Studies in Systems, Decision and Control 358

Diego Oliva
Said Ali Hassan
Ali Mohamed *Editors*

Artificial Intelligence for COVID-19

Springer

Studies in Systems, Decision and Control

Volume 358

The series "Studies in Systems, Decision and Control" (SSDC) covers both new developments and advances, as well as the state of the art, in the various areas of broadly perceived systems, decision making and control–quickly, up to date and with a high quality. The intent is to cover the theory, applications, and perspectives on the state of the art and future developments relevant to systems, decision making, control, complex processes and related areas, as embedded in the fields of engineering, computer science, physics, economics, social and life sciences, as well as the paradigms and methodologies behind them. The series contains monographs, textbooks, lecture notes and edited volumes in systems, decision making and control spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

Indexed by SCOPUS, DBLP, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/13304

Diego Oliva · Said Ali Hassan · Ali Mohamed
Editors

Artificial Intelligence for COVID-19

Springer

Editors
Diego Oliva
Department of Computer Sciences
University of Guadalajara
Guadalajara, Jalisco, Mexico

Said Ali Hassan
Operations Research and Decision Support Department, Faculty of Computers and Artificial Intelligence
Cairo University
Giza, Egypt

Ali Mohamed
Operations Research Department, Faculty of Graduate Studies for Statistical Research
Cairo University
Giza, Egypt

ISSN 2198-4182 ISSN 2198-4190 (electronic)
Studies in Systems, Decision and Control
ISBN 978-3-030-69743-3 ISBN 978-3-030-69744-0 (eBook)
https://doi.org/10.1007/978-3-030-69744-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Artificial Intelligence (AI) is an area of science that emphasizes intelligent machines that work and behave like humans. Machines can often act and react like humans only if they have much information relating to the world. AI has become an essential part of technology and represents incredibly exciting and powerful techniques. It is used to solve many real-world problems of growing interdisciplinary interest and practical importance.

This book contains several different types of articles that provide access to the state of AI art. It covers all major areas of the concepts, methods, tools and leading AI results in the domain of recent developments. The information contained in the book explores different areas of AI; machine and deep learning; advanced image processing; computational intelligence; IoT; robotics and automation; optimization; mathematical modelling; neural networks; information technology; big data, data processing; data mining and likewise. We intend to cover the breadth and depth of AI, presenting general overviews of the scientific issues and detailed discussions of techniques and essential system material in AI for real-world applications. The book's overarching purpose is to provide a collected and connected set of reflections about AI and its influences as the field advance. The participants represent diverse specialties, geographic regions and career stages. The included chapters develop theory, algorithms, programs, policies and systems design to ensure that these systems can inspire innovations and provide intelligent advice to government agencies and non-government organizations to combat against COVID-19. The book is a long-term investigation of AI's field and its influences on defending against the COVID-19 pandemic and broadly benefit individuals and society.

The book features a wealth of real applications, case studies and illustrations for COVID-19 emerging pandemic along with the theoretical concepts, algorithms and procedures. With this diversity and interdisciplinary subject matter, we are trying to build bridges that scientists can cross in other fields. These applications cover a wide variety of different specialties that include many diverse directions in the field of combating against COVID-19 like: relation between numbers of handwashing facilities and COVID-19 deaths; big data analysis for COVID-19; data interpretation of COVID-19 cases using R programming; AI applications to COVID-19; ANN and fuzzy classifier to diagnose COVID-19; QoL barometer for well-being during pandemics; deep and machine learning models for COVID-19; using X-ray images; information and communication technology application in COVID-19 pandemic; robotics and automation for COVID era; COVID-19 machine learning forecasting models; optimization of the international trade activities in COVID-19 period; IoT applications for COVID-19; scheduling disinfection process for COVID-19; forecasting of COVID-19 using SIR; modified SEIR; logistic growth and Holt's models; text mining for COVID-19 analysis; mathematical modelling the spread of COVID-19; information technology in health emergency control; big data in COVID-19 assistance; proactive detection of people infected by COVID-19; online learning problems during COVID-19 period; modelling of population density effect on the infectiousness and mortality rates of COVID-19; AI; chat-bots; solitude; people and a therapeutical blending; prevention guidelines based on cognitively inspired AI and data mining for COVID-19; AI for diagnosis and treatments for COVID-19; factors affecting medical mask purchase decision and Harris hawks optimization for classification of COVID-19 gene.

The book is useful as a valuable reference for AI theoretical and practical professionals because of its comprehensive coverage and a large number of detailed algorithms. It is intended primarily for researchers, decision-makers and practitioners in AI and many application disciplines. It will also be of interest to graduate level for rich case studies and projects.

The editors hope that this book will help the readers share their fascination with Artificial Intelligence and applications. They expect that the proactive guidance stemming from the chapters will have broad global relevance and are making plans for future contributions to expand the scope of the studies.

Finally, we wish to acknowledge the book authors who have made distinguished and fruitful efforts to complete the various chapters of the book. The book contents will enrich the world library in the vital field of Artificial Intelligence with all its multiple specializations and applications.

Guadalajara, Mexico — Diego Oliva
Giza, Egypt — Said Ali Hassan
Giza, Egypt — Ali Wagdy Mohamed
2021

Contents

Simulation of the Relation Between the Number of COVID-19 Death Cases as a Result of the Number of Handwashing Facilities by Using Artificial Intelligence

Sajedeh Norozpour

Abstract Artificial intelligence has recently been one of the most commonly discussed topics of discussion in the world. The need for efficient and reliable energy has been a growing fact in many sectors of the sciences. More importantly, the need for predictive analysis has become essential for most of the forecasting variables. On the other hand, there has been a notable relationship and overlaps between artificial intelligence and simulation due to the decrease in the use of hardware, while the sophistication of the software application increases. Algorithms used in simulation also allow expert systems to think in the direction and line of complex models and systems, which may always change dynamically or that may have an inclusion of the stochastic processes. The following paper begins with an introduction of artificial intelligence and its relationship to modeling and simulation. The study has involved a critical review of different pieces of literature on the value of artificial intelligence in modeling and simulation and also finds out the model of the relationship between the COVID-19 death rate and the number of handwashing materials.

Keywords Artificial intelligence · Simulation and modeling · COVID-19 · Death rate · Handwashing materials

1 Literature Review

The general meaning of AI is the ability of a computing machine to perform activities that would otherwise require human reasoning and decision-making, such as speech recognition, visual distinction, and performing decision-making processes and analytics [1]. However, this definition is inherently misrepresented as to what is logical, and the discussion has always been open to debate. By that definition, an internal controller of the house, such as a thermostat, can be considered intelligent because it can sense and change the temperature [2]. It is not the same as artificial intelligence, with which a drone selects and designs approaches without

S. Norozpour (✉)
Department of Civil Engineering, Istanbul Gelisim University, Istanbul, Turkey
e-mail: snorozpour@gelisim.edu.tr

D. Oliva et al. (eds.), *Artificial Intelligence for COVID-19*, Studies in Systems, Decision and Control 358, https://doi.org/10.1007/978-3-030-69744-0_1

significant intervention from humans, which is usually perceived when it comes to autonomous ordnances system. Additionally, another key factor to consider when talking about autonomous ordinances is the growing inability to liberate independence from the commercial autonomy of military drones. With the growth of the business sector to develop autonomous land and air systems, there is no doubt that there is evidence of a shift in the military's AI capability to the business. Therefore, restricting autonomous innovation for military purposes may not be realistic, as it can also be commercially available in the lower end of the market for consumers. Also, the divergent improvement in the market of the autonomous corporate structure would result in governments and armies that would no longer dominate, resulting in unsafe and dangerous AI systems, both semi-autonomous and autonomous.

One concept that has usually been interlinked with AI is drone swarming, a form of modeling and simulation. The technology behind drone swarming allows drones to make decisions autonomously based on shared decisions from other drones [3]. This has proven to be promising as it can revolutionize drone applications in military warfare. Countries such as China have succeeded in developing drone swarm technology. On the other hand, the US has been testing a ‘ Low-Cost Unmanned aerial vehicle Swarming Technology (LOCUST), a drone swarming program that employs AI on several drones to help in data gathering in the battlefield [3]. Swarms can revolutionize e military as they are applicable in every area of national and homeland security. For example, the marine forces can use swarms of drones to search the ocean for enemy submarines. Drones can also be deployed in large areas to help identify enemy hideouts and help eliminate their air defense systems. Furthermore, drone swarms can act as a shield by helping intercept incoming hypersonic missiles. Furthermore, the drones can be equipped with chemical, biological, radiological, and nuclear (CBRN) sensors, which can help identify and devise solutions to such threats.

The fate of artificial intelligence in combat scenarios is primarily linked to specialists' capacity to develop systems that have an autonomous capability of using knowledge and logical reasoning based on the data being fed [4]. Currently, there are no such systems in operation. Most drones on-site are controlled remotely, which means that a human continues to legitimately control a drone after a certain distance, for example, via an additional virtual link such as a satellite feed. Most unscrewed military aircraft develop just a little, and gradually they gain semi-autonomy, which means that they can navigate and land by themselves without any human input [5]. However, almost all of them require some degree of human input to complete their combat operations. Even drones that take off and fly to a specific target to take pictures and then go back to base still operate on an automated system but cannot be compared to some of the operations of an autonomous let alone a semi-autonomous drone.

Although the current systems are more programmed to be automatic rather than autonomous, more research continues to be carried out within the scope of autonomy. In many countries, airplanes, land, water, and underwater vehicles have gradually improved their military capabilities, with varying results [5]. In the United States, certain types of independent helicopters are under development. A sales representative can operate the aircraft with a cell phone and are currently under development by countries such as the USA, Europe, and China. Autonomous drones are under rapid research and development across the world. However, the offices that make these systems still find it difficult to make a qualitative leap in operational execution [6]. There are many reasons for the failure to develop these innovations, including unexpected costs and unforeseen technical problems. However, authoritarian and social boundaries are also barriers to the implementation of autonomous UAS. The United States, for example, has struggled to launch autonomous drones into operation mainly because of the hierarchical struggle and the prioritization of crewed aircraft [7].

For some soldiers, the drone systems are only suitable for rescue operations, but they put the rule at risk as long as they are qualified to perform the most recognized and advanced tasks. However, other hierarchical problems restrict the operational use of autonomous aircraft, and an increasingly dangerous problem is the commercialization of AI drone technology for businesses. However, cargo delivery services like UPS, FedEx, and DHL are looking into military drone technology in the delivery of cargo [8]. An allegorical arms competition takes place in the company's group to improve autonomous systems. The improvement in autonomous military systems was, at best, moderate and constant and did not reflect the progress made in independent commercial contexts, such as driverless vehicles and, in particular, drones.

Several roboticists and military experts have argued that autonomous drones should not be regarded as a threat on ethical grounds but should be a more preferred option than human fighters. According to roboticist Ronald C. Arkin, future autonomous drones will have the ability to make a better judgment on the battlefield based on the fact that they do not need to be fed with a self-preservation program, which is more common in humans. In this case, the drones will have the ability to make ethical decisions avoiding the need to "shoot first and ask questions later" scenarios [9]. Furthermore, autonomous drones will not be plagued by human emotions, such as fear in a combat scenario. The systems will be able to handle a lot of information being fed compared to humans and thus improve some of the battlefield outcomes.

There are many positive aspects of expanding AI's combat work. Focusing on AI skills increases AI's ability to verify authentic or wrong approaches decisively [5]. Intelligent drones can be designed to reduce the number of acts of war against individuals, e.g., For example, the massacre of civilians or prisoners of war, the destruction of civilian property by non-military personnel, the torture of prisoners, and devastation while limiting collateral damage due to instances of helplessness and confrontation [4]. Artificial intelligence will be more competent than humans and modern combat strategies and methods. Errors are also greatly reduced in the war zone. It is currently possible to accurately calculate and see the degree of destruction of an attack, which will now result in minor injuries due to an unpredictable attack [10]. The additional violation of human rights in the types of aggression, agony, etc. is eliminated with the method at the right time because AI is the predominant method or strategy of combat.

Drone swarms offer customization and flexibility to military commanders allowing them to add or remove drones in the formation when necessary. However, this requires a new form of inter-drone communication protocol that allows for drones to be easily added or removed from a swarm. As such, the drones should easily accommodate drones added or removed from the swarm either from the command or from hostile action [11]. The adaptability of intelligent drone swarms means that the drones can easily adapt to the needs of the situation and alleviate the commanders from small tasks such as managing the drones. The customization of drone swarms means that the size of the swarm can be adjusted depending on the mission.

Furthermore, military commanders can easily customize the drones with sensors, weapons, or other payloads depending on the needs of the mission. Drone swarms can also easily break into small groups or merge as a unit, which can easily change battlefield dynamics [11]. For example, a drone swarm can break into small groups to conduct reconnaissance in a particular area. If they spot an adversary, they can easily merge to eliminate the target.

2 Aim of the Research

Global concerns about the value of artificial intelligence have been raised about modeling and simulation. When simulation and modeling have been used, one of the critical areas is military and security areas. Moreover, previous studies have also added the value of National Language Processing (NLP) as a form of simulation. Modeling has been applied in artificial intelligence to provide a link between human language and computer language. Currently, there has been an increase in the use of Unmanned Aircraft Systems (UASs) in military and commercial environments has led to a heated discussion on the ban on "killer robots" [12]. These autonomous machines, which can operate on the ground or in the air, hypothetically when combined with "artificial intelligence" (AI), can autonomously execute missions. These give rise to the underlying idea of the use of artificial intelligence in simulation and modeling. The debate, which involves many stakeholders and scopes, raises a serious topic of whether AI systems should be allowed to carry out these combat operations, especially when there is a likelihood that there will be human casualties. Considering the complex issue at hand, the functional importance of AI is necessary. However, how artificial intelligence relates to simulation must be discussed deeply.

The value of artificial intelligence ranges from a variety of departments: from security departments to defense departments. The research aims at knowing the importance of artificial intelligence in simulation and modeling. Besides, this article aims to include how artificial intelligence is used in simulation and other modeling programs such as drone swan. The paper also aims at identifying the benefits and applications of artificial intelligence in other governmental and economic departments. For instance, artificial intelligence may be used to model future armed conflicts and mitigations against such conflicts among nations.

2.1 Statement of the Problem

Many studies have been conducted on the future of Information and Technology and programming in the world. This has come up with many issues concerning the future growth of artificial intelligence. Currently, most of the economic sectors in many countries are being run by programs based on artificial intelligence. Moreover, many studies have also been carried out on the importance and negative effects of artificial intelligence. As the world grapples with COVID-19, every ounce of technological innovation and ingenuity harnessed to fight this pandemic brings us one step closer to overcoming it. Artificial intelligence (AI) and machine learning are playing a key role in better understanding and addressing the COVID-19 crisis. Machine learning technology enables computers to mimic human intelligence and ingest large volumes of data to quickly identify patterns and insights.

In the fight against COVID-19, organizations have been quick to apply their machine learning expertise in several areas: scaling customer communications, understanding how COVID-19 spreads, and speeding up research and treatment.

However, this study decided to explore the link between artificial intelligence and modeling and simulation. Previously, little studies had been conducted in that area, and this study finds it as an opportunity to explore such an area. The case example used for this study is the model illustration of the COVID-19 deaths due to the number of handwashing facilities.

2.2 Research Methodology

This part of the study includes a series of methodological approaches for the study. The study on artificial intelligence and its importance on simulation and modeling follows a secondary study model. The study relies on secondary sources of data and follows a qualitative model of analysis. The main source of data for this study is the previous scholarly articles on artificial intelligence. Besides, the study also follows a quantitative approach of study to model and simulate the relationship between the numbers of COVID-19 death cases as a result of the number of handwashing facilities.

2.3 Research Setting and Research Paradigm

The availability of scientific data has allowed many researchers to do further studies on the relationship that might exists between artificial intelligence and simulation and modeling. Both the quantitative and qualitative research paradigm has been implemented for the data collection.

2.4 Limitations of the Method

The method used for this study was based on a qualitative approach to study using secondary sources of data. The fundamental impediment to the use and reliance of secondary data is that the accuracy of the secondary data used may not be fully known because of no accuracy with the information provided. Besides, some of the secondary sources of data might be too old for use in data analysis, and at the same time, the secondary sources may fail to fit the research model and framework. This is because the area of artificial intelligence is dynamic. Besides, the study has also used a quantitative research approach.

3 Results and Discussion

In this study, we wish to know how artificial intelligence may be applied to determine the future number of COVID-19 death cases due to the number of handwashing facilities. In this way, the study wishes to apply simulation and modeling in finding the relationship. The dataset used for the study has been sourced from Coronavirus Pandemic (COVID-19) data. For the model, the dependent variable is the number of COVID-19 deaths, while the independent variable is the number of handwashing facilities. The sample data used is represented in Table 1. (Data retrieved from https://ourworldindata.org/coronavirus-data)

Using the data in Table 1, the study developed a model, which would also be essential for simulation purposes. The model is summarized as follows;

Table 1 Sample coronavirus pandemic (COVID-19) Data

Death rate	Handwashing facilities	Death rate	Handwashing facilities
597.029	37.746	344.094	47.964
278.364	83.741	137.973	83.841
276.045	26.664	303.74	19.351
341.01	94.043	190.968	85.198
559.812	83.241	318.949	4.472
298.003	34.808	266.653	55.182
170.05	88.469	140.448	80.635
176.957	90.083	525.432	89.827
235.848	11.035	167.295	90.65
217.066	79.807	202.812	24.64
204.299	25.383	182.634	7.96
329.635	97.164	331.43	7.876
269.048	11.877	298.245	41.047
293.068	6.144	155.898	76.665
270.892	66.229	336.717	17.45
244.661	2.735	382.474	6.403
435.727	16.603	373.159	77.159
280.995	5.818	430.548	22.863
124.24	65.386	240.208	84.169
261.516	15.574		

Summary output								
Regression statistics								
Multiple R	0.138792							
R Square	0.019263							
Adjusted R Square	−0.00724							
Standard error	112.2751							
Observations	39							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	9161.054	9161.054	0.726739	0.399427			
Residual	37	466411	12605.7					
Total	38	475572.1						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	308.9056	31.33775	9.857299	6.78E-12	245.4093	372.4019	245.4093	372.4019
X Variable 1	-0.45735	0.536492	−0.85249	0.399427	−1.54439	0.629682	−1.54439	0.629682

The model can, therefore, be represented as:

Y = 308.9056 – 0.45735X

which implies that, the rate of COVID-19 fatalities = 308.9056 – 0.45735 (number of handwashing facilities).

The results show that there is a negative relationship between the two variables. I.e., the rate of COVID-19 fatalities will increase with a decrease in the number of handwashing facilities. Given the regression model above, then we can simulate it. The simulation results will show the minimum value of the dependent variable, the first quartile, the median value, mean, third quartile, and the maximum value of the dependent variable. Given the independent variable data, we have the following:

Mean = 47.84

Standard deviation = 33.95

Suppose we simulate from the model;

Y = 308.9056 – 0.45735X + ε; assume ε is approximately normally distributed with mean 0 and X ~ N (47.84;33.95), $\beta_0 = 308.9056$ and $\beta_1 = -0.45735$.

We have the following R output:

```
> set.seed (39)
> x < −rnorm (39, 47.84, 33.95)
> e < −rnorm (39, 0, 2)
> y < −308.9056-0.45735*x + e
> Summary(y)
```

The summary would output the minimum value of the dependent variable, the first quartile, the median value, mean, third quartile, and maximum value. As such, it would be easy to model the rate of fatalities of COVID-19 due to the number of handwashing facilities using artificial intelligence techniques.

4 Conclusion

The idea of artificial intelligence is a critical technological tool in current society. The study has explored the relationship between artificial intelligence and modeling and simulation by exploring the COVID-19 data. Through the study, it is evident that artificial intelligence can be used to model and simulate the relationship between two or more variables. The study has employed both qualitative and quantitative approaches to find the results. More specifically, the study has used a quantitative approach to find out the relationship between the COVID-19 death rate and the number of handwashing materials in different countries. The relationship between the two variables has been represented in a model. Therefore, the data model is essential in simulating the future expectations of the dependent variable (COVID-19 death rates). Therefore, it is true that artificial intelligence has a great interdependence with simulation and modeling of future outcomes and data.

References

1. Stuart, R., Peter, N.: Artificial Intelligence—A Modern Approach 3rd edn. (2016)
2. Rodriguez, C., Robinson, R.: Enhancing Dengue Fever Modeling Through a Multi-scale Analysis Framework–A Case Study in the Central Valley of Costa Rica (2013)
3. Allison, J., Amako, K., Apostolakis, J., Arce, P., Asai, M., Aso, T., Bagli, E., et al.: Recent developments in Geant4. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 835, pp. 186–225 (2016)
4. Kibble, R.: Can an unmanned drone be a moral agent? Ethics and accountability in military robotics. In: Gunkel, D.J., Bryson, J.J., Torrance, S. (eds.), The machine question: AI, ethics and moral responsibility (Proceedings of symposium "Machine Question: AI, Ethics, and Moral Responsibility"AISB/IACAP 2012), pp. 62–67 (2012)
5. Allen, G., Chan, T.: Artificial Intelligence and National Security. Belfer Center for Science and International Affairs, Cambridge, MA (2017)
6. Harwood, R.: The Challenges To Developing Fully Autonomous Drone Technology. Ansys, Com (2019)
7. Spinetta, L., Cummings, M.L.: Unloved Aerial Vehicles: Gutting its UAV Plan, the Air Force Sets a Course for Irrelevance (2012)
8. James, W.: Holding Aloft the Banner of Ethiopia: Caribbean Radicalism in Early-Twentieth Century America. Verso Books (2020)
9. Arkin, Ronald C.: The case for ethical autonomy in unmanned systems. J. Military Ethics **9**(4), 332–341 (2010)
10. Cass, Kelly: Autonomous weapons and accountability: seeking solutions in the law of war. Loy. LAL Rev. **48**, 1017 (2014)
11. Kallenborn, Z.: The Era Of The Drone Swarm Is Coming, And We Need To Be Ready For It - Modern War Institute." Modern War Institute, 2020. https://mwi.usma.edu/era-drone-swarm-coming-need-ready/
12. Human Rights Watch: Arms: New Campaign To Stop Killer Robots. Human Rights Watch. https://www.hrw.org/news/2013/04/23/arms-new-campaign-stop-killer-robots (2013)
13. Ackerman, S.: Pentagon: A Human Will Always Decide When A Robot Kills You. WIRED (2020)
14. Lin, L., Hou, Z.: Combat COVID-19 with artificial intelligence and big data. J Travel Med 27(5) (2020):taaa080
15. Kumar, A., Gupta, P.K., Srivastava, A.: A review of modern technologies for tackling COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews (2020)

Big Data and Data Analytics for an Enhanced COVID-19 Epidemic Management

Joseph Bamidele Awotunde, Rasheed Gbenga Jimoh, Idowu Dauda Oladipo, Muyideen Abdulraheem, Taibat Bolanle Jimoh, and Gbemisola Janet Ajamu

Abstract The transmission features of the COVID-19 epidemic have still not been properly recognized in the sense of global environmental changes. Likewise, the pace of global development, increased population size, nuanced encounters, and lack of medical security in developing countries all contribute to the complexities of COVID-19 prevention and regulation. The COVID-19 seems to have a greater impact on the global economy when compared to severe acute respiratory syndrome (SARS) that occurs in 2013. The bone of contention is how to fight the spread of these infectious diseases and how to contains it spread from human-to-human. Most countries depended on the observation of classic steps to monitor diseases and public health to contain the COVID-19 pandemic, similar to those used with SARS in 2003. Data in the healthcare sector are usually huge and not easy to handle, this results from the enormous way by which data grows in the healthcare sector, the rate at which data are been produced, and the variety of various data in the healthcare system contributed to this growth. Also, the rate at which these data are been captured, stored, analyzed, and retrieved in healthcare has swiftly from the aged paper-based storage technique to the use of digital techniques and methods. Hence, the application of advanced

J. B. Awotunde (✉) · M. Abdulraheem
Department of Computer Science, University of Ilorin, Ilorin, Nigeria
e-mail: awotunde.jb@unilorin.edu.ng

M. Abdulraheem
e-mail: muyideen@unilorin.edu.ng

R. G. Jimoh · I. D. Oladipo · T. B. Jimoh
National Examinations Council, Zonal Office, Ilorin, Nigeria
e-mail: jimoh_rasheed@unilorin.edu.ng

I. D. Oladipo
e-mail: odidowu@unilorin.edu.ng

T. B. Jimoh
e-mail: bolanlejimoh80@gmail.com

G. J. Ajamu
Department of Agricultural Extension and Rural Development, Landmark University, Omu Aran, Nigeria
e-mail: ajamu.gbemisola@lmu.edu.ng

D. Oliva et al. (eds.), *Artificial Intelligence for COVID-19*, Studies in Systems, Decision and Control 358, https://doi.org/10.1007/978-3-030-69744-0_2

technologies which includes virtualization and cloud computing allows for huge and effective data processing in the healthcare system. Thereby rapidly turning the healthcare system into a big data industry. Nevertheless, in these modern days, the improvement in the information and communication technologies (ICT) brings the advancement of varied data from new sources in the healthcare system. Therefore, this chapter discusses the possible application of big data and analytics to improve conventional public health approaches to tackle the COVID-19 pandemic to control, manage, identify and avoid COVID-19 and reduce the effect on health implicitly associated with COVID-19. The significant applications of big data and challenges are presented. Leveraging big data and intelligent analytics becomes expedient in other to enjoy their use for COVID-19 disease outbreak and public health.

Keywords Big data · Coronavirus · Big data analytics · COVID-19 pandemic · Healthcare · Diagnosis

1 Introduction

The scientific research centered on generating new data through the performing of basic experiments to answer specific questions related to any infectious diseases. The collection and storing of a huge amount of data digitally are growing exponentially. This may require patients to receive care at various hospitals within a region, thus underrepresent or over-represent key clinical feature studies single-site. The coronavirus disease 2019 (COVID-19) epidemic poses a growing threat to the world. We are more connected than ever before, our cities are increasingly densely populated and in recent years the world has witnessed successive waves of the new severe acute respiratory syndrome (SARS), Ebola virus, Zika virus and recently COVID-19 and influenza infectious diseases threats global pandemics.

The coronavirus (COVID-19) pandemic is playing thoughtful havoc in socio-communal systems, humanity, and creates economic crises worldwide [1]. Many strategies have been used to managed and curtailed the COVID-19 outbreak, but many countries are still helpless in fighting and containing the outbreak. The epidemiology complexity and the rapidly evolving disease patterns of the COVID-19 demand a couturier approach by government, researchers, and medical experts in dealing with this overwhelming pandemic. Similarly, other methods such as social distancing, the use of facemask, isolated of confirmed patients among others have developed in crucial and manage the COVID-19 pandemic.

The pandemic originated from China in a town called Wuhan, but it has spread to 213 other countries globally. As of August 28, 2020, the World Health Organization (WHO) has reported a total of 24,926,312 COVID-19 cases with 840,662 deaths, and 17,799,210 patients recovered successfully from this deadly disease. This has been confirmed by WHO that COVID-19 spread and transmitted from wild animals illegally sold in the Huanan Seafood Wholesale Market in Wuhan [2]. According to the recent report, ages between 30 and 79 years, approximately 86.60% are susceptible to COVID-19 pandemic of all patients infected so far, thus have a median age of 47 years [3].

The men are noted to exhibit a higher propensity for this outbreak, though researches have not reported significant gender preponderance [4]. Human-to-human has been identified as the major mode of transmission of COVID-19 pandemic, and the National Health Commission ofChina, has identified droplets, direct exposure, feces, respiratory aspirates, and aerosols transmission has a carrier of COVID-19 pandemic [5]. Although research has not provided concrete evidence about sporadic reports about the vertical transmission of SARS-CoV-2 it has been shown that these also can be a possible way of spread COVID-19 [6, 7].

The data science analysis is newly evolving, intending to empower health care systems and organizations to connect to harness information and convert it to usable knowledge and preferably personalized clinical decision making. The application of big data in the area of infectious diseases has introduced several changes in the information accumulation models using analytics. The processing, accumulation, and accrual of big data in healthcare are possible through the rapidly evolving field of Big Data Analytics (BDA), thus offers the capacity to understand, rationalize and use big data to serve different purposes.

Improvement in information technology and data computing has greatly changed researches on population-based by encouraging easy access to a huge amount of data. Sometimes, such database links are referred to as “big data” [8, 9]. In other to make efficient use of these data for researches in clinical health or public health, the researchers need to widen researches further than the traditional surveillance model, as operating with big data differs from focusing on performing narrow analysis, treatment-oriented clinical data. Therefore, leveraging on Big data to reflect accurately on the heterogeneous population it represents becomes expedient [10, 11]. This endeavor needs a swift research environment that can adopt a quick advancement in computing technology to at all-time combine data while making use of new methods to reduce their complexity [9, 10]. Also, trends and patterns which make it easy to diagnose and treat patients are been revealed by big data.

The Mutual collaboration that exists in the healthcare sector results in the generation of a huge amount of data from various sources [10, 12, 13]. This has a result of huge figures of varied sections and departments including physicians of different disciplines, ranging from the nurses, the pathologists, radiologists, and laboratory technologists which collaborates efforts to achieve unified goals towards bringing medical costs and mistakes down to the barest minimum. And also to provide excellent and standardized healthcare services. Various stakeholders that work collaboratively in the health sector get data from different sources. This source includes data gotten from patient’s observation, imaging reports from scans, interview and test results, insurance, and bills, summaries of patient discharges, reports from pharmacists, case notes from physicians, admission notes of the hospitals, feedback from social media and journals of medical articles.

Consequently, the application of advanced technologies which includes virtualization and cloud computing allows for huge and effective data processing in the healthcare system. Thereby rapidly turning the healthcare system into a big data industry. Nevertheless, in these modern days, the improvement in the information and communication technologies (ICT) brings the advancement of varied data from new sources in the healthcare system. The 21st century is an age of big data affecting every aspect of human life, including biology and medicine [14, 15]. The move from paper medical records to Electronic Health Record (EHR) systems has resulted in an unprecedented increase in data [16]. Big data thus offers a great opportunity for doctors, epidemiologists, and specialists in health policy to make evidence-driven decisions that will eventually enhance patient care [17, 18]. "Big data is not only a modern reality for the biomedical scientist but an imperative that needs to be fully grasped and used in the search for new knowledge" [19, 20].

Therefore, this chapter discusses the possible application of big data and BDA to improve conventional public health approaches to tackle the COVID-19 pandemic to control, manage, and identify infected patients, thus reduce the effect on health implicitly associated with the outbreak. The remaining part of this chapter is organized thus: Sector 2 discusses the growth of data in healthcare, the challenges, and the importance of big data in COVID-19. Section 3 presents the big data privacy and ethical challenges in COVID-19. Section 4 discusses BDA in the COVID-19 epidemic. Finally, Sect. 5 concludes the chapter.

2 Big Data and Big Data Analytics for COVID-19

The era of big data impacting all parts of people's lifespan, including genetics and medicine in this 21st century [14, 21]. The transition from paper patient reports to Electronic Health Record (EHR) systems resulted in exponential data development [22]. Big data thus creates a tremendous incentive for physicians, epidemiologists, and public policy experts to take evidence-driven action that will inevitably strengthen patient safety [23, 24]. There are various technologies like IoT devices and sensors in the modern age that contribute greatly to the massive growth of Big Data. Current advances in the fields of computation, storage, and connectivity have created vast data sets, extracting valuable information from this immense volume of data would bring value to science, government, industry, and society.

Through embracing biosensors, wearable devices, and mHealth, the volume of biological data collected has expanded [25]. According to a study by Cisco [26, 27], the total amount of data produced by IoT sensors and devices will hit 847 ZB (Zettabytes) by 2021. The vast volume of data cannot be processed, collected, and controlled with conventional methods. BDA is the method of analyzing broad data sets comprising a range of data types [28] to uncover invisible patterns, latent associations, industry dynamics, consumer desires, and other important business knowledge [29]. The ability to evaluate vast volumes of data can help a company cope with important knowledge that can impact the corporation [30].

The principal aim of BDA is, therefore, to assist industry organizations in strengthening their interpretation of data and thereby to make successful and knowledgeable choices. BDA allows data mineworkers and experts to scrutinize a vast amount of data that cannot be analyzed employing conventional methods [29]. BDA comprises technology and materials that can turn a vast quantity of organized, formless, and semi-organized data into an extra understandable framework for investigative procedures. The procedures employed in these computational methods have to recognize forms, tendencies, and associations in the data across several time prospects [31]. Such methods, after analyzing the results, illustrate the discoveries for effective verdict taking in tables, charts, and three-dimensional maps.

Big data processing is also a major problem for many implementations owing to the sophistication of data and the scalability of fundamental procedures facilitating these systems [32]. Obtaining useful information from big data analysis is a serious issue requiring adaptable investigative procedures and methods to report timely outcomes while existing methods and procedures are incompetent in handling BDA [33]. Therefore, there is a need for broad networks and external technologies to support parallel data. Besides, data foundations, for instance very advanced data streams obtained from various data foundations, have diverse arrangements which make it essential to incorporate multiple analytics solutions [34]. The problem is thus based on the efficiency of existing procedures employed in big data processing, which doesn't grow uninterruptedly with the exponential growth in computing capital [35]. Figure 1 displayed the big data conceptual framework for the COVID-19 Pandemic.

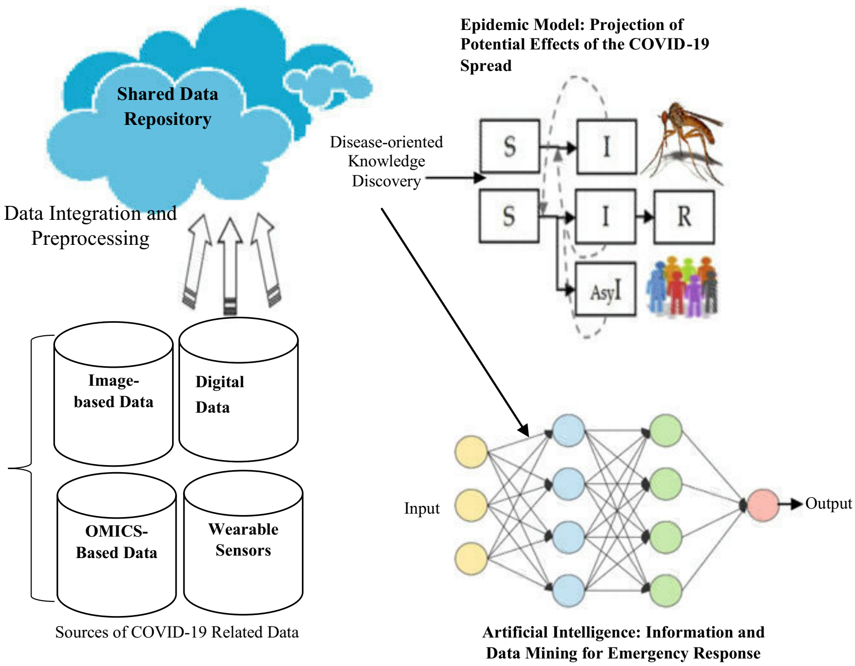

Fig. 1 An integrated big data conceptual model for COVID-19 Pandemic

BDA systems take a large amount of time to provide consumers with input and advice, although only a few applications [36] can handle massive data sets within a short processing period. By comparison, the greatest of the outstanding approaches employ a complex trial-and-error approach to tackle large volumes of data sets and data complexity [37]. There are big-data analytics platforms, for instance, the Investigative Data Examination System [38] is a big-data graphic analytics program that is employed to scrutinize complicated earth structure models of massive quantities of datasets. Big data volume is enormous and thus conventional applications for Database Management can't be used to accumulate and investigate huge data. The resolution comes from modern warehousing databases such as Apache Hadoop that support the analysis of distributed data.

Big data promises more personalized and effective medicine for patients during the COVID-19 outbreak with better precision and earlier diagnosis and tailored therapy to the specific combination of genes, environmental risk, and accurate disease phenotype in an individual. This promise comes from data obtained from various sources, ranging from molecules to cells, to tissues, to individuals and populations, and from the convergence of these data into networks that enhance understanding of health and disease. Big data-driven science could play a role in driving comparative medicine during the COVID-19 pandemic through sharing physiology, pathophysiology, and risk factor for diseases through species.

The integration of EHR data across institutions would provide for earlier unprecedented access to patient data on a scale during the COVID-19 outbreak, thereby facilitating accurate phenotype classification and objective assessment of risk factors and response to COVID-19 outbreak therapy. High-throughput molecular data will provide insight into previously unexplored pathophysiology of molecular and etiology of diseases. Big data analysis and integration from a variety of sources can lead to stronger connections drawn at the molecular level between human and animal disease, allow predictive modeling of infectious dis-ease and identifying key areas of intervention, and promote step-by-step understanding of COVID-19 disease, which can have a major effect on animal and human health. Using big data, however, comes with tremendous challenges. Therefore, in the COVID-19 pandemic, the paper discusses and includes big data opportunities, their limitations, and what's needed to capitalize on big data.

Big Data will aid in the COVID-19 pandemic handling of immense and unparalleled quantities of data from various sources such as situation briefing and updating by government entities and agencies, real-time COVID-19 monitoring and tracking outbreaks, pattern forecasting/now-casting, and health resource utilization knowledge [39]. The big data can spring from different sources: (i) image-based data obtained through radionics or data mining methods by extracting clinical and high-dimensional information from images (ii) Information derived from wearable sensors (iii) Computer and digital big data, such as smartphone, internet and other connected mobile devices (iv) molecular data that be obtained from OMICS-based methods, which includes genomics and post-genomics specialties and by wet-lab techniques [40–42].

2.1 Big Data Analytics Life Cycle

The BDA life cycle is demonstrated in Fig. 2, which encompasses 9 steps primarily beginning with the assessment of the market difficulty, performing an investigation on current approaches to fix the difficulty, describing the data including filtering as well as collecting the data. Established on the utilization obligation, the filtered data is extracted, validated, and cleansed. Through applying correct aggregation techniques, the collected data is suitably interpreted for further analysis. Once data is assembled out of several proveniences, the subsequent move is to perform experimental data scrutiny such that the method is selected. The concluding step in Big Data's life-cycle is to change the method developed and established into data visualization and outcomes scrutiny [43, 44].

Big Data analysis varies from conventional data analysis largely because of the data processes' length, velocity, and variety characteristics. A step-by-step approach

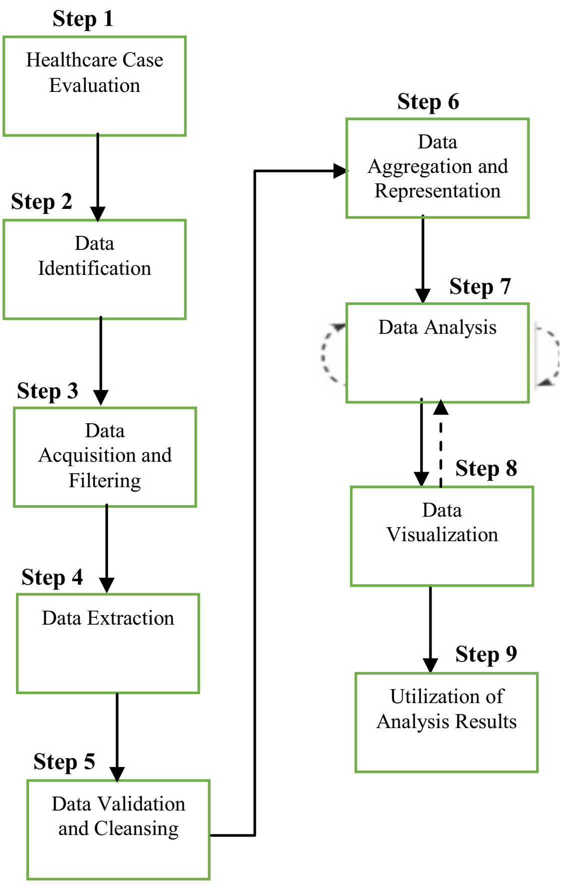

Fig. 2 The Big Data analytics lifecycle

is required to coordinate the processes and tasks involved in collecting, storing, evaluating, and repurposing data to resolve the distinct criteria for carrying out analysis of big data. From the viewpoint of Big Data acceptance and preparation, it is critical that accompanying the life span, attention is given to the matters of data processing recruiting, preparation, accomplishment, and staffing. Based on the complexity of the analytical problems being discussed, "models" that encapsulate new ideas and understandings about the existence of the dynamics and interactions that occur within the data being studied can be generated for the analytical results. A model can look like a math equation or a series of laws. Models can be used to develop the structure of the healthcare process and implementation system, so they can form the base of a new framework or software model.

3 The Opportunities of Big Data and Big Data Analytics in COVID-19 Pandemic

The possibility of fast-spreading diseases that causes uncontrolled death and a bad influence on the economies and sustainability of countries in the world has stretched the importance and needs for developing a quantitative framework to provide supports for making nearly real-time decisions in the public health system. A case study of the 2003 severe acute respiratory syndrome COVID-19 outbreak which originates from China and extends to 29 countries, which by August 2003 becomes healthcare-acquired infection in various regions [45, 46]. The Influenza (HiN1) pandemic of that originates from Mexico and spreads rapidly all over the world through airline network and infected 20 countries, with the maximum hit on the travelers coming from Mexico within a little time of the disease outbreak [47]. The socio-economic implication and outcome of the outbreak of a disease related to 2009 (influenza A(H1N1)) were projected to be a huge amount ranging between $360 billion–$4 trillion [48] been the calculation for the outbreak in the first year. Significantly, as of 28 August 2020, the SARS disease outbreaks at WUHAN in china in 2019 was calculated and projected to be 24,926,312 confirmed cases and 840,662 mortalities across the globe. Ever since the outbreak of COVID-19 in china, people that travel globally to WUHAN imported the disease to various countries of the globe.

The outbreak of severe acute respiratory syndrome disease causes trivia effects on the people and economy all over the world. This effect includes; restrictions and barns on traveling, total closure of shops and market resulting in a huge loss on earnings and proceeds, fear of contracting the disease from the infected person, unpleasant side effects on world tourism at large as many flights were canceled or suspended. Express and huge damage on government finances in various countries as agricultural activities were greatly hampered resulting in food shortage and so many other effects result from the outbreaks. Therefore it becomes imperative to make an immediate response from all angles to fight and contain the virus efficiently. Big data entails the processing of both new and existing data innovatively to proffer meaningful solutions that are of huge business benefits [49–51]. However, to process large data, or varied data will only be a mere technological solution until it is associated with business objectives and goals.

Big data promises More light to be shed on complicated facts and information on the dynamics of how infectious diseases are been transmitted with how to develop new modeling and analytical tools which have far been hidden due to lack of smooth and workable data. More progress will be made in the area of disease forecast when a huge amount of Information on epidemics required for modeling is available. Big data will be best in accomplishing this goal. Improved consistency in reporting diseases and case definition is therefore expedient within time and space. Data obtained from news media sources can give a perfect estimation and evaluation of how the disease is been transmitted. Thus, it's of high importance, especially when comprehensive and well-examined data are not available. In a country with low and middle income, and where studies on disease transmission details are scarce, resources based on internet surveys will provide an opportunity for detailed study and correct assessment on how the disease is been transmitted. For instance, in the crises of infectious diseases COVID -19 when the real-time assessment is expedient.

A well-calibrated computational model tool useful in assessing and analyzing epidemic spread signifies a potent tool to sustain the process of decision-making at a time of emergency epidemic outbreaks. Epidemic models are progressively used to generate predictions of the spatial-temporal progression of a disease outbreak at several spatial scales and for easy access to possible impacts of the various strategies of the intervention [52]. The ability to produce innovative facts or information is greatly expanded by big data. The cost applicable to answer various questions retrospectively and prospectively, by gathering ordered and structured data are exorbitant.

Acquisition of finer data in a computerized manner is achieved by analyzing unstructured data that are within EHRS using various computational techniques such as traditional language processors that extracts clinical concepts from free documents. Analyzing the unstructured data contained within EHRs using computational techniques (e.g., natural language processing to extract medical concepts from free-text documents) permits finer data acquisition in an automated fashion. For example, computerized detection within EHRS with the use of natural language processing is better used to detect complications that arise after the operation when compared to patient wellbeing indicators based on discharge coding [53].

Big data provides the possibility to generate an observable indication base for clinically related questions which may not be achievable but will be helpful with generalizability issues. Generalizability issues restrict the application of conclusions that are derived from random trials carried out on a very narrow scale of participants to patients that shows dissimilar characteristics. Big data helps disseminate knowledge. Many healthcare practitioners strive to be updated with current evidential proof and support that guides clinical practices. Figure 3 depicted the relationship between COVID-19 Pandemic and big data.

The COVID-19 outbreak can be monitored in real-time using big data. The recent pandemic outbreak has been used to create unique open-access databases containing