

Covid-19 Outbreak Analysis

CH V S S Nikil,
Optum,
Telangana, India
saisainikil@gmail.com

Hemlata Dalmia
St Peters Engineering College,
Telangana, India
dalmiahemlata@gmail.com

G Janaki Rama Pavan Kumar
Wipro Technologies,
Telangana, India
janakiramapavankumar@gmail.com

Abstract: “The mind is the attribute of man. When man is born, he comes into existence with only one weapon with him—the reasoning mind”. Now considering the current scenario and seriousness of COVID 19, human is ready to get challenged in this crucial situation. The paper presents the proposal is to do analysis of COVID affected patients. Machine learning models help in identifying the severity and decision making. Understanding pandemic will help in implementing proper measures at right time. Implementing these measures prior to the outbreak will help maintaining country’s health and economy. Here we observe and analyze the behavior of data, which helps in providing better results for decision making. The model proposed in this paper is in adjunct with an interactive application that gives updates regarding the pandemic; indicates the different zones of the local area prescribed by the government and also serves a platform for needy.

Keywords: COVID 19, Machine learning, decision making

I. Introduction

COVID-19 is the biggest challenge that entire world is fighting now. This pandemic is considered as a health emergency across India and few other countries. The rapid pervasiveness of this virus raises the number of cases. There is a need to mitigate the virus which could be achieved by implementing appropriate measures based on the number of cases.

COVID-19 is third novel Coronavirus after SARS-CoV and MERS-CoV which causes diseases in animals, birds and respiratory infections in humans. This was originated from bats and then passed on to humans. Huanan sea food wholesale market, Wuhan, China is identified as origin for covid-19. At present there is no vaccine for this virus. researchers, pharmaceutical scientists, many companies around the globe are working to find the vaccine [1].

Data science plays a pivotal role in understanding the covid-19 pandemic. It is used to analyze the data both structured and unstructured, potential insights could be drawn using this. WHO declared COVID-19 as pandemic, data science helps in Risk assessment, patient prioritization, screening, diagnosis, contact tracing, Automated patient care, vaccine discovery and

economic interventions? Data science provides support in dealing with the pandemic and helps in decision making [2].

With the widespread of the virus, WHO changed the status of covid-19 from epidemic to pandemic. Implementing Deep learning to analyze this situation could be accurate and efficient. Using the data that was built based on computed tomography scan of patients along with the explainable deep learning algorithms and CT scan, prediction models could be developed which help in identifying victim [3].

COVID-19 is spreading rapidly with doubling rate in less time around the world which is a major threat globally especially to old aged people. Few people affected with virus had shown ill health mainly respiratory problems and also asymptomatic symptoms. Vaccine for this pandemic is not ready yet because of the pathological mechanism. Till then people should maintain social distance also improve immunity to fight against virus [4].

Death count will be raising as patients cannot survive till the discovery of vaccine. Deaths due to the virus could be controlled by identifying patients at early stages. Artificial Intelligence can be used for fighting the pandemic. AI-based drug will have high efficiency with minimal cost. Using AI, the time to find the right combination could be reduced [5].

COVID-19 has asymptomatic and symptomatic cases, out of these symptomatic cases have adverse effects on people. WHO announced that the virus is spreading locally because of the close contact and transmission of water droplets from COVID patients? It is advised to maintain 3 feet social distance and active use of sanitizers as preventive measures. As the life span of the virus is up to 72 hours, it is advised to sanitize things and keep up hygiene [6].

II. Related Work

The Middle East Respiratory Syndrome (MeRs-CoV) is a respiratory disease observed first in 2012 in Saudi Arabia, caused by a virus that belongs to the coronavirus’s family. using naive bayes classifier and

J48 decision tree algorithm for the prediction of patients' recovery from the infection. Based on the data published by the Saudi ministry of health, about recovered, deaths and new cases these models were implemented considering age as the primary factor for prediction. These models resulted with accuracies 53.63% and 55.69% respectively [7].

Machine learning models are generally used for classification, prediction and clustering. The best model (standard) are equipped with high accuracy and less error rate. Accuracy of the model depends on the data. Soft computing models like Multilayer perceptron and adaptive neuro fuzzy influence systems are used for predicting the number of cases along with machine learning models. End results proclaim that machine learning models have higher accuracy than the other two models [8].

As advancement in technology is playing key role in almost all sectors in risk assessment. Data driven solutions were most helpful in improving preparedness and minimizing the impact by identifying risks at earlier stages. Rapid growth in digital technologies resulted precision in health care industry and also plays vital role in future. AI tools are used for detection because of their high accuracy and performance. AI algorithms detected the severity of coronavirus and alarmed world 7 days prior to WHO. Adoption to this technology will help in this pandemic situation in order to take decisions accurately [9]. AI is used to deal with challenges in almost all tech companies. It has an important role in fighting with Covid-19. It is used to warn or alert people soon, create interactive, dynamic dashboards, tracking and predicting people, diagnosis, treatment, social control. Efficiency in implementing AI will depend on the data [10].

III. Proposed Algorithm

In our proposed methodology, ICMR Testing details dataset from Kaggle is employed, this dataset provides information about total samples tested, total individuals tested and total positive cases. We also considered another covid-19 dataset which contains statistics of confirmed, active, recovered and dead cases.

K-means clustering algorithm: K-means technique is to classify the data in to a number of clusters. We define centers for each cluster and assign each data point to the nearest cluster center. Compute the new center for each cluster and re assign data points until rearrangements.

Implementing K means clustering algorithm, an unsupervised machine learning technique to identify

the data patterns and that could be useful in making informed business decisions.

Algorithm: K-means

Step 1: obtain the data set and clean the raw data

Step 2: identify the potential features and find the optimal number of clusters required using the Elbow method.

Step 3: Train the K-means model on the dataset

Step 4: Visualize and analyze the results

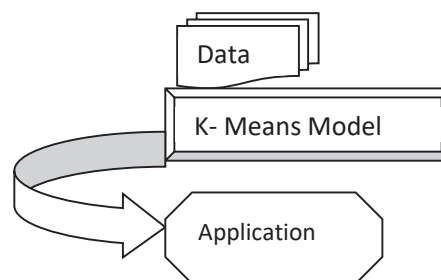


Figure 1: Block diagram of proposed methodology

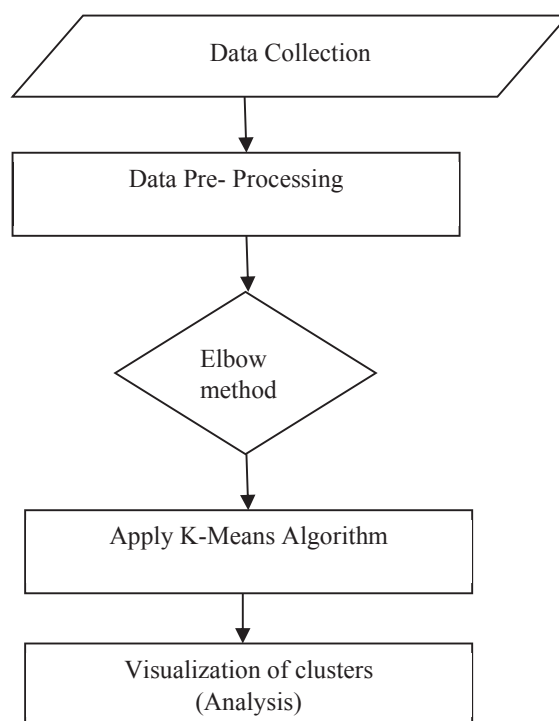


Figure 2: Flow chart

IV. Results

Elbow graph: Elbow method is used to identify the number of clusters (k) in a dataset. To obtain the number of clusters we select the value at the elbow point from the graph.

Tests vs Cases: The tests vs cases plot has 2 clusters,

Cluster 1 has less – average number of tests with average number of cases and
 Cluster 2 has average – high tests with high cases.

Cases vs Deaths: The cases vs deaths plot has 3 clusters,
 Cluster 1 has less confirmed cases with less deaths
 Cluster 2 has average confirmed cases with average deaths
 Cluster 3 has high confirmed cases with more death

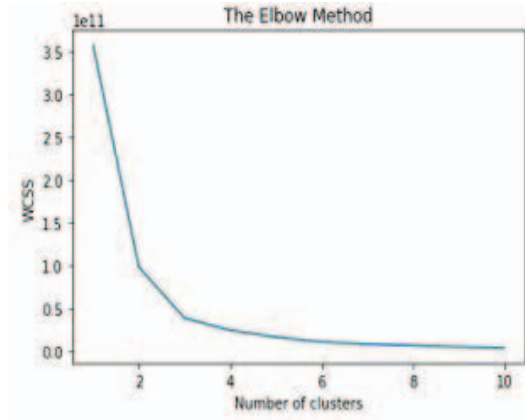


Figure 3: Cases vs Death: Elbow Graph

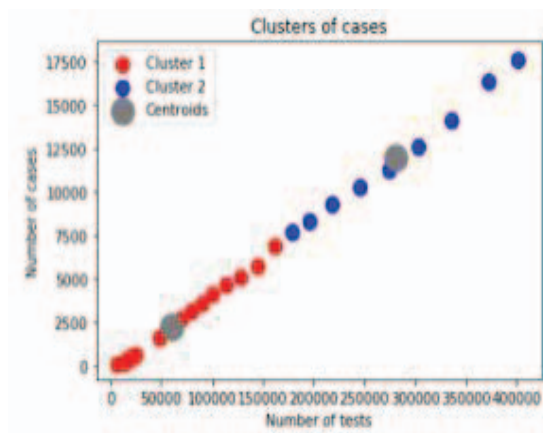


Figure 4: Tests vs cases, Clusters

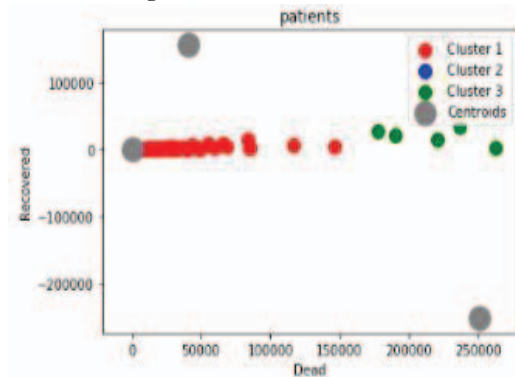


Figure 5: Active vs Dead, Clusters

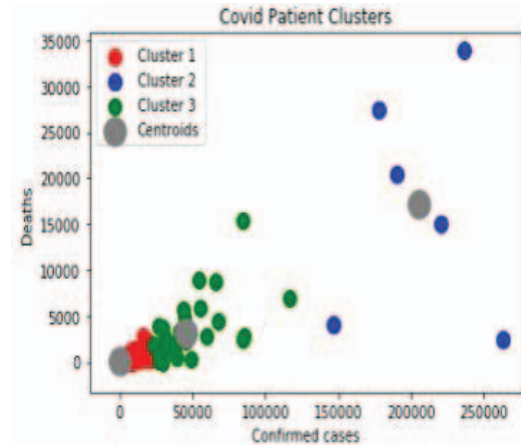


Figure 6: Confirmed Cases vs Deaths, Clusters 1, 2 &3

V. Conclusion

Covid-19 is a serious issue that has to be addressed. Many nations over the globe are working on different aspects implementing restrictions regionally to stop the virus spread. In India there is fourth phase of nationwide lock down right now. Decisions made by government becomes crucial during tough times. In this paper we made analysis using the patient statistics. Using the cluster analysis, we found that different precautions and different conditions are needed for different people. Restrictions based on zones is one good approach to follow. Old age people, people with respiratory problems are at high risk. These people may not recover if they were attacked with virus. This analysis could be further enhanced with adequate data. Testing a greater number of people will help in finding the positive cases at early stages and the data will help in further analysis.

References:

- [1] Alanagreh, Lo'ai, Foad Alzoughool, and Manar Atoum. "The human coronavirus disease COVID-19: its origin, characteristics, and insights into potential drugs and its mechanisms." *Pathogens* 9.5 (2020): 331.
- [2] Latif, Siddique, et al. "Leveraging Data Science To Combat COVID-19: A Comprehensive Review." (2020).
- [3] Soares, Eduardo, et al. "SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification." *medRxiv* (2020).
- [4] Kowalik, Maciej M., et al. "COVID-19—toward a comprehensive understanding of the disease." *Cardiology Journal* (2020).
- [5] Zhavoronkov, Alex, et al. "Potential non-covalent SARS-CoV-2 3C-like protease inhibitors designed using generative deep learning approaches and

reviewed by human medicinal chemist in virtual reality." (2020).

[6] Amiri, Abdalah Shaibu, Muhammad Akram, and M. BEMS. "COVID-19: THE CHALLENGES OF THE HUMAN LIFE."

[7] Al-Turaiki, Isra, Mona Alshahrani, and Tahani Almutairi. "Building predictive models for MERS-CoV infections using data mining techniques." *Journal of Infection and Public Health* 9.6 (2016): 744-748..

[8] Ardabili, Sina F., et al. "Covid-19 outbreak prediction with machine learning." Available at SSRN 3580188 (2020).

[9] Allam, Zaheer, Gourav Dey, and David S. Jones. "Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future Urban health policy internationally." *AI* 1.2 (2020): 156-165.

[10] Naudé, Wim. "Artificial Intelligence against COVID-19: An early review." (2020).