A Study: COVID 19 Pandemic and Role of Big Data

[1]Ms. Daivashala Deshmukh,[2]Mr. Prashant Khosre

[1]Assistant Professor, [2]Assistant Professor,
[1]Computer Science & Engineering Department,
[1]Maharashtra Institute of Technology, Aurangabad, Maharashtra.

Abstract: Big Data can give you insights regarding corona virus. The causes where and how it is spreading faster and at what rate. Big data can be used during effort in battling the COVID-19 pandemic. This should be from predicting and tracking population movements as well as tracking down individuals or groups who have contracted or have had close contact with index cases. It's a treasure trove of data for the epidemiologists to dig. The rapid spread of COVID-19 has brought together doctors, researchers and data scientists to find a solution. Scientists are using sophisticated technologies such as big data analytics, machine learning, and natural language processing for tracking the virus and learning more about it.

As lots of patient data is being stored now, it is difficult to analyze each record and determine a solution to curb the virus. This is where big data comes in. Big data is becoming a powerful tool in analyzing these datasets and identifying patterns that can help in COVID-19 detection and recovery.

IndexTerms **– BigData, Covid 19, pandemic, research, corona virus**

I. INTRODUCTION

The rapid, global spread of COVID-19 has brought advanced big data analytics tools front and center, with entities from all sectors of the healthcare industry. There are several big data components to this pandemic where artificial intelligence can play a big role. Researchers and developers are increasingly using artificial intelligence, machine learning, and natural language processing to track and contain corona virus, as well as gain a more comprehensive understanding of the disease. In the months since COVID-19 hit the US, researchers have been hard at work trying to uncover the nature of the virus why it affects some more than others, what measures can help reduce the spread, and where the disease will likely go next. Research team is going to study on information about how people are moving, the effect of travel restrictions or stay at home orders, how many people have what – that's data we can get. The more details we can get, the better, and a lot of that data is starting to be shared because you don't have to say who the people are, just where the people are. Governments have traditionally worked with health systems to monitor the number of cases as well as the location and timing of infections. Unfortunately, this is a fantastically expensive and labor-intensive process. Using social media to expand our collection of knowledge in real-time is a crucial development in a situation as fast-moving as the COVID-19 pandemic. This point is where social media enters the picture. Multiple research papers have already demonstrated how useful machine learning can be for digesting real-time social media posts and status updates. This technology can also estimate how quickly and where a communicable disease might spread next. Users of Facebook, Twitter and other websites may use language that gives insights into their current or future health. Machine learning equipped with natural language processing can study massive numbers of public messages remarkably quickly. The software can then help researchers come to conclusions about disease spread and ultimately enable public officials to make responsible decisions in-the-moment.

II. PREDICTION PARAMETERS WHERE BIG DATA CAN BE USED TO UNDERSTAND COVID 19 PANDEMIC.

a. Predicting an individual's risk of infection

Many researchers are studying about various risk factors like individual social habbit, his medical condition, age, various hygiene habit, how many times he interact with others these are some factors so that medical people can understand risk factors better and they can predict the model based on this**.** They Enter Hospitals With just a quick scan, the system determines who has a fever and who does not — and signals for further examination when the patient shows potential signs of infection. In remote or understaffed health care settings, this is vitally important for making the best use of limited staff and resources.

b. **Predicting Outcomes of Patient Treatment**

The medical community like health care has long seen big data as a important tool in predicting treatment outcomes and understanding the viability of different treatments. This strategy is a form of high-tech triage, where doctors must weigh the potential risks of a course of treatment against its likelihood to succeed. Every parameter explained above like age and pre health conditions, plays a role in these analytical models. The health care community is exploring many options for using big data and machine learning to screen patients, provide intervention quickly and allocate resources. The platform uses machine learning to predict patient outcomes based on their progression and characteristics. The ability to keep tabs on the number and condition of

patients as they enter health care facilities is crucial no matter the circumstances. In a pandemic, knowing at a glance how many patients require further screening, how many hospital beds are available and which hospital is best equipped to deal with incoming casualties or illness can save lives. These situations are why unified, shared dashboards for health systems — powered by big data have become so vital. Keeping first responders, hospital staff and dispatch professionals informed with accurate, real-time data about the availability of medical staff and resources is essential for allocating limited supplies effectively. It is also crucial for sending patients to facilities that can see them promptly and responding as quickly as possible to the most severe cases.

c. Identifying Promising Drug Candidates

When it comes to drug discovery, computers are far better than human researchers for exploring hundreds, thousands or millions of potential chemical combinations .In the search for effective medications and vaccines efforts which now include the frenzied search for a viable COVID-19 vaccination scientist rely on machine learning to predict how a virus's proteins will interact with existing or novel drugs. This process is known as drug-target interaction prediction. To do this, scientists train neural networks using vast databases of existing DTI data. This method results in a list of drugs or drug combinations that have the highest potential to bind to and inhibit the actions of virus proteins. In March 2020, a team of researchers from Tsinghua University, the Jiangsu Provincial Center for Disease Control and the Shanghai Institute of Materia Medica announced they had found a promising vaccine candidate for COVID-19. Their methodology involved a neural network trained on "knowledge graphs" that had previously been successful in finding Baricitinib, one of the earliest examples of a potential treatment candidate for COVID-19. Scientists build these knowledge graphs using natural language processing and machine learning. The algorithm combs vast medical archives to find connections between, for example, proteins and drugs — or other associations that are meaningful to the medical community.

d. Estimating Real-Time Spread and Forecasting Future Spread

The last example of scientists leveraging big data to flatten the corona virus curve is potentially the most controversial. For anybody reading between the headlines, it's clear the mechanisms under development or already being used to study the pandemic may outlive the crisis and become part of a worldwide surveillance apparatus

III. Survey of COVID-19 and Big Data and the Future of Pandemic Response

a. Detection

The big data role in COVID-19 aid starts from the first step – detection. Toronto-based big data startup BluDot detected some unusual pneumonia cases in Wuhan, China in December 2019. They did it by using their big data algorithm that pulled data from a variety of sources. The algorithm analyzed data from health records, airline ticketing data, government notices, news reports and disease networks to predict the rise of an illness. Researchers are using big data and analytics to better comprehend coronavirus from a number of different angles. The institute recently announced that it would offer government entities, research organizations, and industry access to innovative AI tools, as well as experts in data and public health to help combat COVID-19

b. Analyzing the spread of COVID-19

Using airline ticketing data, BluDot was able to predict the spread of COVID-19 from Wuhan to other Asian cities. Apart from this, mobile phone data is also used for tracking where the virus might spread. Location statistics are also useful.
For example, big data tools can analyze disease data and information about senior citizens, who are at risk of contracting corona virus. The algorithms and tools can track these people down to the postcode level, keeping in mind factors such as obesity or diabetes. These analysis reports will suggest healthcare centers and hospitals where additional medical facilities such as beds will be required. Data scientists are working on mobile applications that will be used for contact tracing. By utilizing the location data on their Smartphone, people can be alerted if they have been exposed to the virus. A team at Southern Illinois University **(SIU)** has developed a data visualization tool that uses GPS data to alert users about locations of COVID-19 cases.

c. Virus tracking dashboards

Big data role in COVID-19 aid is becoming more evident as organizations such as WHO, CDC and Microsoft are creating dashboards based on it. These dashboards pull data from different countries and show confirmed cases, deaths and locations. The dashboards can be used to prepare datasets for big data models. The models can predict possible hotspots and warn the healthcare authorities beforehand.
Another crucial big data process used against COVID-19 is outbreak analytics. This deals with the collection and analysis of outbreak response data. Data including deaths, confirmed cases, tracing people contacted by infected patients, population densities, and much more are used to develop data models for the disease. These models can predict peak infection rates and their impact.

d. Handling risks

Big data can play an important role in analyzing screening data of patients and connecting it with unnamed health issues of hospitalized patients. The results of this analysis will help in determining the key risk factors. So, as more and more data is fed into big data programs and tools, the accuracy of risk predictions will increase. Google's DeepMind artificial intelligence

system uses big data to understand the virus by analyzing its characteristics. This will help doctors to develop drugs and treatment plans.

e. Big data role in COVID-19 aid – Securing our future

After this global pandemic is resolved, big data can help the governments to prevent and battle against future outbreaks. The data from this outbreak can be used to test scenarios and analyze their outcomes to make vital decisions in the future.

IV. Conclusion

In the upcoming years, big data will play an important role in analyzing global data about detected viruses, disease modeling, tracking human activity and visualization of this data. As more and more data pile up into huge datasets, data scientists will have a better shot at preventing such outbreaks. Read more to learn how data science helps us to prevent future pandemics. AI can also help organizations draw on research from the past, applying this knowledge to present and future situations. A lot of work is going on to try to develop a vaccine to find out whether there are any current drugs work against COVID-19. All of those projects require molecular modeling, and many of them are using AI and machine learning to map things we know about the virus to things in pharmacological databases and genomic databases. Several big-name organizations have launched projects like these Amazon Web Services, Google Cloud, and others have recently offered researchers free access to open datasets and analytics tools to help them develop COVID-19 solutions faster.

V. References

[1]. https://healthitanalytics.com/news/understanding-the-covid-19-pandemic-as-a-big-data-analytics-issue

[2]. https://www.smartdatacollective.com/5-ways-big-data-is-being-used-to-understand-covid-19

[3]. https://medium.com/@shrika1994/bigdata-in-covid-19-aid-58100f84fac2

[4]. https://www.upgrad.com/blog/role-of-big-data-in-covid-19-aid/