

Epidemic Diseases Forestall Module using Data Science and SIR Algorithms

Prof. Nivedita Shimbre¹, Tanvi Patil², Rutuja Late³, Kshitij Jagtap⁴, Amogh Patil⁵

¹Computer Science

JSPM's Rajarshi Shahu College of Engineering, Pune, India

nivedita.shimbre@gmail.com

²Computer Science

JSPM's Rajarshi Shahu College of Engineering, Pune, India

tanvispatil223@gmail.com

³Computer Science

JSPM's Rajarshi Shahu College of Engineering, Pune, India

rutujalate1909@gmail.com

⁴Computer Science

JSPM's Rajarshi Shahu College of Engineering, Pune, India

kshitijnjagtap@gmail.com

⁵Computer Science

JSPM's Rajarshi Shahu College of Engineering, Pune, India

amogh.patil@gmail.com

Abstract— This survey paper is intended to prevent epidemic diseases and pandemic diseases. According to the WHO every year in the world over 17 million people die due to this type of disease. Epidemic diseases have lower transmission rate than pandemic diseases and they spread in a bounded area. On the other hand, pandemic diseases have higher transmission rate and it can easily spread in an immense area. We can control this type of disease in its initial stages before it becomes a fatal disease like covid-19. Lack of knowledge in peoples and inefficient systems used by higher authorities in that region are the main reasons to spread diseases in larger areas. But using data science and the epidemic compartment models it's possible to control infectious diseases in its initial stages. For different diseases there are different compartment algorithms that are able to estimate the number of cases in the future. These models often use ordinary differential equations for predicting things. Using data science, we are able to find what are key factors responsible for the spreading of that particular disease.

Keywords- data science, SIR model, pandemic and epidemic diseases, healthcare, big data, infectious disease, transmission prevention

I. INTRODUCTION

The main focal point of this survey paper is to use data science and epidemic compartment models algorithms to control the higher and lower transmission diseases. Using this emerging technology, we can control the epidemic and pandemic diseases in its initial stages.

Data science is important as it gives oversight of infectious disease. Disease like covid-19, Influenza, Swine flu, SARS which spread through the air and diseases like cholera, typhoid fever which spread through the water. Now humans have support of data science which helps humans to study the behavior of the particular virus. On which conditions it spread more and which conditions it spread less also we can predict the number of cases in coming days. If we already know the number of cases in the future. Then we can make sure that there is no shortage of the medicines in that particular time period. This survey paper helps to understand and stop infectious diseases using data science.

Infectious diseases have generated massive data on the number of patients in particular diseases, factors that are responsible for that particular disease and contact tracing. Increasing the availability of data increases the accuracy of the model. That's why data science is a widely adopted technology for infectious data surveillance. Statistics is an important part of data science. Using statistics, we can change data into some useful results and graphs.

This survey paper aims to find and examine the factors that affect and responsible for the disease spreading. The module makes comparisons between the number of patients in particular disease and the factors responsible for spreading the disease. Regression analysis concepts, statistical methods and factor analysis are the techniques used to find the key factors. This comparison gives the information about which factor is most responsible for the disease spreading and which is less responsible. Then it helps to control the spread of disease focusing on those key factors which are most responsible for the spreading of the diseases. For example, in cholera disease

spreading the main factors responsible for the cholera spreading is drinking contaminated water and eating contaminated food but water get contaminated due to the factors like increasing the industry west in rivers, canals and dams, overfishing pollution, rising acidity of water, garbage in water, agricultural waste and flood. Food is contaminated due to over use of pesticides in agriculture, Toxic chemicals used to make veggies look fresh, storage space is unclean. Comparing these factors, we get the exact cause of spreading disease. Then we will be able to control it before it gets worse.

II. RELATED WORK

- Mathematical model of SIR epidemic system with fractional derivative (2021) [3]: In this research paper, researchers took into account the impact of the health system and investigated the SIR (susceptible-infectious-removed) dynamics. They studied the significance of the airport transport network in the COVID-19 outbreak and evaluated the impact of these control efforts. Our findings suggest that on January 13, 2020, the daily risk of exporting at least one SARS CoV-2 case from mainland China via international travel as it approached 95%.
- The role of data science in healthcare advancements (2021) [4]. With the emergence of numerous data-driven applications such as improved pattern recognition and disease forecasting, the researchers most certainly modified the human-machine interfaces and interpretation process to reach a more effective and productive environment. Through their vast range of data in the healthcare industry, including budgetary, clinical, research and development, management, and operational aspects, data scientists can get valuable insights into increasing the productivity of health care.
- A time-dependent SIR model for Covid-19 with undetectable infected persons (2020) [10]. Possible future work in this research paper is to broaden our SIR model to probabilistic models in order to include extreme events that under low prevalence show high variability in the analysis. This will improve accuracy of the prediction outcome even more.

III. MAJOR TECHNIQUES REFERENCED IN THE LITERATURE SURVEY

A. Data science

Disease monitoring, health care administration and management, data confidentiality and risk management, mental health, public health, and pharmacovigilance are six areas in which analytics can be used in the healthcare sector very effectively and efficiently.

The perception of an illness, understanding its condition, genesis, and prevention are all aspects of disease surveillance. The many monitoring approaches would help with service delivery, treatment evaluation, decision-making, and policy analysis. Image processing on healthcare information reveals crucial data about anatomy and organ function, as well as disease and patient health status. AI in image processing will improve prognosis, screening, diagnosis and combining images with genomic information will improve accuracy and speed up disease detection.

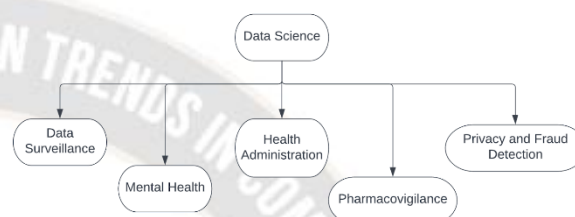


Figure 1. Different applications of data science in healthcare.

B. Data from wearable technology

Apple and Google, for example, are developing health-related apps and wearable technologies as part of a wide range of motion devices for healthcare-related software. The ability to collect accurate data in real-time (e.g., mood, nutrition, sleep patterns, exercise) and correlate it to physiological indications (e.g., bpm, calories burnt, blood glucose level) is likely discrete and global at low cost, unattached to traditional medicine.

Large volumes of data are generated through telemetry and devices that monitor physiological characteristics. Because the data collected is typically kept for a shorter period of time, significant research into the data produced is often overlooked. Advancements in big data in healthcare, on the other hand, aim to assure better data management and improved patient care.

C. Data from the administration of healthcare

Data warehousing and cloud services are largely used to safely and cost-effectively store the growing volume of computer-based patient data to improve medical results. Data storage is used for research, training, teaching, and quality control, in addition to medicinal uses.

The constant updating and exchange of a vast amount of data, as well as the transfer of picture reports to digital medical record systems, hold a lot of potential for radiology practice and research. During a patient's stay in the hospital, patient data management comprises the providing precise delivery of patient care.

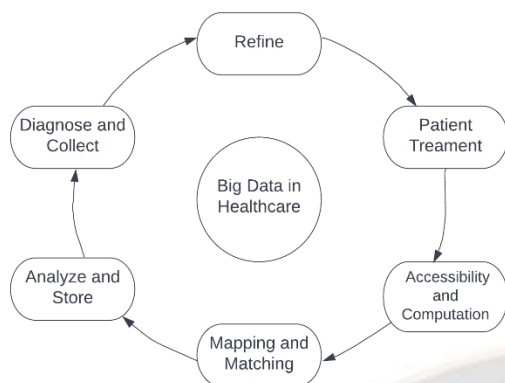


Figure 2. Steps taken to facilitate big data in healthcare.

D. *Fraud detection and Healthcare data privacy*

The privacy of patient data, its protection, and the detection of healthcare fraud are all critical. This necessitates data scientists' efforts to protect huge data from hackers. The application of big data analytics for fraud detection and abuse is beneficial.

E. *Mental Well-Being*

The use of data analytics in the evaluation, diagnosis, or treatment of people with mental health issues differs significantly from the use of analytics to predict cancer or diabetes. The data contexts appear to be more relevant than data volume in this scenario.

F. *Pharmacovigilance*

To ensure patient safety, pharmacovigilance needs recognition of ADRs (adverse drug reactions) and tracking after launch. The estimated annual social cost of ADR events is one billion dollars, indicating that it is a significant part of the healthcare system.

Medical care institutions can provide larger patient datasets that comprise data from laboratory, surveillance, genetics, imaging, and EHRs. To get useful information from this data, it must be properly managed and analyzed. Data science can provide real-time analytics that can be used to get insight into a range of diseases and provide patient-centric treatment. It will aid in the enhancement of researchers' abilities in the fields of science, epidemiological studies, modified medicine according to the needs, and so on. Accuracy of prediction is strongly reliant on effective data integration from several sources in order to be generalized. By forging new avenues in comprehensive medical care, data science can efficiently handle, analyze, and understand huge data.

IV. BASIC INFORMATION FOR DATA SCIENCE MODEL

Nowadays data science is an emerging technology in the world. It aims to collect the data, process it and generate useful results which helps to take any critical decisions. It is also useful in the healthcare sector to prevent epidemic diseases. Current day epidemic diseases are a big problem according to the WHO 17 million of people die every year because of the epidemic diseases.

Epidemic diseases have a high transmission rate. Diseases like Tuberculosis (TB), Flu, Zika, Yellow Fever, Cholera, Plague, Meningitis, MERS, Influenza, Swine flu, etc. are the epidemic diseases. These types of diseases spread through many ways like through air, through water etc. It spread to a larger area in a short time period. If the population is dense then the probability of people getting infected by that particular disease is high. An epidemic is an outbreak of diseases that affects a large population and spreads rapidly. An epidemic is defined as a disease outbreak that affects a large number of individuals in a brief span of time, usually less than two weeks. Epidemic diseases transmission rate is lower than pandemic diseases. Using the data science concepts like statistical analysis, data analysis we are able to generate the results which helps to control the epidemic diseases. Every epidemic disease has certain key factors which are responsible for the spreading of the particular disease.

In data science the data collection is an initial and important step in data science. The basic data collection methods are survey and data collection using the virtual interface like websites, mobile application in this case user provides the data. To control the epidemic diseases various types of data are required like number of patients in that particular region, environmental conditions, population density, location etc. This data is available with the help of different surveys and virtual interfaces. Cleaning the data is also an important step. In this step we remove some errors in the data or some different data which do not fulfill the conditions and requirements of that particular factor.

After that we perform the statistical operations like factor analysis, regression analysis etc. which generate some results which display in the form of graphs and charts which provide the exact cause of the epidemic disease which helps to prevent epidemic diseases. Once the model is fixed and data modeling is finalized then we have to train the model efficiently. The training can be done in parts: a) the important parameters can be modified further to get accurate output as required, b) then in the production phase, the model is subjected to actual data and its output is monitored.

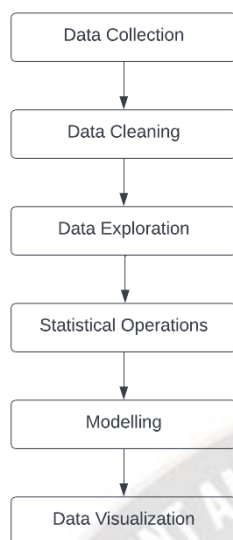


Figure 3. Steps describing the process of big data model creation.

V. EPIDEMICAL COMPARTMENT MODELS

A. Basic SIR model

A. G. McKendrick and W. O. Kermack created the SIR model for that they make three compartments first compartment of susceptible peoples. Susceptible peoples are those peoples who are not infected yet by that particular disease. Second compartment is of infected peoples and the last compartment is the recovered peoples. They consider fixed population means no death and birth occur in this time period. This is the main disadvantage of the SIR model after that some modifications happened in this model according to the diseases and some other conditions.

Flow of compartments is considered as follows:

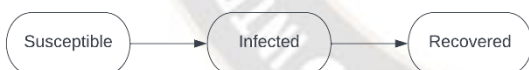


Figure 4. Basic SIR model

$S(t)$: Rate of change of susceptible people with respect to time

$I(t)$: Rate of change of infected people with respect to time

$R(t)$: Rate of change of recovered people with respect to time

$$N = S(t) + I(t) + R(t) \quad (1)$$

$$dS/dt = -\beta SI / N \quad (2)$$

$$dI/dt = -\beta SI / N - \alpha I \quad (3)$$

$$dR/dt = \alpha I \quad (4)$$

Here N is the total population mean. The population remains constant in all processes. Initially, we consider at $t=0$ the number of people in each compartment is zero. Here β is the

coefficient of transmission. Means at what rate a particular disease is spreading in that particular area. α is the recovery rate. It gives information about how many people are recovering from this disease.

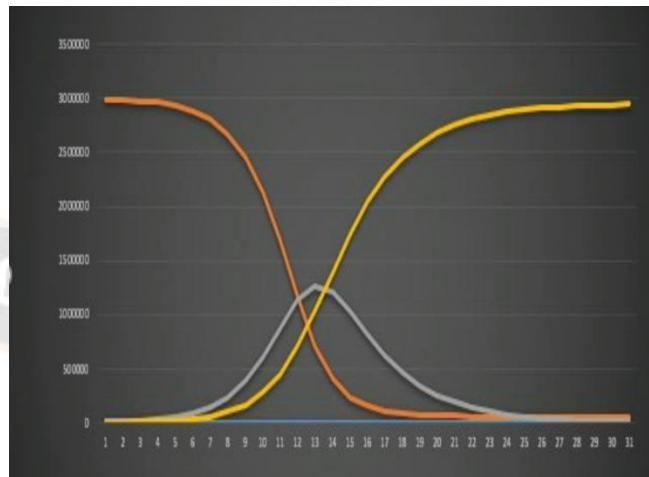


Figure 5. Graph of SIR model showing the relation between S, I and R.

B. SIS Model

The SIS model differs from the SIR model in that infected persons recover with no immunity and become immediately susceptible to reinfection. Diseases like gonorrhea, the common cold, and tuberculosis follow the SIS model.



Figure 6. Different steps of the SIS model

$S(t)$: Rate of change of susceptible people with respect to time

$I(t)$: Rate of change of infected people with respect to time

$R(t)$: Rate of change of recovered people with respect to time

N : Total population

$$dS/dt = -\beta SI / N \quad (5)$$

$$dI/dt = -\beta SI / N - \alpha I \quad (6)$$

C. SIRD Model

This type of model includes four compartments susceptible, Infected, Recovered, Deceased, or dead. According to this model, susceptible people come in contact with infected persons and susceptible people become infected. These infected people recover or die due to disease. Here μ is the rate of change of mortality means how many people die within a certain time.



Figure 7. Different steps of the SIRD model

S(t): Rate of change of susceptible people with respect to time

I(t): Rate of change of infected people with respect to time

R(t): Rate of change of recovered people with respect to time

D(t): Rate of change of dead people with respect in time

N: Total population

$$dS/dt = -\beta SI / N \quad (7)$$

$$dI/dt = -\beta SI / N - \alpha I - \mu I \quad (8)$$

$$dR/dt = \alpha I \quad (9)$$

$$dD/dt = \mu I \quad (10)$$

VI. CHALLENGES

For EHR-based profiling consisting of multiple clinical features inside the EHRs, millions of data sets are available. Managing and controlling the whole data of millions of people, like sequencing data, is a big task. The main challenges faced include:

- Standardization of data, privacy, effective storage, and transfers all demand a large amount of workforce to consistently scan/record and ensure that the requirements are met.
- Since there are no standards for handling bioinformatics, data deposition, build next-generation sequencing (NGS) data, or supporting medical decision-making, merging genomic data into medical investigations is crucial.
- When working with data, there is a language barrier.
- It's tough to find the right stability between retaining victim-centric data and assuring its standard and availability.
- The data gathered was generally disorganized or erroneous, making it difficult to get insight into it.

“(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

VII. CONCLUSION

We use this model as smart medical management is able to predict and find the factors responsible for the spread of diseases like COVID-19, zika, TB, and cholera. In this paper,

we propose and execute SIR, SIS, and SIRD, models with machine learning algorithms. We represent both mathematical analyses and carbon results. These models will aid in the development of scientific and epidemiological research capabilities. The primary aim of this study is to evaluate the models for providing smart medical management that is able to anticipate infectious diseases like covid-19, influenza, Ebola, and swine flu.

VIII.OBSERVATION OF LITERATURE SURVEY

TABLE I. OBSERVATIONS

Author	Methodology	Summary	Result
1. Qingpeng Zhang (2021)	Intra -MSA percolation, Inter-MSA percolation, and Time series of the threshold predictability, these techniques are used in this paper.	Human mobility, contact tracking, analytic records, and other data have been created as a result of infectious disease outbreaks.	Using different data science techniques, we were able to make surveillance of causes of diseases and control the diseases before it creates a pandemic like situation. We can monitor the mobility network of humans in particular regions.
2. Qingpeng Zhang, Jianxi Gao, Joseph T. Wu, Zhidong Cao and Daniel Dajun Zeng (2021)	Patient data mining, contact tracing using the GPS and Bluetooth.	The new techniques are compared to the traditional epidemiological research. It analyzed the COVID-19 spread, and explored the benefits and drawbacks of using data science to combat time ahead infectious disease outbreaks.	It proved that using data science techniques we can do drug screening, contact tracing, and medical imaging.
3.Alqahtani, R.T. (2021)	In this paper, the dimensional model, the model with fractional derivative the numerical techniques are used for solving	The susceptible-infectious-recovered (SIR) dynamics are explored and evaluated, considering the impact of the	It proved the uniqueness, extent, and bound of the model. It studied all details of stability and possible

	the differential equations.	health system. A general occurrence rate function and a healing rate function are considered as functions of the no. of hospital beds. When a fundamental reproduction number $R_0 < 1$, the free steady state is locally stable, but when $R_0 > 1$, it is unsteady. The phenomenon of backward bifurcation occurs, according to the study, when $R_0 = 1$.	steady-state infusion of the model derived from the reproductive number.		Procedures, such as wearing masks and maintaining social distancing.	infectious respiratory diseases that has been conducted in response to the growing concern about the transmission of these diseases.	efficient form the conventional approaches.
				6. Paoluzzi, Matteo & Gnan, Nicoletta & Grassi, Francesca & Salvetti, Marco & Vanacore, Nicola & Crisanti, Andrea (2021)	Susceptible-Infected-Removed (SIR) mathematical model.	These models depict the transmission of infectious illnesses across homogenous compartments with varying levels of health. The SIR model is the most basic, in which the susceptible compartment transforms into infected at a rate of infection.	This paper proved that by using mathematical models we are able to predict the number of patients in the future.
4. Subrahmanya, S.V.G., Shetty, D.K., Patil, V. et al. (2021)	Data mining approaches such as data correlation, data grouping, data clustering, and regression, sequential patterns mining, and storage of data are discussed in this research study.	It is proven that real-time prediction analytics are utilized to better understand disease processes and provide patient-centric treatment. They aid in the enhancement of researchers' abilities in the fields of customized medicine, epidemiological studies and proves that prediction accuracy is strongly reliant on effective data integration from several sources in order to be standardized.	This research paper showed big data analytics and data science can provide relevant insight and aid in determining the strategies in health-care decisions. It contributes to the development of a holistic picture of consumers, patients, and physicians. Determining data-driven decisions brings up new avenues for improving healthcare quality.	7. Piu Samui, Jayanta Mondal, Subhas Khajanchi (2020)	They examined the susceptibility indices of the fundamental reproductive number R_0 in the SAIU model for COVID-19 because R_0 identifies the start of sickness transmission and susceptibility indices allow us to determine the relative importance of different parameters in coronavirus transmission.	This paper proposed the SIR model that gives transmission of disease. A sensitivity analysis is performed to determine the best effective treatment for this condition.	According to predictions, COVID-19 will reach a peak in India in roughly 60 days, after which the curve will flatten, however coronavirus diseases will persist for a long time.
5. Leite GS, Albuquerque AB, Pinheiro PR (2021)	In this research paper, deep learning techniques are used to detect Standard Operating	Several researchers focused on the application of technology in the prevention of emerging	This paper proved that some approaches in data science that are different and	8. Manotosh Mandal, Soovoojeet Jana, Swapam Kumar Nandi, Anupam Khatua, Sayani Adak, T.K. Kar (2020)	This paper teaches us about optimum problem control, numerical simulations, and systems design for fixed control.	In this paper, we learn about the SIR model, which may help us determine whether these intervention approaches are the best approaches for illness control	To dynamically study the disease, investigate the disease, manage the destructive character of disease, and decrease transmission

		and how they may influence the disease dynamic.	of infectious disease.				outbreak in China.
9. Gostic, Katelyn and Gomez, Ana CR and Mummah, Riley O and Kucharski, Adam J and Lloyd-Smith, James O (2020)	In this paper we study about epidemiological factors, screening programs to catch the disease, predicting worst-case scenario, questionnaires based screening	Given existing information of essential COVID-19 life history and epidemiological factors, we assess the effect of various screening programs. We predict that testing will miss more than half of sick patients, even under the best-case scenario.	The Covid-19 screening approach aids in the detection of symptoms that emerge before questionnaire-based exposure risk screening, and discovered instances do not advance to the next step. This allows us to measure the percentage of instances found using symptom or risk screening at entry or exit.	11. S. Latif et al. (2020)	Artificial intelligence (AI), natural language processing (NLP), machine learning (ML), algorithms, modeling, statistics simulation, and any other scientific approaches that learn from organized and unstructured data are all examples of artificial intelligence (AI). We recognise the value of insights from the social sciences, ethics, history, and other disciplines, but they are outside the scope of this project.	This paper tries to systemize the varied COVID-19 analysis activities leverage knowledge science, wherever knowledge science is outlined loosely to comprehend the varied ways and tools that may be wont to store, process, and extract insights from knowledge, together with those from computer science, machine learning, statistics, modeling, simulation, and knowledge image	This report relies on current evaluations and perspectives to assist systematize available resources and aid the research community in developing COVID-19 pandemic solutions. They sought to be thorough, but in a continuously growing topic like this, exhaustiveness is impossible to achieve.
10. Yi-Cheng Chen and Ping-En Lu and Cheng-Shang Chang and Tzu-Hsuan Liu (2020)	In this research paper, they have used SIR models, Differential Equations for the Tracking Transmission Rate and Recovering Rate by Ridge Regression. Tracking the no. of Recovered Persons and the no. of Infected Persons from COVID-19, collected the data from the National Health Commission of the People's Republic of China (NHC).	In this paper, we study mathematical and numerical analyses to answer questions such as when the epidemic will peak and when it will cease, and how asymptomatic infections impact disease propagation.	Numerical results for the dataset collected from the People's Republic of China's National Health Commission (NHC), show that one-day prediction errors for the no. of infected people X_t and the no. of recovered people R_t are (nearly) 3%. Furthermore, by following the features of the transmission rate and the recovery rate with respect to time t , we may predict the future growth of the COVID-19	REFERENCES <ol style="list-style-type: none"> [1] Q. Zhang, "Data science approaches to infectious disease surveillance," vol. 380, no. 2214. Royal Society, Nov. 22, 2021. DOI:https://doi.org/10.1098/rsta.2021.0115. https://royalsocietypublishing.org/doi/10.1098/rsta.2021.0115 [2] Q. Zhang, Jianxi Gao, J. T. Wu, Z. Cao, and D. D. Zeng, "Data science approaches to confronting the COVID-19 pandemic: a narrative review," vol. 380, no. 2213. Royal Society, Nov. 22, 2021. DOI: https://doi.org/10.1098/rsta.2021.0127. https://royalsocietypublishing.org/doi/10.1098/rsta.2021.0127 [3] R. T. Alqahtani, "Mathematical model of SIR epidemic system (COVID-19) with fractional derivative: stability and numerical analysis," vol. 2021, no. 1. pp. 2-, Jan. 04, 2021. DOI:10.1186/s13662-020-03192-w. https://advancesindifferenceequations.springeropen.com/articles/10.1186/s13662-020-03192-w. [4] Ana Oliveira, Yosef Ben-David, Susan Smit, Elena Popova, Milica Milić. Enhancing Data-driven Decision Making with Machine Learning in Decision Science. Kuwait Journal of Machine Learning, 2(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/200 			

- [5] S. V. G. Subrahmanya *et al.*, "The role of data science in healthcare advancements: applications, benefits, and future prospects." Aug. 16, 2021. doi: 10.1007/s11845-021-02730-z. <https://link.springer.com/article/10.1007/s11845-021-02730-z>
- [6] G. S. Leite, A. B. Albuquerque, and P. R. Pinheiro, "Applications of Technological Solutions in Primary Ways of Preventing Transmission of Respiratory Infectious Diseases-A Systematic Literature Review.," vol. 18. Oct. 14, 2021. DOI: 10.3390/ijerph182010765. <https://www.mdpi.com/1660-4601/18/20/10765>.
- [7] Chinmay Rawat. (2023). AI for Effective use of Water in India for Crop Cultivation. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3s), 266–270. Retrieved from <https://ijisae.org/index.php/IJISAE/article/view/2683>.
- [8] M. Paoluzzi, N. Gnan, F. Grassi, M. Salvetti, N. Vanacore, and A. Crisanti, "A single-agent extension of the SIR model describes the impact of mobility restrictions on the COVID-19 epidemic." Dec. 28, 2021. doi: 10.1038/s41598-021-03721-x. <https://www.nature.com/articles/s41598-021-03721-x>
- [9] P. Samui, J. Mondal, and S. Khajanchi, "A mathematical model for COVID-19 transmission dynamics with a case study of India," vol. 140. pp. 110173-, Nov. 01, 2020. DOI: <https://doi.org/10.1016/j.chaos.2020.110173>. <https://www.sciencedirect.com/science/article/pii/S0960077920305695>
- [10] M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak, and T. K. Kar, "A model based study on the dynamics of COVID-19: Prediction and control," vol. 136. pp. 109889-, Jul. 01, 2020. DOI: <https://doi.org/10.1016/j.chaos.2020.109889>. <https://www.sciencedirect.com/science/article/pii/S0960077920302897>.
- [11] Prof. Naveen Jain. (2013). FPGA Implementation of Hardware Architecture for H264/AV Codec Standards. *International Journal of New Practices in Management and Engineering*, 2(01), 01 - 07. Retrieved from <http://ijnpme.org/index.php/IJNPME/article/view/11>.
- [12] K. Gostic, A. C. Gomez, R. O. Mummah, A. J. Kucharski, and J. O. Lloyd-Smith, "Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19," vol. 9. eLife Sciences Publications, Ltd, Feb. 24, 2020. doi: 10.7554/eLife.55570. <https://elifesciences.org/articles/55570>
- [13] Y.-C. Chen, Ping-En Lu, C.-S. Chang, and T.-H. Liu, "A Time-Dependent {SIR} Model for {COVID}-19 With Undetectable Infected Persons," vol. 7. Institute of Electrical and Electronics Engineers (IEEE), pp. 3281–3294, Oct. 0, 2020. doi: 10.1109/tNSE.2020.3024723. <https://ieeexplore.ieee.org/document/9200529>
- [14] S. Latif *et al.*, "Leveraging Data Science to Combat COVID-19: A Comprehensive Review," vol. 1. pp. 85–103, 0 0, 2020. doi: 10.1109/TAI.2020.3020521. <https://ieeexplore.ieee.org/document/9184922>.