

Review on Big Data Analysis on COVID-19

Vinay Kailash Upadhyay¹, Dr. Dinesh Dattatray Patil² and Prof. Yogesh Patil³

¹Department of Computer science & Engineering, Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal, INDIA

²Associate Professor and Head, Department of Computer science & Engineering, Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal, INDIA

³Assistant Professor, Department of Computer science & Engineering, Shri Sant Gadge Baba College of Engineering and Technology, Bhusawal, INDIA

¹Corresponding Author: vinayupadhyay158@gmail.com

ABSTRACT

Big data analytics is in transforming stage and it will continue to grow and contribute in different area like health. In recent past we have been gone through biggest health issue of the century and it taught us a lesson that in this modern world, technology with relevant data can help to us reduce any human challenge. COVID 19 is that health issue which we faced and suffered a lot. Big data provided us a platform which help us to create an accurate and most trusted data base to analyze, track and tackle this situation. Big data provides an elaborate set of attributes, details of infected patient in very explanatory.

Keywords-- Big Data Analytics, 2019 Novel Coronavirus Disease (COVID-19), Processing

storage, management, and processing tools, analytics tools and methods, and visualization and evaluation tools to the different phases of the decision-making process.

1.2 Creation of Big Data Storage and Management

One of the first things that have to manage when dealing with big data, is where and how this data will be stored once it is acquired. The traditional methods of structured data storage and retrieval include relational databases, data marts, and data warehouses.

The data is uploaded to the storage from operational data stores using Extract, Transform, Load (ETL), or Extract, Load, Transform (ELT), tools which extract the data from outside sources, transform the data to fit operational needs, and finally load the data into the database or data warehouse. Thus, the data is cleaned, transformed, and catalogued before being made available for data mining and online analytical functions [3].

1.3 Big Data Analysis Processing

There are four critical requirements for big data processing:

1. Fast data loading. Since the disk and network traffic interferes with the query executions during data loading, it is necessary to reduce the data loading time.
2. Fast query processing.
3. The third requirement for big data processing is the highly efficient utilization of storage space.
4. Fourth requirement is the strong adaptivity to highly dynamic workload patterns. As big data sets are analyzed by different applications and users, for different purposes, and in various ways, the underlying system should be highly adaptive to unexpected dynamics in data processing, and not specific to certain workload patterns.

1.4 Advantages

With detailed data capturing capability, big data can be used gainfully to minimize the risk of spreading this virus.

Big data technology can store a massive amount of information about these people infected with this COVID-19 virus. It helps in understanding the nature of

I. INTRODUCTION

Over a million of people got infected with the COVID-19 and it is still spreading. The size, variety, and rapid change of such data require a new type of big data analytics, as well as different storage and analysis methods. Such sheer amounts of big data need to be properly analyzed, and pertaining information should be extracted.

The contribution of this paper is to provide an review analysis of the available literature on big data analytics. Accordingly, some of the various big data tools, methods, and technologies which can be applied are discussed, and their applications and opportunities provided in several decision domains are portrayed.

Due to long review process of journals, most of the papers discussing covid analysis based random data. While big data analytics is being researched in academia, several of the industrial advancements and new technologies provided were mostly discussed.

1.1 Existing system

Proposed the Big – Data, Analytics, and Decisions (B-DAD) framework which incorporates the big data analytics tools and methods into the decision-making process [8]. The framework maps the different big data

this virus in detail.

Big data provides a massive amount of information to the scientists, health workers, epidemiologists and help them to make informed decision to fight with the COVID-19 virus. These data can be used to track the virus on a global basis continuously and to create innovation in the medical field

1.5 Objective

To store a large amount of data of the cases, using different data storage technologies. These data are used to undertake research and development about the virus, pandemic and measures to fight this virus and its after-effects. Big data is an innovative technology which can digitally store a large amount of data of these patients. It helps to computationally analyze to reveal patterns, trends, associations and differences. It can also help in revealing the insights into the spread and control of this virus.

II. REVIEW OF LITERATURE

2.1 Health

Research and development have leveraged advances in data science and Big Data technology to predict future events. Various studies related to virus transmission were carried out to predict:

- (a) The spread of the virus [1];
- (b) The person suspected of being infected [3]
- (c) New infection areas [4];
- (d) The likelihood of the second and third waves of the epidemic [5];
- (e) COVID-19 contamination scenario based on people movement [6];
- (f) The increased number of cases [7].

Studies aimed to find effective treatments without side effects are still ongoing in pharmacology and medicine. Analysis of chloroquine derivatives showed improving clinical outcomes and the reduction of mortality in COVID-19 patients.

Additionally, data from the National Health Insurance Service showed that patients taking medication for high blood pressure have a lower risk of exposure.



Figure 1: Virus as a Data

III. METHODOLOGY

The analysis of remote patient monitoring data can assist in estimating the number of patients in a specific area to optimally plan for containing any expected increase in the number of patients beyond the hospital capacity. Moreover, health data is growing exponentially, making it difficult to use traditional representation methods such as tables. The employment of artificial intelligence alongside data analytics tools has a role in addressing this challenge,

and it can help in the extraction and representation of data in real-time.

IV. KEY CHALLENGES

Several challenges may hinder the beneficial outcome from the application of big data analysis tools in the health sector that have been encountered when designing solutions to address the COVID-19 epidemic, which will be discussed in the following subsections.

V. CONCLUSION

The volume of data increases dramatically over time, especially data generated on the global pandemic caused by COVID-19. Such volume of data requires utilizing big data analytics tools along with AI techniques to make sense of the pandemic and control its spread in a timely manner. In this study, we presented a review of several data analysis applications for COVID-19, providing a taxonomy structure which classified the potential applications of COVID-19 into four categories, namely diagnosis, estimate or predict risk score, healthcare decisionmaking, and pharmaceutical.

REFERENCES

- [1] Abouelmehdi, K., Beni-Hssane, A., Khaloufi, H. & Saadi, M. (2017). Big data security and privacy in healthcare: A review. *Procedia Comput. Sci.*
- [2] Schmidt, B.-M., Colvin, C.J., Hohlfield, A. & Leon, N. (2020). Definitions, components and processes of data

harmonisation in healthcare: A scoping review. *BMC Med. Inform. Decis. Mak.*

- [3] Radcliffe, K., Lyson, H.C., Barr-Walker, J. & Sarkar, U. (2019). Collective intelligence in medical decision-making: A systematic scoping review. *BMC Med. Inform. Decis. Mak.*

- [4] Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., He, K., Shi, Y. & Shen, D. (2021). Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE Rev. Biomed. Eng.*

- [5] Abdel-Basst, M., Mohamed, R. & Elhoseny, M. A. (2020). Model for the effective covid-19 identification in uncertainty environment using primary symptoms and CT scans. *Heath Inform. J.*

- [6] Drew, D.A., Nguyen, L.H. & Steves, et al. (2020). Rapid implementation of mobile technology for real-time epidemiology of covid-19. *Science.*

- [7] A.K. Chanda, et al. (2017). A new framework for mining weighted periodic patterns in time series databases. *ESWA*, 79, 207-224.