author used clustering to minimize energy dissipation in WSNs.

Michael R. Loken et. al. in [11] proposed a system to investigate the relationship between the presenting immunophenotype by applying hierarchical clustering. This system will response to therapy in a large, controlled study of pediatric AML patients. Patients with similar diagnostic IEPs are grouped which is defined mathematically from the analysis of unsupervised hierarchical clustering. By minimizing within-cluster variation, appropriate number of clusters was accomplished.

Dac-Tu Ho et. al. in [12]. proposed Particle Swarm Optimization (PSO) as an optimization method to find the optimal clusters. This system reduces the energy consumption, bit error rate (BER), and UAV travel time. Low Energy Adaptive Clustering Hierarchy (LEACH) is commonly used to conserve energy in conventional wireless sensor networks (WSNs). Conversation of energy is highly challenging for large-scale deployments than many other things.

Andy Podgursky and Charles Yang in [13] presented a new approach to estimate software reliability by reducing the manual labor required. Partition testing methods along with those of stratified sampling are combined to reduce the sample size necessary to estimate reliability with a given degree of precision. Automatic cluster analysis is used to stratify program execution and group them with similar features.

Though the South Asian Countries are less fatal concerning COVID-19, this fact is not revealed in any published work as per the authors' best knowledge. However, in very few researches, it is considered that the death toll in South Asian countries is less than the rest of the world. Jennifer Beam Dowd and Liliana Andriano et al. in [14] showcased the impact of demographic structure on the mortality rate by taking the examples of Italy and South Korea. Similarly, authors like Yueling Ma et. Al. in [15] and Abhibhav Sharma et al. in [16] have portrayed the significance of temperature and BCG vaccine in correspondence to the mortality rate due to COVID-19 in their respective publications. However, the authors of this paper claim to have brought upon the different critical factors leading to the lower death count in SAARC nations on the basis of the analysis of the risk factor, which is so far a totally unique work.

Causes of Less Disaster Concerning COVID-19 in South Asian Counties

The biggest threat posed by COVID-19 [17, 18] is its spread rate. Despite the high transmission rate, conversion rate of confirmed cases to death is not same everywhere. Significant difference has been noticed for the SAARC countries with respect of confirm cases to death.

In Fig. 1, the statistical visualization has been presented of the death count in eight different countries which includes four SAARC countries depicted with dashed lines. The death count is so low in the SAARC countries that the four different lines are clustered in same space. Though the death count in India is quite high concerning other SAARC countries, at the same time, the total population of India needed to be considered, which is the second highest in the world.

Fundamental factors responsible for less fatality in South Asian Countries

The fundamental factors include 1. Bacillus Calmette-Guerin (BCG) vaccine, 2. Average Temperature, 3. Average Age, 4. Critical Days, 5. Herd Immunity and are described below elaborately.

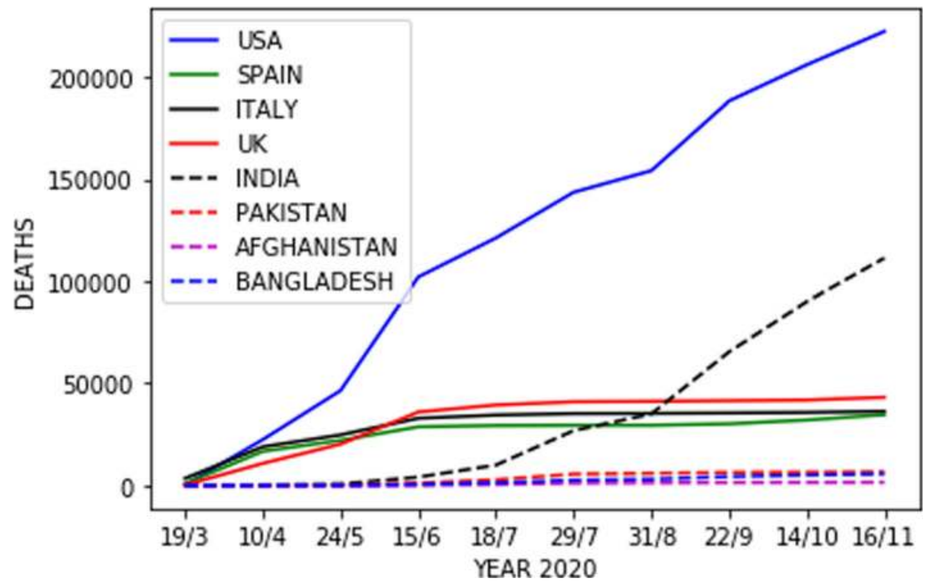

Fig. 1 World death count due to COVID-19 as of November 16, 2020

1. Bacillus Calmette-Guerin (BCG) Vaccine: BCG vaccine is mainly used to protect against tuberculosis [19]. Even though this vaccine does not protect against COVID-19, it strengthens the immune systems. Studies [20, 21] have proved that the recovery chances from COVID-19 increasing in persons with a strong immune system (Table 2). The vaccine is taken in every SAARC nation, but none of the countries with high death count takes it (see Table T3).

2. Average Temperature: The major cities with the highest death rates include Zurich, London, Berlin and Paris. During the months of February, March and April all these countries had an average temperature of 5.60 °C to 6.10 °C [22]. For the SAARC nations, the temperature increased to 21.30 °C to 29.80 °C. Studies have shown that every 10 °C increase in Average Temperature leads to a decrease in the cumulative number of cases by 0.86 [23].

Figure 2 gives a clear visualization of dependency of death per million with respect to the Average Temperature. The graph is plotted with respect to the data collected from 165 countries till April 31. The Average Temperature of the whole year has been considered given the fact that the pandemic is up and running. Since three SAARC countries fall in the Average Temperature range of 70 °C to 150 °C and display a bimodal distribution with respect to rest of the countries in the world. The weight given to average

Table 2 Distribution of BCG vaccine taken by (**a**) countries with highest deaths and (**b**) SAARC countries

	Countries	Confirmed	Deaths	BCG taken
(a)	USA	9,802,374	239,853	No
	Russia	1,712,858	29,509	No
	Spain	1,356,798	38,118	No
	UK	1,099,059	47,742	No
	Brazil	5,590,941	161,170	Yes
	Germany	597,453	11,035	No
	Turkey	384,509	10,558	Yes
	France	168,693	38,674	No
(b)	India	8,364,086	124,354	Yes
	Pakistan	338,875	6893	Yes
	Bangladesh	416,006	6021	Yes
	Afghanistan	41,925	1554	Yes
	Sri Lanka	12,187	29	Yes
	Maldives	11,822	38	Yes
	Nepal	182,923	1052	Yes
	Bhutan	358	0	Yes

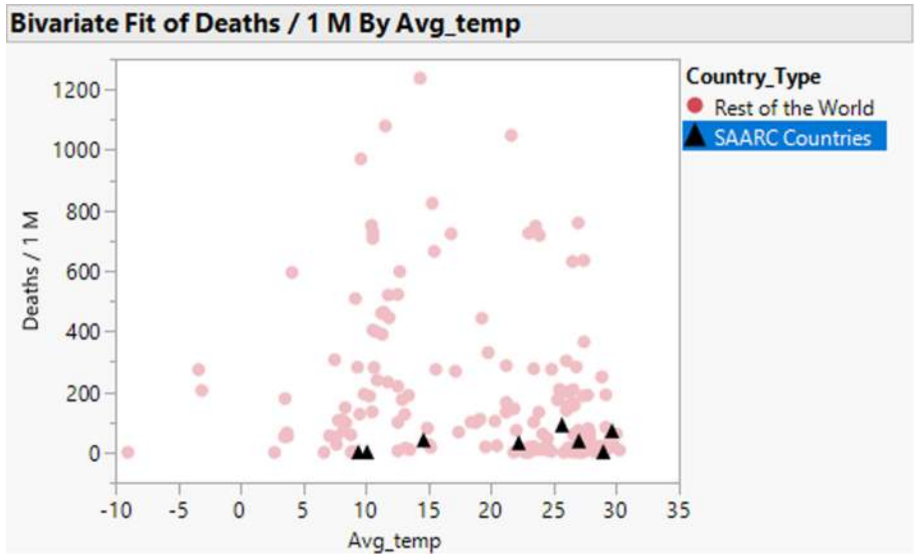

Fig. 2 Death per million vs. Average Temperate, as of November 13, 2020

temperature during calculation of risk factor (RF) in Eq. (1) will have the least value which makes our prediction more robust and scalable.

3. Critical Days: The Critical Days signify the number of days taken by the government to take action, i.e., localized recommendation, national recommendation, localized lockdown, national lockdown [24] after the appearance of first COVID-19 confirm case in that country. Sri Lanka has taken action even before the country had its first confirm case. India declared the lockdown in March 25, 2020, when the overall death count of the country was below 50. By contrast, most of the states in the USA did not have a proper lockdown [2]. European countries like Sweden did not impose a complete lockdown. When compared to the COVID-19 spread on the scale of population density, the SAARC countries have shown better results. The impact of the Critical Days is relatively high in the case of COVID-19 compared to other infectious diseases is because of its high transmission rate.

The countries in the left section of Fig. 3 are all SAARC nations who has taken faster action compared to the countries in the right section of Fig. 3. This visualization shows the low margin of Critical Days in the SAARC nations compared to the countries with high death counts [25].

4. Average Age: Reports have shown that older population respond poorly to COVID-19 compared to a younger population (Fig. 4). One of the main reasons could be the strong immune system of the younger people. Another reason is the underlying diseases that the aged person carries tend to weaken their immune system and thus they become more susceptible to get COVID-19 infection. The share of deaths is a low rate of 3.9% for the age range of 18–44 years. The rate jumps to 24.9% for 65–74 years old

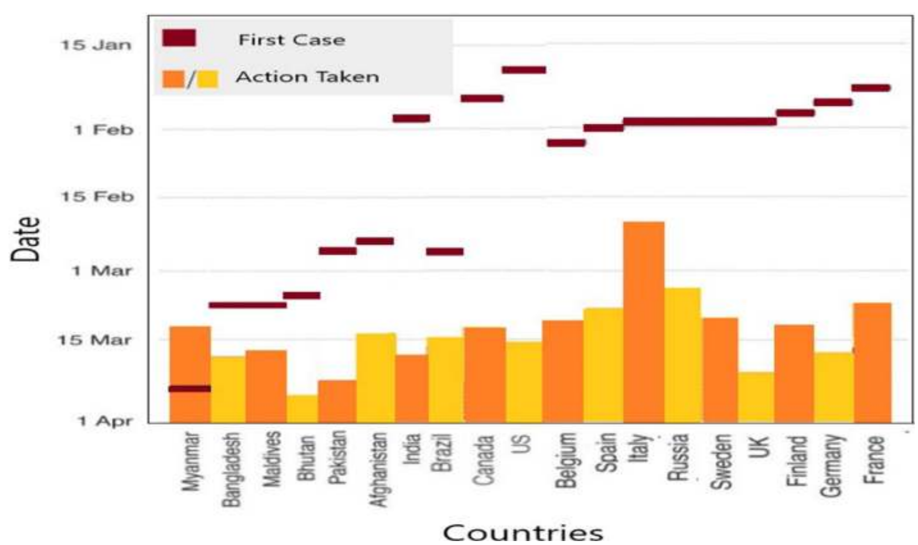

Fig. 3 Duration between first step taken and first COVID-19 confirm case

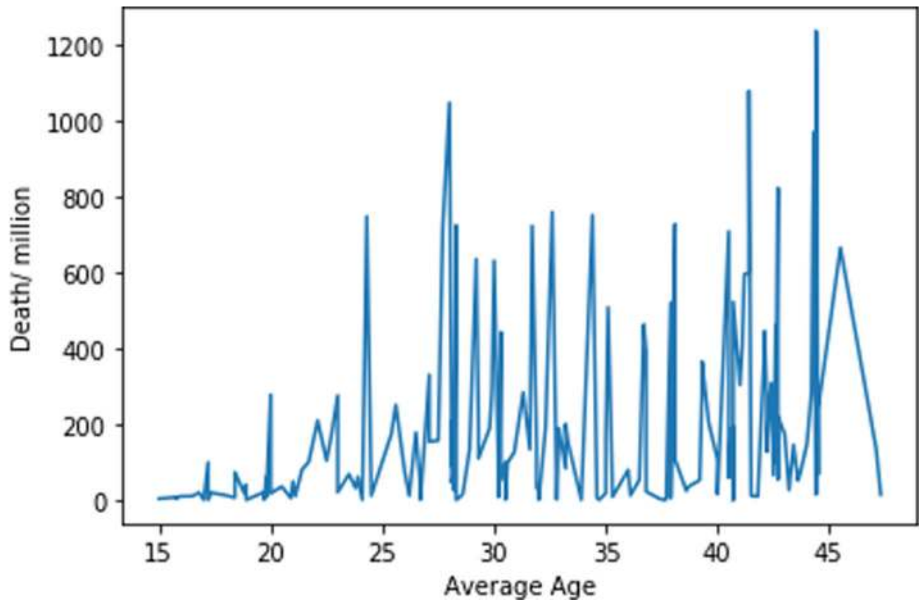

Fig. 4 Death per million vs. average age of a country, as of November 13, 2020

and to 48.7% for above 75 years of age [26]. From these statistics (Table 3), it is clear that the younger population is having a significantly low death rate concerning the aged population. The average age of an Indian is 26.8, while the number is even low for Bangladesh, Nepal and Pakistan. In contrast, countries like Italy, Germany, France and UK [27] all with high death counts have an average age of over 40.

5. Herd Immunity: South Asian countries are exposed to more infectious diseases and germs compared to the western world countries. Since these nations are exposed to more pathogens, the white blood cells develop a broader memory that can trigger an immune response when a virus is recognized. With past history [28] of different infectious diseases like cholera, malaria, dengue, SARS-COV-1 [29, 30], etc., the people tend to possess a greater variation in the leukocyte antigen genes, responsible for immune response. As a result, the immune system fights back in a better way by producing proteins called antibodies. This is a critical factor to observe in our study. The low number of conversions from confirmed to death cases in SAARC countries also signifies that their immune system is churning out more antibodies to fight against COVID-19. Apart from the SAARC nations, it has been observed that the African countries who have also been a victim of lot of past diseases like Ebola, Zika, have a significantly low death count. Their genetic diversity has offered a better protection against COVID-19 compared to the European countries and American countries.

The facts stated and displayed above reflects their significance on the COVID-19 death count. These factors have been applied for each country corresponding to their death count to generate three clusters, namely low risk, moderate risk and high risk. To determine this, we have applied analytical hierarchy process to generate the weights of these five factors based on their individual impact towards COVID-19 deaths. In the paper, a relation has been established between each of these factors and the death rate to generate the risk of COVID-19 for each of the 165 countries. Finally, the hierarchical clustering has been applied to plot the three clusters based on their risk index and death per million counts. The suggested methodology is depicted in Fig. 8.

Enumeration of risk factor (RF)

To investigate the situation holistically with respect to the above attributes, the authors have tried to calculate 'risk factor (RF)' associated with individual country. High Average Temperature, usage of BCG and immunity earned or Herd Immunity are all inversely proportional to the death count and high average age and high value of Critical days are directly proportional to the death count. Hence, RF is formulated as follows:

$$\begin{aligned} RF = {} & 0.419 \times (\text{Avg. Age}) + 0.263 \times (1/\text{BCG}) + 0.16 \\ & \times (1/\text{Herd Immunity}) \\ & + 0.097 \times (\text{Critical Days}) + 0.062 \\ & \times (1/Avg.Temperature) \end{aligned} \tag{1}$$

where the local weights of the five factors, i.e., Avg. Age, BCG given in % of the total population, Immunity earned, Critical Days, and Avg. Temperature of the country, are obtained through MCDM [31–34], i.e., multiple criteria decision-making technique [50–61], particularly analytical hierarchy process (AHP) [35–38], defined as Y, is presented below:

Table 3 Statistical data of COVID-19 for (**a**) countries with highest deaths/1 M and (**b**) SAARC countries concerning Avg. Age as of November 13, 2020

	Countries	Deaths/ I M	Avg. Age
(a)	USA	727	38.1
	Russia	205	39.6
	Spain	823	42.7
	UK	708	40.5
	Brazil	759	32.6
	Germany	134	47.1
	Turkey	126	30.9
	France	598	41.4
(b)	India	90	28.1
	Pakistan	31	23.8
	Bangladesh	37	26.7
	Afghanistan	40	18.9
	Sri Lanka	1	32.8
	Maldives	70	28.2
	Nepal	0	24.1
	Bhutan	0	26.7

$$Y = \begin{array}{c c} & \begin{array}{ccccc} x1 & x2 & x3 & x4 & x5 \end{array} \\ \begin{array}{c} x1 \\ x2 \\ x3 \\ x4 \\ x5 \end{array} & \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1/2 & 1 & 2 & 3 & 4 \\ 1/3 & 1/2 & 1 & 2 & 3 \\ 1/4 & 1/3 & 1/2 & 1 & 2 \\ 1/5 & 1/4 & 1/3 & 1/2 & 1 \end{pmatrix} \end{array} \quad (2)$$

In this paper, the Average Age was recognised is the most significant deciding factor, BCG given in % of the total population is the second, Immunity earned is the third, Critical Days is the fourth, and Average Temperature of the country is the fifth vital deciding factor. Equation 2 shows the relative weight of the deciding factors.

The normalized weights of the factors are calculated by using an online computing software [39] as W = {0.419, 0.263, 0.16, 0.097, 0.062}, which shows weights of the Average Age, BCG given in % of the total population, Immunity earned, Critical Days, and Average Temperature of the country, sequentially. The Average Temperature of the country is the least significant and the Average Age is the most important criterion, it is cleared from Fig. 5.

To calculate the risk factor, we have generated a formula based on the weights calculated and the one-to-one relationship between each of the factors and the death rate. High Average Temperature, usage of BCG and Immunity earned or Herd Immunity are all inversely proportional to the death count. High average age and high value of Critical Days are directly proportional to the death count.

Clustering countries into different risk regions through hierarchical clustering algorithm

The countries are grouped into three clusters like, high risk, low risk and moderate risk based on the RF calculated using Eq. 1 on the dataset produced. This grouping of countries into different risky zones is done in this correspondence employing *hierarchical clustering algorithm*.

Hierarchical clustering is the most popular and widely used clustering techniques in Machine Learning [40–45]. Here nodes are compared with one another based on their similarity and instead of producing a single clustering; they produce a hierarchy of clustering. Larger groups are built by joining groups of nodes based on their similarity. A criterion gets introduced to compare nodes based on their relationship. There are mainly two types of hierarchical clustering approaches prevail; namely: *Agglomerative approach* and *Divisive approach*. The second one gets utilized in this study. This is a top-down approach in contrast to bottom-up approach of agglomerative method. In Divisive approach initially, all nodes are supposed to belong to the same cluster; gradually, each node forms its own cluster. The hierarchical clustering technique can be visualized using a dendrogram which is basically a tree-like structure that portrays the sequences of merges or splits (Fig. 6).

Step-by-step procedure of the proposed methodology

Through this paper, the authors want to establish that South Asian countries are less fatal concerning COVID-19. In

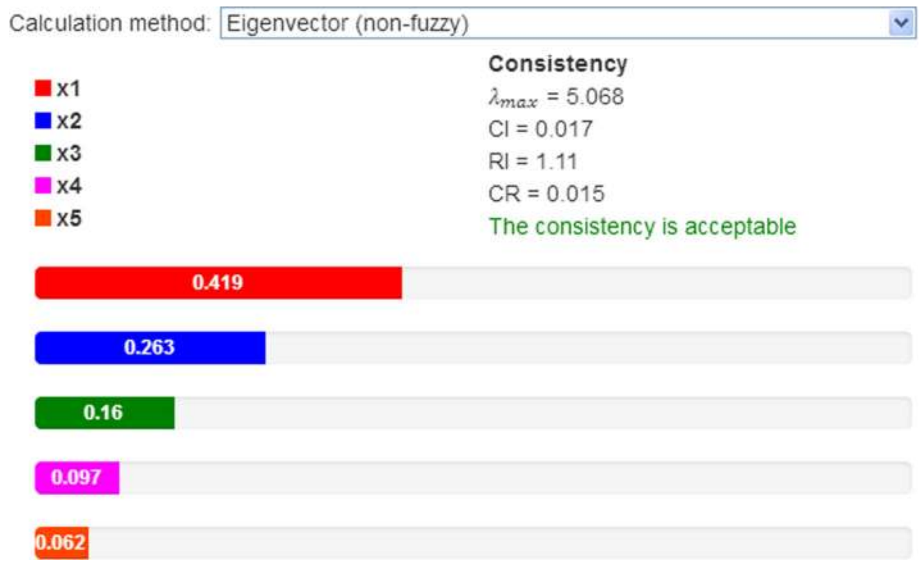

Fig. 5 Weights of the deciding factors using Saaty's AHP

Fig. 7, the step-by-step procedure of the proposed methodology is depicted. Firstly, the fundamental factors responsible have been selected for less fatality in south Asian countries, i.e., BCG, Average Temperature, Critical Days, Average Age, Herd Immunity. Secondly, the weights of the individual attributes are calculated through AHP. Thirdly, the attributes are to be labeled with individual weights. Fourthly, the risk factor needs to be formulated combining labeled attributes and also calculated using the dataset for every individual country. Then, countries are clustered into low risk, moderate risk, and high risk through the hierarchical clustering technique and the SAARC countries are analyzed based on the Clustering concerning other countries.

Experimental results and discussion

The objective of our study is based on the fact that in spite of being such a densely populated area, SAARC countries still have a low death count when compared to the first world countries. The dataset used by the authors for this report has been coined from different online sources [19, 25, 46–49]. The authors have framed a dataset containing the attributes – population density (per km^2), death/million and whether the country belongs to SAARC or not (see Table 4). The dataset includes eight SAARC countries, countries with high amount of deaths, first world countries along with Singapore and South Korea whose early counter COVID-19 strategy inspired the World.

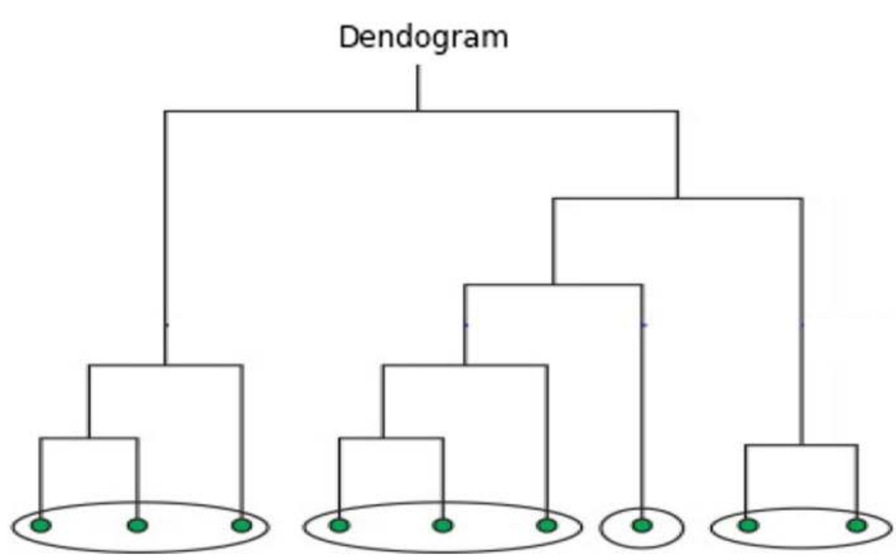

Fig. 6 Dendrogram: an example of hierarchical clustering approach

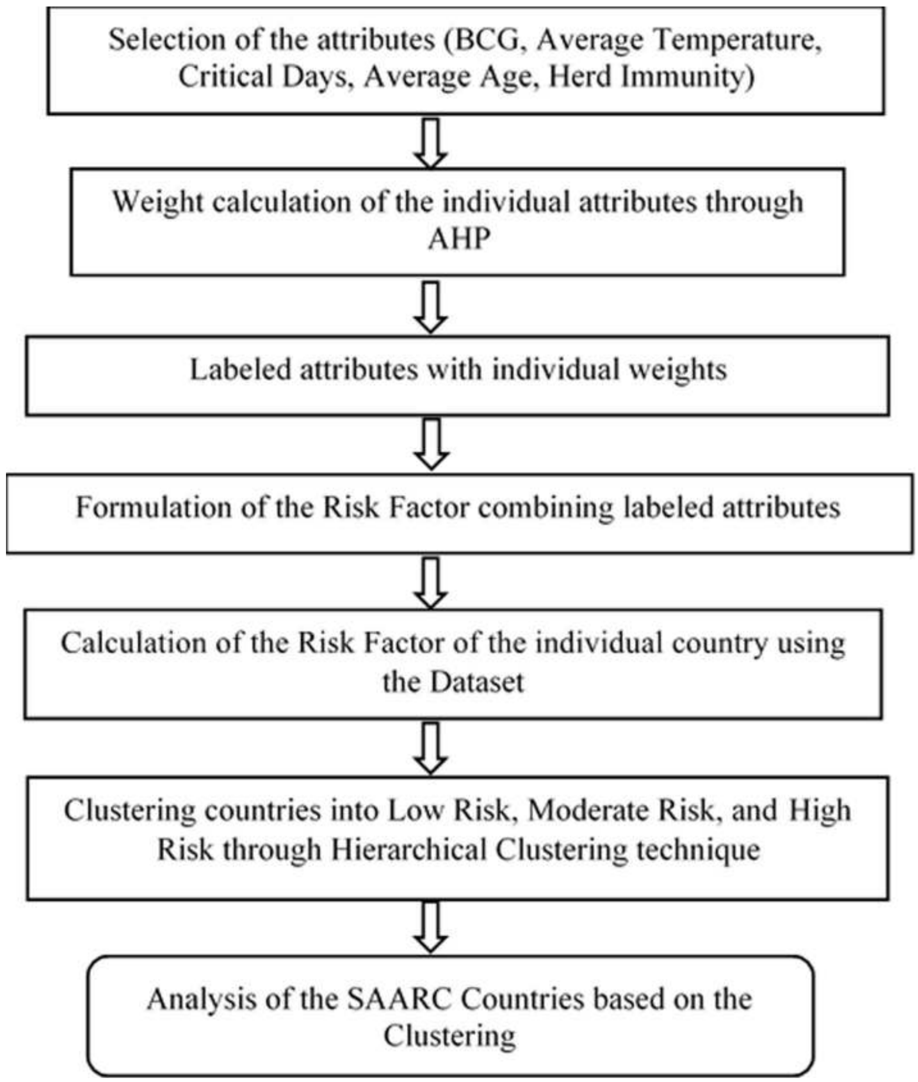

Fig. 7 Step-by-Step Procedure of the Proposed Methodology

Table 4 Dataset-1 snapshot

Country	Density (km^2)	Deaths / 1 M	SAARC
Maldives	1802	70	Yes
Bangladesh	1265	37	Yes
Sint Maarten	1261	303	No
Bermuda	1246	145	No
Channel Islands	907	275	No
San Marino	566	1237	No
S. Korea	527	9	No
India	464	90	Yes

Figure 8 depicts that even with high population density in SAARC countries like Bangladesh and Maldives, the death count of COVID-19-infected person is low. In the next half of this section, we present the experimental results to showcase how the five factors introduced in section 3 governs the COVID-19 death count and why SAARC enjoys such a low death rate.

The dataset that has been created for the study of the impact of risk factor and death count contains the data of 165 countries have been considered till November 13, 2020.

The dendrogram has been plotted (Fig. 9) taking the risk factor (RF) as x-axis and death/million as y-axis. The dendrogram illustrates the arrangement of the clusters produced by the corresponding analyses. A horizontal line is passed though the center of the longest vertical line which in our case is the blue line. Since the horizontal line cuts through three vertical lines, the optimum number of clusters for this particular dataset is three.

The dataset has been feeded to the hierarchical clustering model [29] and set the parameter of number of clusters to 3.

In Fig. 10, the countries are grouped into three clusters: high risk, low risk and moderate risk. The figure clearly explains that the countries with the low risk factor also have low death count. The countries with high death count that includes UK, USA, France, and Italy are all grouped together in the high-risk cluster. Some of the countries with

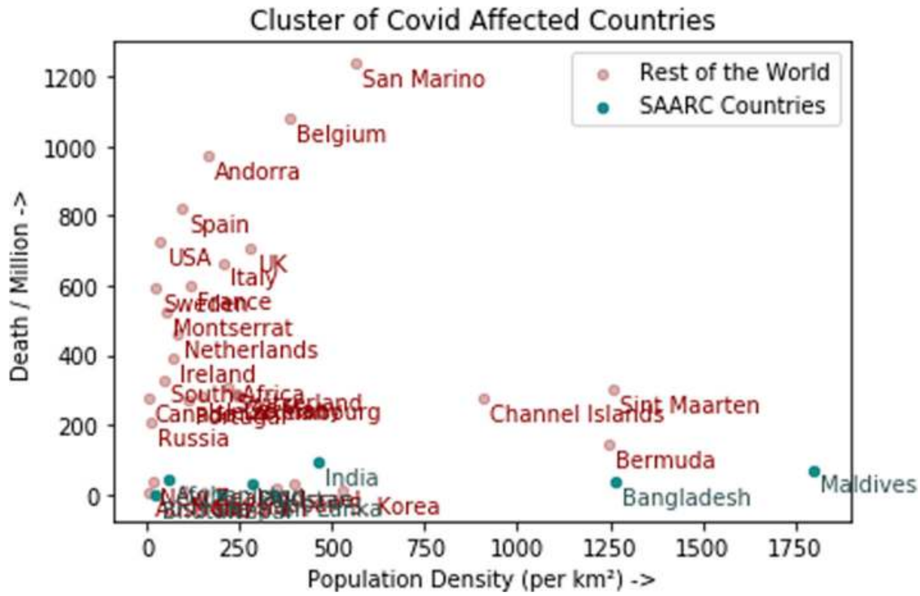

Fig. 8 Distribution of countries on the basis of death/million vs. population density, as of November 13, 2020

Table 5 Dataset-2 Snapshot

Country	BCG	Avg_temp	CriticalStage	Avg_Age	Immunity Earned	Deaths / 1 M
San Marino	No	14.33	17.0	44.4	No	1237
Belgium	No	11.55	11.55	41.4	No	1079
Andorra	No	9.6	8.0	44.3	No	970
Spain	No	15.3	38.0	42.7	Yes	823
Italy	No	15.45	20.0	45.5	Yes	665
UK	No	10.55	51.0	40.5	Yes	708
France	No	12.7	44.0	41.4	Yes	598
Sint Maarten	No	26.0	18.0	41.0	No	303

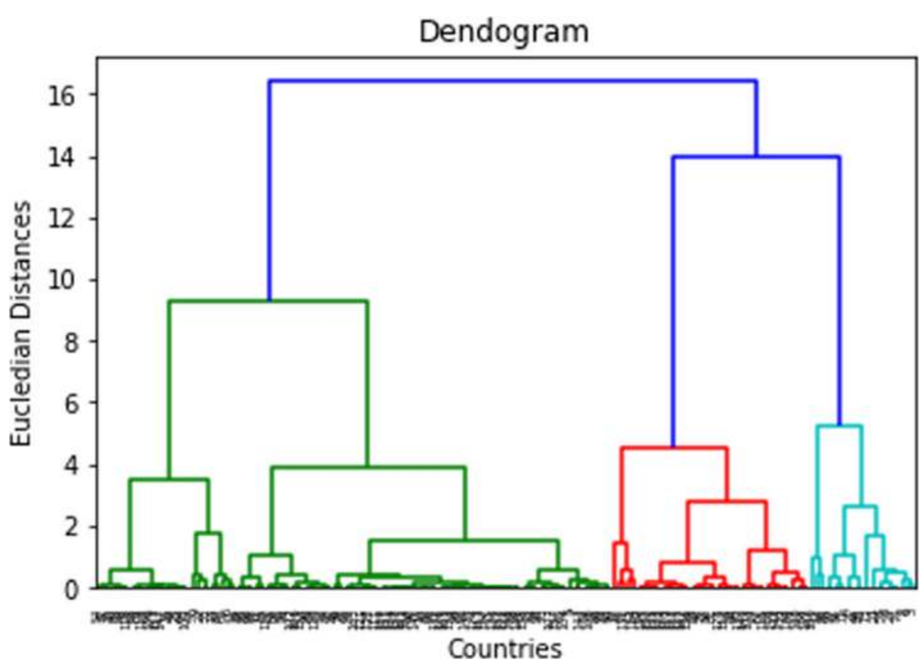

Fig. 9 Dendrogram

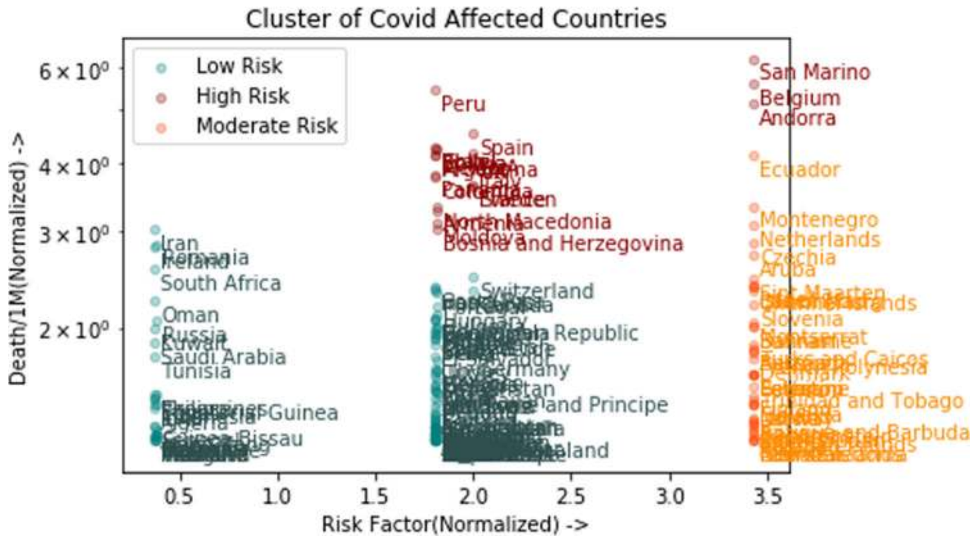

Fig. 10 Cluster of COVID-19-affected countries (death/ million vs risk factor), as of November 13, 2020

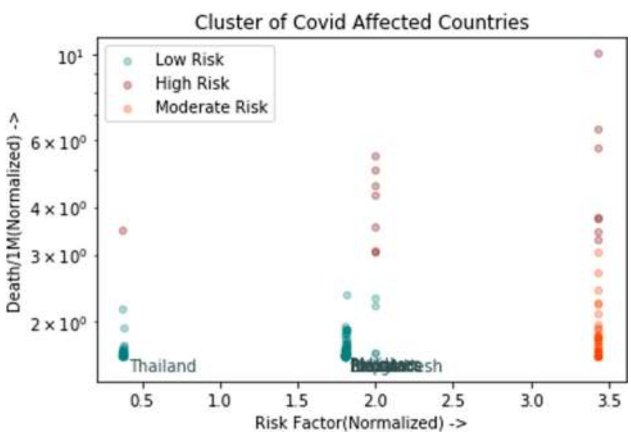

Fig. 11 Cluster of affected SAARC countries (death/ million vs risk factor), as of November 13, 2020

high risk factor have low deaths which is mainly due to their low population. To have a clear view where the SAARC countries stand with respect to Fig. 10, an extra filter has been applied to label the SAARC countries only.

The distribution of the SAARC countries (labeled and colored blue) is depicted in Fig. 11. It is clearly observable that the SAARC countries enjoy a low risk factor for which the death count is also low. The results prove the impact of the factors that the authors have taken into consideration. Apart from the SAARC countries, the countries that satisfy the ideal conditions of the five governing factors also display a low death count as shown in Fig. 10.

Concluding remarks

COVID-19 virus is mutating and evolving day-by-day. The number of infected cases is increasing at a very high rate. Many of the countries like USA, France, and Italy have reached the stage of community spread. In contrast, India, Pakistan, and Bangladesh with such high population density are in the local transmission stage. Other SAARC countries have a very low infected case. This paper has tried to establish the factors that segregate the SAARC countries from the countries with high infected cases and high death counts. The factors have been considered based on their individual impact on 165 different countries. Each of the factors has been weighted after studying their interrelation with death count. The risk index calculated using AHP shows the proper result when mapped with the real data. The objective of the study has been justified with cluster graphs which helps us to visualize where the SAARC countries stand with regard to the COVID-19 death count when compared to the other developed nations with superior medical facilities.

Appendix

Analytical hierarchy process (AHP)

The algorithm of AHP mainly having four steps as follows [31–39]:

Table T1 1 to 9 AHP scale (Saaty proposed)

Saaty's rating	Description
1	Equal significance
3	Moderate significance
5	Essential or strong significance
7	Very strong significance
9	Extreme significance
2, 4, 6, 8	Intermediate values of the two adjacent judgments

Table T2 Values of RI

Average Random Index (RI)										
a	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Table T3 Distribution of BCG vaccine taken worldwide

Country	BCG
USA	No
Spain	No
Italy	No
UK	No
France	No
Germany	No
Russia	Yes
Turkey	Yes
Brazil	Yes
Iran	Yes
China	Yes
Canada	No
Belgium	No
Peru	Yes
India	Yes
Netherlands	No
Switzerland	No
Ecuador	No
Saudi Arabia	Yes
Portugal	Yes
Mexico	Yes
Sweden	No
Ireland	Yes
Pakistan	Yes
Chile	Yes
Singapore	Yes
Belarus	Yes
Israel	No
Qatar	Yes
Austria	No
Japan	Yes
UAE	Yes
Poland	Yes
Romania	Yes
Ukraine	Yes
Indonesia	Yes
S. Korea	Yes
Bangladesh	Yes
Denmark	No
Serbia	No

Table T3

Philippines	Yes
Dominican Republic	Yes
Norway	No
Czechia	No
Colombia	Yes
Panama	Yes
Australia	No
South Africa	Yes
Egypt	Yes
Malaysia	Yes
Finland	No
Kuwait	Yes
Morocco	Yes
Argentina	Yes
Algeria	Yes
Moldova	Yes
Kazakhstan	Yes
Luxembourg	No
Bahrain	No
Hungary	Yes
Thailand	Yes
Afghanistan	Yes
Oman	Yes
Greece	Yes
Nigeria	Yes
Armenia	Yes
Iraq	Yes
Uzbekistan	Yes
Ghana	Yes
Croatia	Yes
Azerbaijan	Yes
Bosnia and Herzegovina	Yes
Iceland	No
Estonia	Yes
Bulgaria	Yes
Cuba	Yes
Bolivia	Yes
North Macedonia	Yes
New Zealand	No
Slovenia	No
Lithuania	Yes
Slovakia	No
Ivory Coast	Yes
Senegal	Yes
Djibouti	Yes
Honduras	Yes
Hong Kong	Yes
Tunisia	Yes
Latvia	No
Cyprus	No
Kyrgyzstan	Yes
Albania	Yes
Niger	Yes
Andorra	No
Lebanon	No
Costa Rica	Yes
Sri Lanka	Yes
Guatemala	Yes
Uruguay	Yes
Georgia	Yes
San Marino	No
Mali	Yes
El Salvador	Yes
Channel Islands	No

Table T3

Maldives	Yes
Kenya	Yes
Malta	Yes
Jamaica	Yes
Jordan	Yes
Taiwan	Yes
Paraguay	Yes
Venezuela	Yes
Palestine	No
Mauritius	Yes
Montenegro	No
Isle of Man	No
Equatorial Guinea	Yes
Vietnam	Yes
Guinea-Bissau	Yes
Faeroe Islands	No
Cabo Verde	Yes
Myanmar	Yes
Madagascar	Yes
Gibraltar	No
Ethiopia	Yes
Brunei	No
Zambia	Yes
Togo	Yes
Trinidad and Tobago	No
Bermuda	No
Eswatini	No
Aruba	No
Uganda	Yes
Haiti	Yes
CAR	Yes
Bahamas	No
Guyana	Yes
Barbados	Yes
Liechtenstein	No
Mozambique	Yes
Sint Maarten	No
Nepal	Yes
Cayman Islands	No
Libya	Yes
French Polynesia	No
South Sudan	No
Malawi	Yes
Mongolia	Yes
Angola	Yes
Zimbabwe	Yes
Antigua and Barbuda	No
Timor-Leste	Yes
Botswana	Yes
Belize	Yes
New Caledonia	No
Gambia	Yes
Curaçao	No
Sao Tome and Principe	Yes
Burundi	Yes
Turks and Caicos	No
Montserrat	No
Greenland	Yes
Suriname	No
Mauritania	Yes
Bhutan	Yes

Step 1 Based on the given problem, the decision hierarchy needs to be formed with the sub-problems or the independent factors.

Step 2 Decision-maker needs to calculate and decide the weights to the judgment factors. Pairwise comparison executed on each decision factor.

Step 3 It is very essential to confirm that the consistency ratio is maintained based on the calculation.

Step 4 The final ranking of each alternative is obtained considering the synthesized overall result.

Considering the pairwise comparison of each judgment factor, an evaluation matrix Y is to be formed. Experts will choose the value of the weight factors from the Saaty's 1 to 9 fundamental scale (Table T1). Now, square matrix Y can be written as,

$$Y = \left(y_{pq}\right)_{a\times a} = \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{pmatrix} \tag{3}$$

pth and the qth judgment factor is expressed by y_{pq} and here, a number of decision factors is considered. Each entry of the matrix can be written as the pairwise ratio, i.e., $y_{pq} = \mathrm{w_p}/\mathrm{w_q}$ where p, $\mathrm{q} = 1, 2,...,\mathrm{a}$. In AHP, each entry of the matrix maintains the reciprocal property $y_{\mathrm{pq}} = 1/y_{qp}$. The following condition needs to meet for becoming matrix Y as a consistent matrix;

$$y_{pr}y_{rq} = y_{\mathrm{pq}}\ ,\ \mathrm{p,\ q,\ r} = 1, 2,..........,\mathrm{a} \tag{4}$$

The equation $Yw_{AHP} = \lambda_{\max}w_{AHP}$ provides the eigenvalue $\lambda_{\max}$ and eigenvector w_{AHP}. Priority vector $w_{AHP} = \{\mathrm{w}_1, \mathrm{w}_2, ...\mathrm{w}_m\}$ expresses the local weights of the criteria. Due to the inconsistency of the matrices, it is necessary to find out the decision error, which is can be measured by the consistency ratio (C.R.). C.R. is simply a ratio of consistency index (C.I.) to random index (R.I.), viz, $C.R. = \mathrm{C.I.} : \mathrm{R.I.}$. Table T2 shows all the R.I. values corresponding to the various number of decision factors. Again, C.I. can be described as $C.I. = \frac{(\lambda_{\max} - \mathrm{a})}{(\mathrm{a\ -\ 1})}$.

Where,

$$\lambda_{\max} = \left(\frac{1}{a}\right)\sum_{p=1}^{a}\frac{\left(Yw_{AHP}\right)_p}{w_{AHPp}} \tag{5}$$

Normally, $C.I. = 0$ is expected for consistent matrix and the accepted value of the $C.R. \le 0.1$, otherwise when $C.R. > 0.1$, adjustment is required pairwise comparison. Using Saaty's AHP, the final ranking of the alternatives is expressed by global ranking $w_{AHP\,Global}$ where w_{AHPq} are the local weights,

$$w_{AHP\,Global} = w_{AHP\,\mathrm{intermediate}\,q} \cdot w_{AHPq} \tag{6}$$

References

1. WHO-China Joint Mission, Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), (2020). https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf, Accessed 13 Nov 2020
2. https://en.wikipedia.org/wiki/Covid-19, Accessed 13 Nov 2020
3. https://theprint.in/health/why-south-asia-has-20-of-worlds-population-but-less-than-2-of-covid-19-cases/408471/, Accessed 13 Nov 2020
4. Y. Zhao, G. Karypis, U. Fayyad, Hierarchical clustering algorithms for document datasets. Data Min Knowl Disc **10**(2), 141–168 (2005)
5. Bar-Joseph, Z., Gifford, D. K., & Jaakkola, T. S. (2001). Fast optimal leaf ordering for hierarchical clustering. Bioinformatics, 17(suppl_1), S22-S29.
6. F. Luo, K. Tang, L. Khan, in *Third IEEE Symposium on Bioinformatics and Bioengineering, 2003. Proceedings*, Hierarchical clustering of gene expression data, (IEEE, New York, 2003), pp. 328–335
7. D. Cai, X. He, Z. Li, W.Y. Ma, J.R. Wen, in *Proceedings of the 12th annual ACM international conference on Multimedia*, Hierarchical clustering of WWW image search results using visual, textual and link information, (2004) pp. 952–959
8. S. Bandyopadhyay, E.J. Coyle, in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, An energy efficient hierarchical clustering algorithm for wireless sensor networks, (IEEE, New York, 2003), pp. 1713–1723
9. R. Cheng, G.W. Milligan, Mapping Influence Regions in Heirarchical Clustering. Multivar Behav Res **30**(4), 547–576 (1995)
10. T. Tiwari, N.R. Roy, in *International Conference on Computing, Communication & Automation*, Heirarchical clustering in heterogeneous wireless sensor networks: a survey, (IEEE, New York, 2015), pp. 1385–1390
11. M.R. Loken, A.P. Voigt, R.B. Gerbing, L.E. Brodersen, A.J. Menssen, L. Pardo... & S. Meshinchi, Heirarchical clustering of immunophenotypic cell surface antigen expression identifies clinically meaningful cohorts in childhood AML: A report from the Children's Oncology Group protocol AAML0531. (2015)
12. D.T. Ho, E.I. Grøtli, P.B. Sujit, T.A. Johansen, J.B. Sousa, in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, Cluster-based communication topology selection and UAV path planning in wireless sensor networks, (IEEE, New York, 2013), pp. 59–68
13. A. Podgurski, C. Yang, Partition testing, stratified sampling, and cluster analysis. ACM SIGSOFT Softw Eng Notes **18**(5), 169–181 (1993)
14. J.B. Dowd, L. Andriano, D.M. Brazel, V. Rotondi, P. Block, X. Ding ... & M.C. Mills, Demographic science aids in understanding the spread and fatality rates of COVID-19. Proc. Nat. Acad. Sci. **117**(18), 9696–9698 (2020)
15. Y. Ma, Y. Zhao, J. Liu, X. He, B. Wang, S. Fu, …B Luo, , Effects of temperature variation and humidity on the death of COVID-19 in Wuhan China. Sci Total Environ **724**, 138226 (2020)
16. A. Sharma, S.K. Sharma, Y. Shi, E. Bucci, E. Carafoli, G. Melino, …G. Das, BCG vaccination policy and preventive chloroquine usage: do they have an impact on COVID-19 pandemic? Cell death dis. **11**(7), 1–10 (2020)
17. S. Guhathakurata, S. Kundu, A. Chakraborty, J.S. Banerjee, in *Data Science for COVID*-19, A novel approach to predict COVID-19 using support vector machine, (Elsevier, New York, 2021)
18. S. Biswas, L.K. Sharma, R. Ranjan, J.S. Banerjee, Go-COVID: An Interactive Crossplatform Based Dashboard for Real-time Tracking of COVID-19 using Data Analytics. J. Mech. Continua. Math. Sci. **15**(6), 1–15 (2020)
19. http://www.bcgatlas.org/, Accessed 6 May 2020
20. https://www.bloomberg.com/news/articles/2020-04-02/fewer-coronavirus-deaths-seen-in-countries-that-mandate-tb-vaccine, Accessed 13 Nov 2020
21. https://www.newindianexpress.com/opinions/2020/apr/29/coronavirus-and-the-bcg-vaccine-2136722.html, Accessed 13 Nov 2020
22. J. Wang, K. Tang, K. Feng, W. Lv, High temperature and high humidity reduce the transmission of COVID-19. Available at SSRN 3551767, (2020)
23. M. Wang, A. Jiang, L. Gong, L. Luo, W. Guo, C. Li, Y. Chen, Temperature significant change COVID-19 Transmission in 429 cities. MedRxiv. (2020)
24. https://academic.oup.com/jtm/article/27/3/taaa037/5808003, Accessed 13 Nov 2020
25. https://www.bbc.com/news/world-52103747, Accessed 10 May 2020
26. https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/, Accessed 13 Nov 2020
27. https://www.vox.com/2020/3/12/21173783/coronavirus-death-age-covid-19-elderly-seniors, Accessed 13 Nov 2020
28. https://www.jhsph.edu/covid-19/articles/achieving-herd-immunity-with-covid19.html, Accessed 13 Nov 2020
29. https://en.wikipedia.org/wiki/2002%E2%80%932004_SARS_outbreak, Accessed 13 Nov 2020
30. https://www.worldometers.info/coronavirus/. Accessed 13 Nov 2020
31. S. Paul, A. Chakraborty, J.S. Banerjee, in *Cognitive Informatics and Soft Computing*, The extent analysis based fuzzy AHP approach for relay selection in WBAN, (Springer, Singapore, 2019), pp. 331–341
32. S. Paul, A. Chakraborty, J.S. Banerjee, in *2017 4th International Conference on Opto-Electronics and Applied Optics (Optronix)*, A fuzzy AHP-based relay node selection protocol for wireless body area networks (WBAN), (IEEE, New York, 2017), pp. 1–6
33. J.S. Banerjee, A. Chakraborty, A. Chattopadhyay, in *Advances in Electronics, Communication and Computing*, Relay node selection using analytical hierarchy process (AHP) for secondary transmission in multi-user cooperative cognitive radio systems, (Springer, Singapore Springer, Singapore, 2018), pp. 745–754
34. J.S. Banerjee, A. Chakraborty, A. Chattopadhyay, in *Advances in Optical Science and Engineering*, Fuzzy based relay selection for secondary transmission in cooperative cognitive radio networks, (Springer, Singapore, 2017), pp. 279–287
35. O. Saha, A. Chakraborty, J.S. Banerjee, in *2017 4th International Conference on Opto-Electronics and Applied Optics (Optronix)*, A decision framework of IT-based stream selection using analytical hierarchy process (AHP) for admission in technical institutions, (IEEE, New York, 2017), pp. 1–6
36. O. Saha, A. Chakraborty, J.S. Banerjee, in *Emerging Technologies in Data Mining and Information Security*, A fuzzy AHP approach to IT-based stream selection for admission in technical institutions in India (Springer, Singapore, 2019), pp. 847–858
37. J.S. Banerjee, A. Chakraborty, A. Chattopadhyay, Reliable best-relay selection for secondary transmission in co-operation based cognitive radio systems: a multi-criteria approach. J Mech Continua Math Sci **13**(2), 24–42 (2018)
38. J.S. Banerjee, A. Chakraborty, A. Chattopadhyay, A novel best relay selection protocol for cooperative cognitive radio systems using fuzzy AHP. J Mech Continua Math Sci **13**(2), 72–87 (2018)
39. P. Holecek, J. Talašová, in *Proceedings of the 34th International Conference on Mathematical Methods in Economics*, A free

software tool implementing the fuzzy AHP method (2016), pp. 266–271
40. J. Banerjee, S. Maiti, S. Chakraborty, S. Dutta, A. Chakraborty, J.S. Banerjee, in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Impact of machine learning in various network security applications (IEEE, New York, 2019), pp. 276–281
41. A. Chakraborty, J.S. Banerjee,, An advance Q learning (AQL) approach for path planning and obstacle avoidance of a mobile robot. Int. J. Intell. Mech. Robot. (IJIMR) IGI Glob. **3**(1), 53–73, (2013)
42. I. Pandey, H.S. Dutta, J.S. Banerjee, in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. WBAN: a smart approach to next generation e-healthcare system (IEEE, New York, 2020), pp. 344–349
43. J. Chattopadhyay, S. Kundu, A. Chakraborty, J.S Banerjee, in *International Conference on Computational Vision and Bio Inspired Computing*, Facial expression recognition for human computer interaction (Springer, New York, 2020), pp. 1181–1192
44. S. Biswas, L.K. Sharma, R. Ranjan, S. Saha, J.S. Banerjee, in *Recent Trends in Computational Intelligence Enabled Research*, Smart farming & water saving based intelligent irrigation system implementation using IoT. (Elsevier, Netherlands, 2021)
45. P. Saha, S. Guhathakurata, S. Saha, A. Chakraborty, J.S. Banerjee, in *2020 1st Global Conference on Artificial Intelligence and Applications (GCAIA)*, Application of machine learning in app-based cab booking system: a survey on indian scenario (Springer, New York, 2021)
46. https://en.wikipedia.org/wiki/List_of_countries_by_median_age. Accessed 13 Nov 2020
47. https://www.weather-atlas.com/en/climate. Accessed 13 Nov 2020
48. https://en.wikipedia.org/wiki/Pandemic. Accessed 13 Nov 2020
49. https://en.wikipedia.org/wiki/Hierarchical_clustering
50. A. Chakraborty, J.S. Banerjee, A. Chattopadhyay, "Malicious node restricted quantized data fusion scheme for trustworthy spectrum sensing in cognitive radio networks. J Mech Contin Math Sci **15**(1), 39–56 (2020)
51. A. Chakraborty, J.S. Banerjee, A. Chattopadhyay, Non-uniform quantized data fusion rule for data rate saving and reducing control channel overhead for cooperative spectrum sensing in cognitive radio networks. Wireless Pers Commun **104**(2), 837–851 (2019)
52. A. Chakraborty, J.S. Banerjee, A. Chattopadhyay, in *2017 IEEE 7th International Advance Computing Conference (IACC)*, Non-uniform quantized data fusion rule alleviating control channel overhead for cooperative spectrum sensing in cognitive radio networks, (IEEE, New York, 2017), pp. 210–215
53. R. Roy, S. Dutta, S. Biswas, J.S. Banerjee, in *Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019) (2020)*. Android things: a comprehensive solution from things to smart display and speaker., nitttr chandigarh, india (Springer, Singapore, 2020) pp. 339–352
54. D. Das, I. Pandey, A. Chakraborty, J.S. Banerjee, Analysis of implementation factors of 3D printer: the key enabling technology for making prototypes of the engineering design and manufacturing. Int J Comput Appl **975**(8827), 8–14 (2017)
55. D. Das, I. Pandey, J.S. Banerjee, in *Proc. MICRO 2016 conference on microelectronics, circuits and systems*, An in-Depth Study of Implementation Issues of 3D Printer (2016) pp. 45–49.
56. J.S. Banerjee, D. Goswami, S. Nandi, in *Proc International Conference on Contemporary Challenges in Management, Technology & Social Sciences*, OPNET: A new paradigm for simulation of advanced communication systems (SEMS, Lucknow, 2019) pp. 319–328.
57. J.S. Banerjee, A. Chakraborty, Fundamentals of software defined radio and cooperative spectrum sensing: a step ahead of cognitive radio networks, in *Handbook of Research on Software-Defined and Cognitive Radio Technologies for Dynamic Spectrum Management*. ed. by N. Kaabouch, W. Hu (IGI Global, USA, 2015), pp. 499–543
58. J.S. Banerjee, A. Chakraborty, Modeling of software defined radio architecture & cognitive radio, the next generation dynamic and smart spectrum access technology, in *Cognitive Radio Sensor Networks: Applications, Architectures, and Challenges*. ed. by M.H. Rehmani, Y. Faheem (IGI Global, USA, 2014), pp. 127–158
59. J.S. Banerjee, A. Chakraborty, K. Karmakar, in *Cognitive Radio Technology Applications for Wireless and Mobile Ad Hoc Networks*, ed. by N. Meghanathan,, Y.B. Reddy Architecture of cognitive radio networks, (IGI Global, USA, 2013), pp. 125–152
60. J.S. Banerjee, K. Karmakar, A comparative study on cognitive radio implementation issues. Int J Comput Appl **45**(15), 44–51 (2012)
61. S. Guhathakurata, S. Saha, S. Kundu, A. Chakraborty, J.S Banerjee, J. S, in *Cyber-physical systems: AI and COVID-19*, A New Approach to Predict COVID-19 using Artificial Neural Networks (Elsevier, New York, 2021 in press)