



Review Article

Data Driven Models for Contact Tracing Prediction: A Systematic Review of COVID-19

Saravanan Muthaiyah ¹, Thein Oak Kyaw Zaw ^{1*}, Kalaiarasi Sonai Muthu Anbananthen ²,
Byeonghwa Park ³, Myung Joon Kim ⁴

¹ Faculty of Management, Multimedia University, Selangor, Malaysia.

² Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia.

³ Department of Management and Marketing, Valdosta State University, Georgia, United States.

⁴ Department of Business Statistics, Hannam University, Daejeon, South Korea.

Abstract

The primary objective of this research is to identify commonly used data-driven decision-making techniques for contact tracing with regards to Covid-19. The virus spread quickly at an alarming level that caused the global health community to rely on multiple methods for tracking the transmission and spread of the disease through systematic contact tracing. Predictive analytics and data-driven decision-making were critical in determining its prevalence and incidence. Articles were accessed from primarily four sources, i.e., Web of Science, Scopus, Emerald, and the Institute of Electrical and Electronics Engineers (IEEE). Retrieved articles were then analyzed in a stepwise manner by applying Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISM) that guided the authors on eligibility for inclusion. PRISM results were then evaluated and summarized for a total of 845 articles, but only 38 of them were selected as eligible. Logistic regression and SIR models ranked first (11.36%) for supervised learning. 90% of the articles indicated supervised learning methods that were useful for prediction. The most common specialty in healthcare specialties was infectious illness (36%). This was followed closely by epidemiology (35%). Tools such as Python and SPSS (Statistical Package for Social Sciences) were also popular, resulting in 25% and 16.67%, respectively.

Keywords:

Coronavirus;
Data Driven Decision Making;
Prediction;
Pandemic;
Covid-19;
SARS-CoV-2.

Article History:

Received:	14	May	2022
Revised:	24	July	2022
Accepted:	15	August	2022
Published:	24	September	2022

1- Introduction

Pandemics have always been a source of concern for medical practitioners globally. The potential danger of major infectious diseases spreading across the globe before an adequate response can be made is a contentious topic. The wide spread of the pandemic has a significant impact on any country's healthcare system. The world has experienced this during the SARS, H1N1 and SARS-CoV-2 outbreaks. The toll it takes on contingency policies is indeed significant [1]. The Coronavirus disease (COVID-19) caused by the SARS-CoV-2 virus was the most recent in a long series of virus-related diseases that have significantly impacted the world. Coronavirus disease (COVID-19) was first detected on December 8, 2019, in the city of Wuhan, China [2]. Given the fact that Coronavirus (nCoV) is a variant of the Coronavirus that has not been observed yet, containing the outbreak was nearly impossible [3, 4].

* **CONTACT:** rahman93.iu@gmail.com

DOI: <http://dx.doi.org/10.28991/ESJ-2023-SPER-02>

© 2023 by the authors. Licensee ESJ, Italy. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Lack of understanding, coupled with the absence of a cure, pushed experts to work tirelessly to develop novel therapies and vaccinations for infected patients. Despite this, data scientists started working on producing data-driven decision models and algorithms to understand how transmissible the virus was and how it could be traced effectively. In general, contact tracing is the process of identifying close contacts of an infected person, whether they are symptomatic or asymptomatic. In this case, the Centers for Disease Control (CDC) and Prevention have given two main guidelines to be followed for close contact identification and they are: (1) Distance requirement of less than 2-feet and a minimum duration of 15 minutes [5, 6]. Since the beginning of the digital health revolution [7, 8], a massive amount of health-related data has been acquired and processed through efficient data modelling techniques for better predictive analysis. Nonetheless, the accuracy of these models was heavily dependent on the available training data sets [9, 10]. Researchers are constantly developing algorithms to better understand and improve control as well as administer solutions for pandemic crisis through numerous data-driven decision-making techniques. Considering the critical nature of the COVID-19 pandemic, surveying the right and competent data-driven decision making methods will help to narrow down appropriate and plausible methods for better combating the disease. The results will also assist us in determining unascertainable characteristics of the new virus and the subsequent development throughout the pandemic. The aim of this paper was to gather, summarize, and analyze articles published in the context of data-driven decision-making applications, particularly for contact tracing applications. The following are the specific research propositions: (RP1) is to identify data sources for contact tracing solutions around the globe; (RP2) is to understand the number of studies that addressed issues with regards to the COVID-19 outbreak up to the end of 2021; (RP3) investigates the usage of data-driven decision making methods and tools used for contact tracing; (RP4) identifies resources of information and data; and lastly (RP5) determines the most used techniques and tools in terms of frequency. The rest of this paper is organized as follows: The second section provides methods; the third section reports results; the fourth section presents discussion of the results; and the fifth section provides the conclusion of the study.

2- Methods

2-1-Data Sources for Contact Tracing

Identifying the data sources for COVID-19 is crucial as it gives an overall picture where the information originates from. In return, it gives an idea on how hard or easy it is to conduct data driven decision making in order to do predictions. Furthermore, studies on a selected 30-countries from Legender et al. [11] and Rahman [12] in Table 1 shows an overview of contact tracing solutions from these countries while Figure 1 illustrates world map for the countries studied. In general, what the data shows is that most of the contact tracing solutions require their information to be locked and stored in the mobile phones in which, it requires special permission from the users. It also illustrates that most of the latest solutions have been using BLE-based technology that captures mainly time, duration and IDs of other users, however, does not record the medical condition of the patient. In return, data driven decision making for COVID-19 becomes much harder even if there is more information being captured on the applications as the data is being stored in mobile phones. Therefore, data driven decision making for COVID-19 tracing prediction will need to heavily rely on clinical information from healthcare centers.

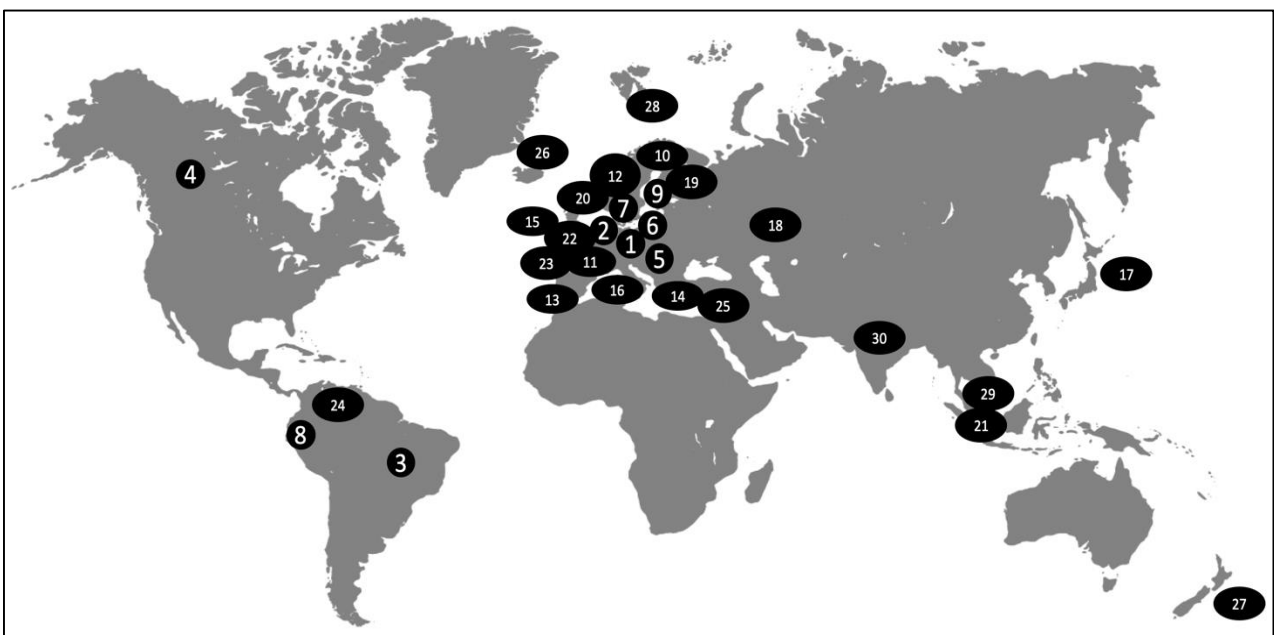


Figure 1. World Map for 30-Countries Studied

Table 1. Contact Tracing Solution Information by 30-selected Countries

No.	Country	Data Source	Application Name	Technology Used	Symptomatic Tracing	Asymptomatic Tracing
1.	Austria		Stopp Corona			
2.	Belgium		Coronaalert			
3.	Brazil		Coronavirus - SUS			
4.	Canada		COVID Alert			
5.	Croatia		Stop COVID-19			
6.	Czech Republic		eRouška			
7.	Denmark		Smittestop			
8.	Ecuador		ASI			
9.	Estonia		Hoia			
10.	Finland		Koronavilkku			
11.	France		TousAntiCovid	BLE		
12.	Germany		Corona-Warn App			
13.	Gibraltar		Beat-Covid Gibraltar			
14.	Greece	Phone storage	Exo		Yes	Yes
15.	Ireland		Covid Tracker			
16.	Italy		Immuni App			
17.	Japan		COCOA – COVID – 10 Contact App			
18.	Kazakhstan		Saqbol			
19.	Latvia		Apturi Covid Latvia			
20.	Lithuania		Korona Stop LT			
21.	Singapore		TraceTogethe	BLE and tokens		
22.	Switzerland		SwissCovid	BLE		
23.	UK		NHS Covid-19 app	BLE		
24.	Uruguay		Coronavirus UY	BLE		
25.	Israel		Hamagen	Global Positioning System (GPS)		
26.	Iceland		Rakning C19	BLE and GPS		
27.	New Zealand		NZ Covid Tracer App	BLE and QR Code		
28.	Norway		Smittestopp	BLE (previously GPS)		
29.	Malaysia	Ministry of Health	MySejahtera	QR code		No
			MyTrace	BLE		Yes
			GerakMalaysia	GPS		No
30.	India	Indian Council of Medical Research (ICMR)	Aarogya Setu App	Bluetooth and GPS		Yes

2-2-Literature Review

To ensure the presence of relevant papers, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria checklist was adopted [13]. Subsequently, to categorize the primary features of studies, a synthesis of relevant literature was undertaken concerning the main characteristics. From 2020 to August 16, 2021, literature search was conducted on four main sources i.e., Web of Science, Scopus, Emerald, and IEEE. Key terms used were "data driven decision making", "prediction", "techniques", "COVID-19" and "Coronavirus". Search terms were also used to create a Boolean search algorithm which was then iteratively applied on every database source.

2-3-Enclosure and Elimination Criteria for Research Selection

Articles were incorporated if they matched the subsequent requirements: (a) the studies or assessments with regards to COVID-19 and (b) papers that discussed the application of data driven decision making techniques or data exploration methods. These types of approaches were chosen based on methods described by Patel and Patel [13]. Due to the vast diversity of methods available in this field, studies reviewed were only limited by papers that were available in the English language. Papers were considered ineligible if they matched any of the following requirements: (I) the paper doesn't refer to any topics on pandemics or COVID-19 and this includes the title, abstract and also the references, (II)

other than papers and conference publications other materials such as book chapters, letters, short abstract, briefing comments were excluded, (III) papers in any other language other than English and (IV) image processing techniques were also excluded. Lastly, (V) the full text documents that were unavailable were also excluded.

2-4-Data Extraction Phase

A total of 390 articles were retrieved from searching the databases i.e. IEEE, Science Direct, Scopus, and Emerald. The entire process is highlighted by the PRISMA approach illustrated in Figure 2. Before beginning search, few inclusions and exclusion criterias were applied for the purpose of filtering papers that were relevant. In the first step, headings and summaries of retrieved papers were reviewed to identify significance for inclusion. In second step screening was applied to remove any duplication. This step consists of a comprehensive screening process. During this stage, the entire texts of relevant studies have been carefully reviewed and filtered. In the third step full text eligibility is determined and after complete analysis, the decision of selecting papers that are relevant is completed. After completing steps (I to IV) there were 38 manuscripts that still remained eligible and relevant. An extraction form was created to keep track of the articles during the reviewing process. Generic information and specific information were both included in this classification system. General information comprises of author names, publication date and publisher. The specific information covers the primary purpose of research; Data driven decision making approaches, application of the Data driven decision making method and tools, disciplinary, primary outcomes, assessment results, data resources, sample sizes, software used, and the study's nation place. The attributes of the articles included were extracted and compared to the previously defined classification. All the information that was retrieved, was re-examined to obtain an agreement.

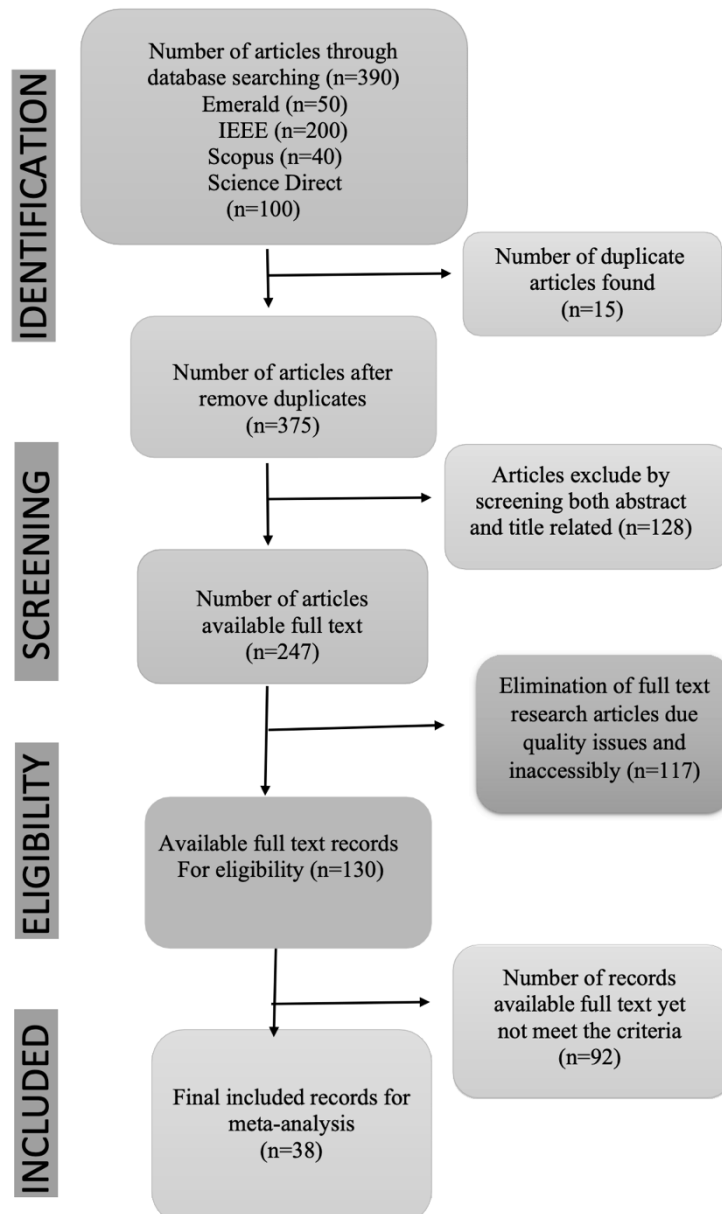


Figure 2. Systematic Literature Review Process (PRISMA)

3- Results

An initial search of scientific databases resulted in 390 citations, which was significant. At the first screening stage, three articles were deleted from consideration due to duplication. Following that, 128 publications were removed since their abstracts and titles were deemed irrelevant by the reviewers. The second screening stage, defined as the complete test screening stage, resulted in the removal of 117 articles due to their inaccessibility and poor overall quality. In the final filtering of full-text articles, it is crucial to distinguish between those relevant to the issue and those that are not, such as medical, anatomy, economic impact, and others. As a result, 38 papers were recognised as being potentially connected.

3-1- Study Characteristics

There were twenty eight journal publications and ten conference papers that fulfilled our eligibility requirements and criteria. Table 2 shows the distribution of studies by year, with the most recent studies appearing first. As can be seen, most of the studies were published in the year 2020.

Table 2. Articles Published According to Year

Year	Number of Articles Published
2020	21
2021	17

3-2- The Distribution of Papers by Countries

Overall, the selected research article is distributed by 17 different nations, which is a large number when broken down by country (Table 3). In addition, the essay makes use of global statistics on the sickness pandemic that has swept the world. Compared to other countries, India has the highest frequency of occurrences.

Table 3. Characteristics of Paper Based on Country Origin

Countries	Number of Articles
India	9
China	5
USA	5
Saudi Arabia	3
Bangladesh	2
Indonesia	1
Iran	1
Iraq	1
Italy	1
Jordan	1
Kuwait	1
Morocco	1
Poland	1
Portugal	1
South Africa	1
Taiwan	1
Turkey	1

3-3- The Usage of Data Driven Decision Methods Assessed through Articles

Following the findings in research conducted by Patel & Patel [13], frequency of methods used in predicting pandemics has been investigated. The primary goal of this paper is to determine the extent of usage of data driven decision making techniques in predicting outbreak or number of cases during pandemics. Table 4 presents a summary on usage of appropriate data driven decision making methods in the assessed publications, based on the number of articles analysed. According to the findings, methodologies have been categorised into 22 major categories. The decision tree was the most often employed technique among the examined publications (13.64%). It was ranked top in effectiveness of determining the relationship between independent factors and a single dichotomous dependent variable [14]. One important thing to note is the researcher's usage of multiple types of data driven decision making techniques during their investigation.

Table 4. The Rate of Occurrence of Data Driven Decision Making Methods in the Selected Articles

DD Methods	Rate of Occurrence	%	Sources
Logistic Regression	10	11.36%	[15-23, 54]
Sir Model	10	11.36%	[4, 8, 20, 29, 31-36]
Decision Tree	9	10.23%	[17, 20-21, 24-28, 54]
Support Vector Machine (SVM)	9	10.23%	[2, 4, 20, 23-25, 27-28, 54]
K-Nearest-Neighbour	7	7.95%	[13, 22-24, 28, 29, 54]
Random Forest	7	7.95%	[19-21, 24, 28, 30, 54]
Naïve Bayes	6	6.82%	[20, 22, 24, 28, 52, 54]
Neural Network	6	6.82%	[9, 25, 29, 35, 36, 54]
LSTM	3	3.41%	[8, 30, 34]
Time Series	3	3.41%	[32-33, 37]
(CNN)	3	3.41%	[2, 4, 23]
GIS	2	2.27%	[4, 38]
XGBoost	2	2.27%	[19, 35]
Regression models	2	2.27%	[7, 42]
RNN	2	2.27%	[8, 34]
Fuzzy	1	1.14%	[39]
k-means algorithms	1	1.14%	[37]
NetNCSP	1	1.14%	[40]
Polynomial Regression	1	1.14%	[41]
SVR	1	1.14%	[30]
(TWC)	1	1.14%	[43]
Wang-Mendel Neuro-Fuzzy Rule Inference	1	1.14%	[3]

3-4- The Distribution of Data Driven Decision Making Software in Reviewed Articles

Data driven decision making approaches necessitate the use of specialised tools and an appropriate platform (Table 5). Review on the frequency of various devices employed in this research has been listed in this section. Among other tools, Python had the most significant percentage (25%), followed by SPSS with six publications (16.67%) and WEKA five articles (13.88%). Others that had a small percentage were Matlab, Orange, R, Fuzzy, SQL and 2 other privately owned softwares. Eight other articles from the total publications (22.22%) did not specify the tools used.

Table 5. Tools used in Data Driven Decision Making

Techniques	Frequency
Python	9
Not mentioned	8
SPSS	6
WEKA	5
Orange	2
R Software	2
Fuzzy	2
GLS	1
Matlab	1
SQL	1
Hybrid Undefined	1

4- Discussion

The principal purpose of this research is to offer a complete overview about data-driven data driven decision making methods in managing viruses' transmission. After screening of 337 articles, 38 publications were selected and analysed.

This study describes the study's findings and conclusions. In terms of country, most of the research has been performed in India. This could be justified by the reality that large number of outbreaks happened in India. Social media is an additional source of information in the modern era, capable of creating substantially more data in a shorter period than traditional sources [2]. Accessing this type of data is convenient in comparison to other traditional data in terms of information tracking. The qualitative assessment finds out that researchers tend to use organised approaches such as regression to build analytical models to understand unknown pandemics and diseases better. All these tactics have been successfully implemented in a variety of medical sectors with outstanding results [44]. Additionally, research indicates that categorization algorithms are being deployed at a higher rate than expected. Researchers can determine the optimal approach for implementing accurate prediction models for unknown diseases. These models may then be used to forecast significant outcomes [45, 46]. Following that, developing predictive models can benefit physicians, health policymakers, and society. Since most studies were conducted in India, these models may require further testing. None of the research, however, recommended the implementation of developed models in real-world settings. On the other perspective, many authors expressed confidence regarding the development of predictive models. Maakoul et al. [18] perspective on forecasting model construction is compatible with the research findings. Multiple researchers, including [46-48], performed a systematic review of COVID-19 analytical models. They determined that the projected models lack sufficient documentation and are hence prone to prejudice. The data indicated that the fundamental objective of pandemic sickness is to halt the spread of infectious diseases [31]. When a new disease forms as part of a pandemic, the disease's nature is often unknown, and scientists' primary focus is determining the disease's characteristics. As a result, the great majority of research is devoted to elucidating the disease's characteristics. This is explained by the fact that scientists should concentrate diagnosis over other tasks during a pandemic sickness outbreak [49]. The second critical factor to consider in pandemic diseases is the virus's transmission. As a result, around 10% of studies have focused on forecasting the disease's development. On the other hand, the test size of datasets is quite varied and diversified, owing to usage of varies methodologies. The findings indicated that most of the research uses a variety of data resources but small in sample size (Table 6).

Table 6. Overview on Reviewed Articles

No.	Source	Key Approaches	Medical Scope	Sample Size	Data Source
1.	[2]	Prevention and Risk Management	Biological	Daily cases and death growth	WHO
2.	[3]	Prevention and Risk Management	Social Management	Creation of Model	Not Applicable
3.	[4]	Machine learning	Social Management	n=80,735 cases	National Health Commission of the People's Republic of China.
4.	[7]	Info veillance	Infectious Disease	n=100,000 (Hubei Province)	WHO
5.	[8]	Info veillance	Social Behaviour	n=507	Northeast University
6.	[9]	Outbreak prediction	Epidemiology	literature review	WHO
7.	[24]	Disease Characteristics	Infectious Disease	literature review	Literature Review
8.	[13]	Patient Monitoring	Epidemiology	NA	NA
9.	[15]	Disease Characteristics	Respiratory Machine	n=315 questionnaire, 3375 images	Social Media and www.kaggle.com
10.	[17]	Risk factors	Social Management	Probability Sampling Technic n=14,667	Portuguese Health Ministry
11.	[18]	Infoveillance	Social Behaviour	n=80	Own data set
12.	[19]	Prevention and Risk Management	Machine Learning	n=5644	Hospital Israelita Albert Einstein, Brazil
13.	[20]	Early Diagnosis	Virology	literature review	WHO
14.	[21]	Outbreak Prediction	Epidemiology	n=5434	Korean Centre for Disease Control
15.	[22]	Infoveillance	Machine Learning	Not Mentioned	kawalCovid19.com and kawalcorona.com
16.	[23]	Tracking and Prevention Management	Image processing	n=349 Covid 19 CT scan, 463 non covid 19 CT scan)	WHO
17.	[25]	Tracking and Prevention Management	Infectious Disease	n=3799	Indian Ministry of Health Website
18.	[26]	Tracking and Prevention Management	Machine Learning	n=2342	Ministry of Health and Family Welfare, India
19.	[27]	Info veillance	Social Behaviour	Literature	N/a

20.	[28]	Tracking and Prevention Management	Infectious Disease	Not Mentioned	www.kaggle.com
21.	[29]	Outbreak prediction	Epidemiology	n=290	Iraqi Hospital
22.	[30]	Active case Prediction	Epidemiology	n=30 days	Institute of Epidemiology, Disease Control and Research, Bangladesh
23.	[32]	Tracking and Prevention Management	Infectious Disease	n= 100 days before and after lockdown	WHO
24.	[33]	Outbreak prediction	Infectious Disease	n=10 top countries	Johns Hopkins University Centre for System Science and Engineering
25.	[34]	Prediction	Engineering	n=300	own Experiment using 3 sensor Temperature devices
26.	[35]	Outbreak prediction	Public Health	Not Mentioned	Johns Hopkins University Centre for System Science and Engineering
27.	[37]	Prevention and Risk Management	Infoveillance	n=513	Global
28.	[38]	Outbreak prediction	Epidemiology	literature review	US Government
29.	[39]	Outbreak prediction	Social Behaviour	n=124 days	Turkey's Republic Health Ministry
30.	[40]	Prevention and Risk Management	Machine Learning	Not Mentioned	Not Mentioned
31.	[47]	Emergency Management	Machine Learning	NA	NA
32.	[52]	Predict Outbreak	Machine Learning	n= top 10 active cases countries	WHO
33.	[41]	Tracing Transmission	Epidemiology	50 US state	Covid tracking Project - 50 US, US Census Bureau
34.	[42]	Medical Imaging	Respiratory Machine	n=6995	2 hospital in China
35.	[43]	Info veillance	Epidemiology	n=48	Geographic Profiling Approach
36.	[44]	Outbreak prediction	Respiratory Specimen	n=3711	Ministry of Health, Labour and Welfare of Japan
37.	[50]	Outbreak prediction	Epidemiology	Not Mentioned	www.kaggle.com
38.	[51]	Tracing Transmission	Infectious Disease	Not Mentioned	Own Data
39.	[53]	Prediction	Machine Learning	n=600 medical reports	www.aha.org
40.	[54]	Risk factors	Ontology	n=71	Italian Society of Medical and Intervention Radiology (SIRM)Repository

The usage of enormous data collections can boost the intensity of the findings and the precision of the model's expectations, which can aid experts in their efforts to better understand and combat this emerging illness. As a result, researchers are encouraged to employ massive datasets for their studies, even if they are conducted on a global scale, to make better diagnostic and beneficial recommendations. In terms of pandemic lead diseases, most of the activities are concentrated under the COVID-19 umbrella. Those who came in second place discussed issues that were relevant to influenza pandemics. Given the high prevalence of these two disorders, it is reasonable to expect this outcome. The use and retrieval of huge amounts of data provided by automated systems as a main resource makes it easier and more convenient to access data. As an outcome, recently performed data-driven studies are more expedient than previously. Most of the papers excluded past pandemics such as Asian flu and H1N1 since the authors of these papers did not consider these diseases to be epidemics. Furthermore, there were certain drawbacks to this investigation, which have been discovered.

In today's world, studies on COVID-19 are published daily. The entire screening and filtering looked at the literature between the years 2020 and 2021. As a result, certain studies may overlook the period between publication and the publication of this article. Therefore, additional studies will be required to complete the findings. Another drawback of this study is that only four journal databases were used as electronic search engines. Due to the access restriction, the other databases that may have quality academic journals should be integrated in future studies. The present research paper provides researchers with a valuable framework for upcoming work by allowing them to grasp the overall framework of data-driven decision-making performances in pandemics and for their submissions to recognize the disease better. Yet to come, research might include the construction of search algorithms in larger datasets or the examination of data-driven decision-making applications as a broader idea. An article on the analysis and incorporation of non-English written papers using machine translation techniques may be adopted in the near future. At the very least, it would be fascinating to incorporate non-English-based papers.

5- Conclusion

This research is expected to make it easier for academics to identify published papers on data-driven decision making techniques and the Covid-19 pandemic. The PRISMA method for systematic review and meta analysis was useful for inclusion and exclusion criteria, thus allowing researchers to evaluate the quality of papers to be included or otherwise for the study. Healthcare related decisions, especially during a pandemic, should be based on well-informed and readily available research evidence. Systematic research provides evidence-based medicine that introduces evidence from ongoing contact tracing methods that are mostly effective. PRISMA provides such evidence with good reviews. All data-driven decision-making techniques utilised in global pandemics are listed throughout this paper. However, the bulk of these techniques were developed to avoid and forecast the COVID-19 pandemic during its current phase. According to the study findings, the primary goal of data-driven decision-making applications is to increase the characteristics of the condition being treated. Medical practitioners don't really have the time to synthesize the huge number of articles being written in this space. As such, the value this paper brings to the healthcare community is to identify, summarize, and evaluate studies that will make available evidence more accessible to healthcare practitioners. Additionally, it can aid politicians and decision-makers in making more informed choices about the identification and management of severe pandemics in the countries where they operate.

6- Declarations

6-1-Author Contributions

Conceptualization, S.M. and T.O.K.Z.; methodology, S.M., T.O.K.Z. and K.S.M.A.; validation, S.M. and T.O.K.Z.; formal analysis, T.O.K.Z, B.P. and M.J.K.; investigation, S.M., T.O.K.Z. and B.P.; writing—original draft preparation, S.M., B.P. and M.J.K.; writing—review and editing, S.M. and M.J.K.; supervision, S.M. and K.S.M.A.; project administration, K.S.M.A. All authors have read and agreed to the published version of the manuscript.

6-2-Data Availability Statement

Data sharing is not applicable to this article.

6-3-Funding and Acknowledgements

This study is supported by Faculty of Management (FOM), Multimedia University, Malaysia.

6-4-Institutional Review Board Statement

Not applicable.

6-5-Informed Consent Statement

Not applicable.

6-6-Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

- [1] Jain, V., Duse, A., & Bausch, D. G. (2018). Planning for large epidemics and pandemics: Challenges from a policy perspective. *Current Opinion in Infectious Diseases*, 31(4), 316–324. doi:10.1097/QCO.0000000000000462.
- [2] Meo, S. A., Al-Khlaiwi, T., Usmani, A. M., Meo, A. S., Klonoff, D. C., & Hoang, T. D. (2020). Biological and epidemiological trends in the prevalence and mortality due to outbreaks of novel coronavirus COVID-19. *Journal of King Saud University - Science*, 32(4), 2495–2499. doi:10.1016/j.jksus.2020.04.004.
- [3] Wang, R., Hu, G., Jiang, C., Lu, H., & Zhang, Y. (2020). Data Analytics for the COVID-19 Epidemic. 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC). doi:10.1109/compsac48688.2020.00-83.
- [4] Zhou, C., Su, F., Pei, T., Zhang, A., Du, Y., Luo, B., Cao, Z., Wang, J., Yuan, W., Zhu, Y., Song, C., Chen, J., Xu, J., Li, F., Ma, T., Jiang, L., Yan, F., Yi, J., Hu, Y., ... Xiao, H. (2020). COVID-19: Challenges to GIS with Big Data. *Geography and Sustainability*, 1(1), 77–87. doi:10.1016/j.geosus.2020.03.005.
- [5] Zaw, T. O. K., Muthaiyah, S., Anbananthen, K. S. M., & Soe, M. T. (2022). A Novel Approach on Covid-19 Contact Tracing – Utilization of Low Calibrated Transmission Power & Signal Captures in BLE. *Emerging Science Journal*, 6, 181–192. doi:10.28991/esj-2022-sper-013.

- [6] Centers for Disease Control and Prevention (CDC). (2022). Contact Tracing for COVID-19. Centers for Disease Control and Prevention, U.S. Department of Health & Human Services, Washington, United States. Available online: <https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/contact-tracing.html> (accessed on July 2022).
- [7] Salzberger, B., Glück, T., & Ehrenstein, B. (2020). Successful containment of COVID-19: the WHO-Report on the COVID-19 outbreak in China. *Infection*, 48(2), 151–153. doi:10.1007/s15010-020-01409-4.
- [8] Sun, K., Chen, J., & Viboud, C. (2020). Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *The Lancet Digital Health*, 2(4), e201–e208. doi:10.1016/S2589-7500(20)30026-1.
- [9] Heymann, D. L., & Shindo, N. (2020). COVID-19: what is next for public health? *The Lancet*, 395(10224), 542–545. doi:10.1016/S0140-6736(20)30374-3.
- [10] Gulyaeva, M., Huettmann, F., Shestopalov, A., Okamatsu, M., Matsuno, K., Chu, D. H., Sakoda, Y., Glushchenko, A., Milton, E., & Bortz, E. (2020). Data mining and model-predicting a global disease reservoir for low-pathogenic Avian Influenza (A) in the wider pacific rim using big data sets. *Scientific Reports*, 10(1), 1–11. doi:10.1038/s41598-020-73664-2.
- [11] Legendre, F., Humbert, M., Mermoud, A., & Lenders, V. (2020). Contact Tracing: An Overview of Technologies and Cyber Risks. *Cryptography and Security*. doi:10.48550/arXiv.2007.02806.
- [12] Rahman, M. (2021). List of countries using Google and Apple's COVID-19 contact tracing API. XDA Developers. Available online: <https://www.xda-developers.com/google-apple-covid-19-contact-tracing-exposure-notifications-api-app-list-countries/> (accessed on April 2022).
- [13] Patel, S., & Patel, H. (2016). Survey of Data Mining Techniques used in Healthcare Domain. *International Journal of Information Sciences and Techniques*, 6(1/2), 53–60. doi:10.5121/ijist.2016.6206.
- [14] Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593.
- [15] Alodat, M. (2021). Using deep learning model for adapting and managing covid-19 pandemic crisis. *Procedia Computer Science*, 184, 558–564. doi:10.1016/j.procs.2021.03.070.
- [16] Rai, P. K., Sonne, C., Song, H., & Kim, K.-H. (2022). The effects of COVID-19 transmission on environmental sustainability and human health: Paving the way to ensure its sustainable management. *Science of The Total Environment*, 838, 156039. <https://doi.org/10.1016/j.scitotenv.2022.156039>.
- [17] Ferreira, A. T., Fernandes, C., Vieira, J., & Portela, F. (2021). Pervasive intelligent models to predict the outcome of COVID-19 patients. *Future Internet*, 13(4). doi:10.3390/fi13040102.
- [18] Maakoul, O., Boucht, S., EL HACHIMI, K., & Azzouzi, S. (2020). Towards Evaluating the COVID'19 related Fake News Problem: Case of Morocco. 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS). doi:10.1109/icecocs50124.2020.9314517.
- [19] Podder, P., & Mondal, M. R. H. (2020). Machine Learning to Predict COVID-19 and ICU Requirement. 2020 11th International Conference on Electrical and Computer Engineering (ICECE). doi:10.1109/icece51571.2020.9393123.
- [20] Rakhra, A., Jain, I., Gupta, R., & Bhatia, M. (2021). Predicting the Prevalence Rate of COVID-19 Falsity on Temperature. 11th International Conference on Cloud Computing, Data Science & Engineering. doi:10.1109/confluence51648.2021.9377198.
- [21] Rao, V. C. S., Gampa, S., Rama, V., Anumala, H., & Gadepally, A. (2021). Extracting Insights and Prognosis of Corona Disease. 2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP2021). doi:10.1109/icccsp52374.2021.9465516.
- [22] Romadhon, M. R., & Kurniawan, F. (2021). A Comparison of Naive Bayes Methods, Logistic Regression and KNN for Predicting Healing of Covid-19 Patients in Indonesia. 2021 3rd East Indonesia Conference on Computer and Information Technology (EIconCIT). doi:10.1109/eiconcit50028.2021.9431845.
- [23] Tomar, M., Arora, T., Sabitha, A. S., & Hasteer, N. (2021). Comparative Study of Deep Learning Models for COVID-19 Diagnosis. 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC). doi:10.1109/icsccc51823.2021.9478139.
- [24] Alballa, N., & Al-Turaiki, I. (2021). Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. *Informatics in Medicine Unlocked*, 24, 100564. doi:10.1016/j.imu.2021.100564.
- [25] Alsmadi, T., Alqudah, N., & Najadat, H. (2021). Prediction of Covid-19 patients states using Data mining techniques. 2021 International Conference on Information Technology (ICIT). doi:10.1109/icit52682.2021.9491716.
- [26] Gupta, V. K., Gupta, A., Kumar, D., & Sardana, A. (2021). Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model. *Big Data Mining and Analytics*, 4(2), 116–123. doi:10.26599/BDMA.2020.9020016.

- [27] Tripathi, A., Kaur, P., & Suresh, S. (2021). AI in Fighting Covid-19: Pandemic Management. *Procedia Computer Science*, 185, 380–386. doi:10.1016/j.procs.2021.05.039.
- [28] Villavicencio, C. N., Macrohon, J. J. E., Inbaraj, X. A., Jeng, J. H., & Hsieh, J. G. (2021). Covid-19 prediction applying supervised machine learning algorithms with comparative analysis using weka. *Algorithms*, 14(7). doi:10.3390/a14070201.
- [29] Jaleel, R. A., Burhan, I. M., & Jalookh, A. M. (2021). A Proposed Model for Prediction of COVID-19 Depend on K-Nearest Neighbors Classifier:Iraq Case Study. 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE). doi:10.1109/icecce52056.2021.9514171.
- [30] Mohammad Masum, A. K., Khushbu, S. A., Keya, M., Abujar, S., & Hossain, S. A. (2020). COVID-19 in Bangladesh: A Deeper Outlook into the Forecast with Prediction of Upcoming per Day Cases Using Time Series. *Procedia Computer Science*, 178, 291–300. doi:10.1016/j.procs.2020.11.031.
- [31] Adhikari, S. P., Meng, S., Wu, Y.-J., Mao, Y.-P., Ye, R.-X., Wang, Q.-Z., Sun, C., Sylvia, S., Rozelle, S., Raat, H., & Zhou, H. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infectious Diseases of Poverty*, 9(1). doi:10.1186/s40249-020-00646-x.
- [32] Burhanuddin, A., & Kurniawan, F. (2021). Analysis of the Spread of COVID-19 in Local Areas in Indonesia. 2021 3rd East Indonesia Conference on Computer and Information Technology (EICoCIT). doi:10.1109/eiconcit50028.2021.9431906.
- [33] Khan, S., & Alfaifi, A. (2020). Modeling of coronavirus behavior to predict it's spread. *International Journal of Advanced Computer Science and Applications*, 11(5), 394–399. doi:10.14569/IJACSA.2020.0110552.
- [34] S, M., V.R, N., P.S, R., S, R. R. M., & L, N. (2020). Pervasive computing in the context of COVID-19 prediction with AI-based algorithms. *International Journal of Pervasive Computing and Communications*, 16(5), 477–487. doi:10.1108/IJPC-07-2020-0082.
- [35] Vadyala, S. R., Betgeri, S. N., Sherer, E. A., & Amritphale, A. (2021). Prediction of the number of COVID-19 confirmed cases based on K-means-LSTM. *Array*, 11, 100085. doi:10.1016/j.array.2021.100085.
- [36] Albastaki, A., Naji, M., Lootah, R., Almeheiri, R., Almulla, H., Almarri, I., Alreyami, A., Aden, A., & Alghafri, R. (2021). First confirmed detection of SARS-COV-2 in untreated municipal and aircraft wastewater in Dubai, UAE: The use of wastewater based epidemiology as an early warning tool to monitor the prevalence of COVID-19. *Science of The Total Environment*, 760, 143350. <https://doi.org/10.1016/j.scitotenv.2020.143350>.
- [37] Plessen, M. G. (2020). Integrated Time Series Summarization and Prediction Algorithm and its Application to COVID-19 Data Mining. 2020 IEEE International Conference on Big Data (Big Data). doi:10.1109/bigdata50022.2020.9377823.
- [38] Cook, A. H., & Cohen, D. B. (2008). Pandemic disease: A past and future challenge to governance in the United States. *Review of Policy Research*, 25(5), 449–471. doi:10.1111/j.1541-1338.2008.00346.x.
- [39] Cihan, P. (2020). Fuzzy Rule-Based System for Predicting Daily Case in COVID-19 Outbreak. 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). doi:10.1109/ismsit50672.2020.9254714.
- [40] Wu, Y., Zhu, C., Li, Y., Guo, L., & Wu, X. (2020). NetNCSP: Nonoverlapping closed sequential pattern mining. *Knowledge-Based Systems*, 196, 105812. doi:10.1016/j.knosys.2020.105812.
- [41] Zhan, F. (2021). COVID-19 Case Number Prediction Utilizing Dynamic Clustering With Polynomial Regression. 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC). doi:10.1109/ccwc51732.2021.9375966.
- [42] Elbasi, E., Mathew, S., Topcu, A. E., & Abdelbaki, W. (2021). A Survey on Machine Learning and Internet of Things for COVID-19. 2021 IEEE World AI IoT Congress (AIIoT). doi:10.1109/aiiot52608.2021.9454241.
- [43] Buscema, P. M., Della Torre, F., Breda, M., Massini, G., & Grossi, E. (2020). COVID-19 in Italy and extreme data mining. *Physica A: Statistical Mechanics and Its Applications*, 557, 124991. doi:10.1016/j.physa.2020.124991.
- [44] Zhang, Y., Guo, S. L., Han, L. N., & Li, T. L. (2016). Application and exploration of big data mining in clinical medicine. *Chinese Medical Journal*, 129(6), 731–738. doi:10.4103/0366-6999.178019.
- [45] Alanazi, H. O., Abdullah, A. H., Qureshi, K. N., & Ismail, A. S. (2018). Accurate and dynamic predictive model for better prediction in medicine and healthcare. *Irish Journal of Medical Science*, 187(2), 501–513. doi:10.1007/s11845-017-1655-3.
- [46] Nithya, B., & Ilango, V. (2017). Predictive analytics in health care using machine learning tools and techniques. 2017 International Conference on Intelligent Computing and Control Systems (ICICCS). doi:10.1109/iccons.2017.8250771.
- [47] Asadzadeh, A., Pakkhou, S., Saeidabad, M. M., Khezri, H., & Ferdousi, R. (2020). Information technology in emergency management of COVID-19 outbreak. *Informatics in Medicine Unlocked*, 21, 100475. doi:10.1016/j.imu.2020.100475.
- [48] Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., ... van Smeden, M. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ*, m1328. doi:10.1136/bmj.m1328.

- [49] Kelly-Cirino, C. D., Nkengasong, J., Kettler, H., Tongio, I., Gay-Andrieu, F., Escadafal, C., Piot, P., Peeling, R. W., Gadde, R., & Boehme, C. (2019). Importance of diagnostics in epidemic and pandemic preparedness. *BMJ Global Health*, 4, 1–8. doi:10.1136/bmjgh-2018-001179.
- [50] Kumari, R., Kumar, S., Poonia, R. C., Singh, V., Raja, L., Bhatnagar, V., & Agarwal, P. (2021). Analysis and predictions of spread, recovery, and death caused by COVID-19 in India. *Big Data Mining and Analytics*, 4(2), 65–75. doi:10.26599/BDMA.2020.9020013.
- [51] M, T., M, G., S.R, S., M, A., R, K., & P.S, R. (2020). Detecting coronavirus contact using internet of things. *International Journal of Pervasive Computing and Communications*, 16(5), 447–456. doi:10.1108/IJPCC-07-2020-0074.
- [52] Bhatia, S., & Malhotra, J. (2021). Naïve Bayes Classifier for Predicting the Novel Coronavirus. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). doi:10.1109/icicv50876.2021.9388410.
- [53] Estiri, H., Strasser, Z. H., & Murphy, S. N. (2021). Individualized prediction of COVID-19 adverse outcomes with MLHO. *Scientific Reports*, 11(1). doi:10.1038/s41598-021-84781-x.
- [54] Oyelade, O. N., & Ezugwu, A. E. (2020). A case-based reasoning framework for early detection and diagnosis of novel coronavirus. *Informatics in Medicine Unlocked*, 20(May), 100395. doi:10.1016/j.imu.2020.100395.