



BIBLIOMETRIC ANALYSIS OF COVID-19-BASED MODELS OF INDIA

*Ameet**

*Chhavi Rana***

*Sanjeev Kumar****

*Yogesh Kumar*****

*Research Scholar, Department of C.S.E, U.I.E.T, M D University Rohtak, India

**Associate Professor, Department of C.S.E, U.I.E.T, M D University Rohtak, India

***Research Scholar, Department of C.S.E, U.I.E.T, M D University Rohtak, India

****Assistant Professor, Department of C.S.E, U.I.E.T, M D University Rohtak, India

[*ameet.rs.uiet@mdurohtak.ac.in](mailto:ameet.rs.uiet@mdurohtak.ac.in)

[**chhavi.uiet@mdurohtak.ac.in](mailto:chhavi.uiet@mdurohtak.ac.in)

[***Sanjeevbumbra07@gmail.com](mailto:Sanjeevbumbra07@gmail.com)

[****dryogeshkumar,uiet@mdurohtak.ac.in](mailto:dryogeshkumar,uiet@mdurohtak.ac.in)

Article History

Volume 6, Issue 13, July 2024

Received: 04 June 2024

Accepted: 05 July 2024

doi:

10.48047/AFJBS.6.13.2024.4797-4806

ABSTRACT

This research paper reviews the publication trends of the Web of Science(WOS) and Scopus databases from 2020 to 2023 topics related to the prediction of COVID-19 numbers in India. Through bibliometric analysis, we investigate trends in publications and citations of COVID-19, highly cited papers, most frequent authors, and their affiliated institutions and network visualisation is performed using various types of analysis, i.e. co-authorship of authors, co-occurrences of keywords, bibliographic coupling, and co-citation of cited sources. Various biometric tools has been analysed Citations per publication(C/P), Cited Publications (CP), citation density and the finding of the study indicate that lots of research has been done during this period (2020-23) on the prediction of COVID-19 using machine learning and deep learning techniques. Useful insights in the findings provide greater chances to explore more in the field of prediction of the spread of viruses using machine learning techniques.

Keywords: Bibliometric analysis, COVID-19, Scopus, Web of Science, network visualisation.

1. INTRODUCTION

A common and reliable way to look at and analyze a lot of scientific data is through bibliometric analysis. It helps us understand how a certain field has changed over time and highlights emerging areas in that particular field. Bibliometric analysis can reveal trends in article and journal performance, collaboration patterns, and research elements, as well as the intellectual structure of a certain area in the literature[1]. The objective of the bibliometric analysis is to focus on the performance of articles and journals based on the number of citations, publications, occurrences of keywords, or research work. Understanding the present research on COVID-19 requires a study of the literature related to COVID-19, and the study is of major relevance. In this paper, Web of Science(WOS) and Scopus databases have been considered for bibliometric analysis. This study examines the numbers of COVID-19 documents, annual trends in publications, and citations, as well as highly cited articles, most frequent authors, and affiliations and Institutions. This paper presents a detailed year-wise analysis of COVID-19 publications, including trends in the number of publications, citations, and highly cited papers. This temporal analysis helps in understanding the evolution of research focus over the specified period. The novelty in the paper's approach lies in its holistic and detailed examination of COVID-19 prediction research, considering the regional context of India and employing a variety of bibliometric tools for a thorough analysis. The integration of multiple tools, the focus on prediction models, and the inclusion of regional aspects contribute to the uniqueness of the paper.

2. METHODOLOGY

The collection of COVID-19 papers is from two renowned databases, Web of Science (WOS) and Scopus, from 2020 to 2023. The search strategy for the selection of keywords is “Prediction model” AND (“Deep Learning” OR “Machine Learning”) AND “COVID-19” AND “India”. A total of 635 publications have been identified from Scopus database and 23 publications from the WOS database. The selection of publications focuses on the prediction-based model using both deep learning and machine learning techniques, and publications must consist of Indian authors or case studies from India.

3. RESULTS

3.1. NUMBERS OF COVID-19 PUBLICATIONS

Total 658 publications have been identified that fulfilled the search strategy. Out of 658 publications, 635 are from Scopus and the remaining 23 are from WOS database. For Scopus, maximum of 275 documents are submitted in 2022, 225 documents are submitted in 2021, 68 documents are submitted in 2023 and at least 67 documents are submitted in 2020. For WOS, a maximum of 11, 7, 3, and 2 are submitted in 2021, 2022, 2023, and 2020.

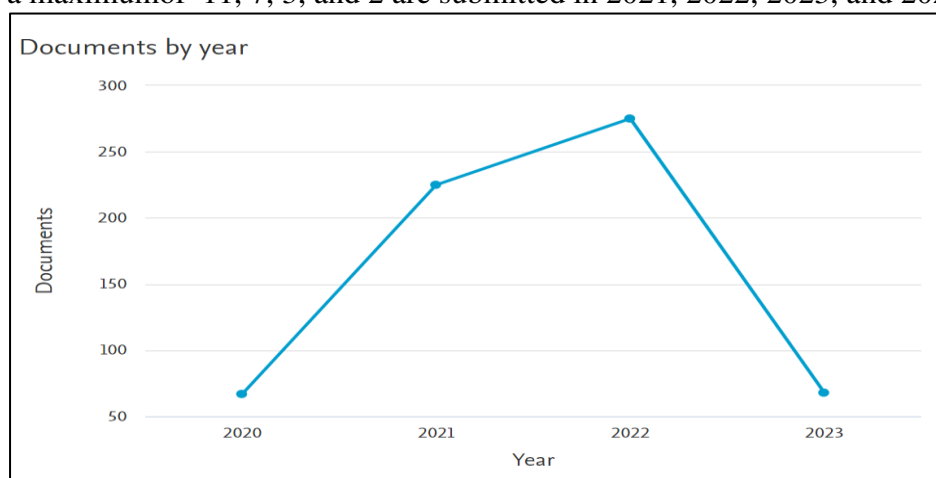


Figure 1: Covid-19 publications by year

3.2. ANNUAL TRENDS OF PUBLICATIONS AND CITATIONS OF COVID-19

In the beginning of 2020, the outbreak of COVID-19 in major countries was found. In 2020, 69 documents will be published (P) of which 60 of be cited as publications (CP). In that year, the total citations (C) of all publications related to COVID-19 articles were 2411 with citations per publication (C/P) is 34.94. Further, 15 articles are above or equal 50 to citations, 7 articles are above or equal to 20 citations, 9 articles are above or equal to 10 citations, and 12 articles are above or equal to 5 citations. In 2021, 136 documents are published (P) in which 175 of them are cited publication (CP). In that year, total citations (C) of all publications related to COVID-19 articles were 2576 with citations per publication (C/P) is 10.91. Further, 12 articles are above or equal 50 to citations, 29 articles are above or equal to 20 citations, 28 articles are above or equal to 10 citations, and 27 articles are above or equal to 5 citations. In 2022, 282 documents are published (P) in which 132 of them are cited publication (CP). In that year, total citations (C) of all publications related to COVID-19 articles are 621 with citations per publication (C/P) of 2.20. Further, 8 articles are above or equal to 20 citations, 11 articles are above or equal to 10 citations, and 19 articles are above or equal to 5 citations. In 2023, 71 documents are published (P) in which 16 of them are cited publication (CP). In that year, total citations (C) of all publications related to COVID-19 articles were 43 with citations per publication (C/P) of 0.60. And only 1 article is above or equal to 10 citations. It is observed a 58.20% of the documents are cited. Other interesting facts have been observed a total of 4287 document results that cite these 635 articles. A Total of 170, 1141, 2208, and 768 documents are cited; these selected 635 documents are in 2020, 2021, 2022, and 2023, respectively.

Table No.1: Annual trends of publications and citations of COVID-19

Year	P	CP	C	C/P	>=50	>=20	>=10	>=5
2020	69	60	2411	34.94	15	7	9	12
2021	236	175	2576	10.91	12	29	28	27
2022	282	132	621	2.20	0	8	11	19
2023	71	16	43	0.60	0	0	1	0
Total	658	383	5651		27	44	49	58
%		58.20						

3.3. HIGHLY CITED PAPERS OF COVID-19

Table No. 2 depicted the top 10 highly cited documents related to COVID-19 between 2020 to 2023. The table also considers citation density which total numbers of citations divided by a number of years since the publication of a document. Rank 1 document with 227 citations having a citation density of 75.66 “A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images” published in 2020[2]. The article proposed a detection model to detect COVID-19 cases using X-ray images by implementing GRAD-CAM color visualisation techniques. This model detected in less than 2 seconds that detected faster than the RT-PCR test. Rank 2 document with 214 citations “Forecasting and planning during a pandemic: COVID-19 growth, supply chain disruptions, and governmental decisions” focused on forecasting COVID-19 growth rates[3]. The third highest citation document (204 citations) was published in 2020 “A machine learning forecasting model for COVID-19 pandemic in India” which developed a prediction model based on machine learning techniques that predict COVID-19

cases in India[4]. Next fourth highest cited document(198 citations) was published in 2020 “Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India” also developed a prediction model that uses deep learning algorithms to forecast COVID-19 cases in India[5]. 5th highest cited document(181 citations) “Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics” published in 2020 used Logistic regression and other machine learning models to predict and analyse the COVID-19 trends[6].The sixth highest cited document (167 citations) “Automated medical diagnosis of COVID-19 through Efficient Net convolutional neural network” used EfficientNet based on the CNN approach that classifies the patients into COVID and non-COVID patient and shows a 99.63% recall value[7]. Seventh most cited document with 164 citations “Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020” published in 2020 that discussed the how the weather has influenced the spread of COVID-19 in the USA[8]. Eighth highest cited document with 140 citations “Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study” predicted the COVID-19 situation in India-USA using deep learning models[9]. Ninth highest cited document(121 citations) is “Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming”. The paper performed time series analysis and forecast COVID-19 cases using a new form of algorithm known as the genetic algorithm[10]. Tenth highest cited document(104 citations) “Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries” discussed the mutation of the spread of coronavirus in most affected countries[11].

Table No.2: Highly cited papers of COVID-19

Rank	Citations	Citation Density	Title	Author	Year
1	227	75.66	A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images[2]	Panwar, H., Gupta, P.K., Siddiqui, M.K., Morales- Menendez, R., Bhardwaj, P., Singh, V.	2020
2	214	107	Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions[3]	Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., Vasilakis, C.	2021
3	204	68	A machine learning forecasting model for COVID-19 pandemic in India[4]	Sujath, R., Chatterjee, J.M., Hassanien, A.E.	2020
4	198	66	Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India[5]	Arora, P., Kumar, H., Panigrahi, B.K.	2020
5	181	60.33	Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics[6]	Wang, P., Zheng, X., Li, J., Zhu, B.	2020
6	167	55.66	Automated medical diagnosis of COVID-19 through Efficient Net	Marques, G., Agarwal, D., de	2020

			convolutional neural network[7]	la Torre DÃ-az, I.	
7	164	54.66	Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020[8]	Gupta, S., Raghuvanshi, G.S., Chanda, A.	2020
8	140	46.66	Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study[9]	Shastri, S., Singh, K., Kumar, S., Kour, P., Mansotra, V.	2020
9	121	40.33	Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming[10]	Salgotra, R., Gandomi, M., Gandomi, A.H.	2020
10	104	52	Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries[11]	Wang, R., Chen, J., Gao, K., Wei, G.-W.	2021

3.4.MOST FREQUENT AUTHORS

Around 2100 authors contributed to 658 documents over a period of more than 3 years from 2020-2023. According to analysis, Meenu Gupta affiliated with Chandigarh University, Mohali, India has topped the most frequent authors who work on 4 documents engaged with 18 other authors[12]. Her Collaborative Index(CI) is 5.25 and CI is calculated as a number of contributing authors per total publication. Her two publications named as "AI-enabled COVID-19 outbreak analysis and prediction: Indian states vs. union territories" and "Real-time measurement of uncertain epidemiological appearances" have more than 20 citations[13][12]. Two other authors who also published 4 documents, Rajesh Ittamulla with affiliation from the Indian Institute of Technology, Hyderabad, India, and Rachna Jain affiliated with Bhagwan Parshuram Institute of Technology, NewDehli, India. Rachna Jain scored the highest CI of 5.5 and engaged with 19 other authors. Ittamulla's publications have a highest citation of a 92 with Citation/Publication value of 23 and having 2 publications with more than 20 citations and 2 publications with more than 5 citations. The above discussion is depicted in table.3.

3.5. MOST FREQUENT AUTHORS' AFFILIATING INSTITUTIONS IN COVID-19

Most of the top institutes of India have contributed to COVID-19 research. Table depicted the top six Indian institutes who engaged with researchers to publish a good number of publications. Top ranked institute among all is Amity University, Noida, India published the highest 20 documents with a citation of 86. C/P is 4.3 and two Publications are above or equal to 20 citations, 1 publication are above or equal to 10 and 5 citations. The second highest publication institute is Delhi University, New Delhi which published 11 publications during 2020-2023 with citations of 77 and 3 and two publications above or equal to 20 citations, 1 each publication above or equal to 10 and 5 citations. The third highest publication institute is Delhi Technical University, New Delhi which published 8 documents with citations of 10.

Table No.3:Most frequent Authors

Rank	Author	Affiliation	City	P	N	CI	C	C/P	>=50	>=20	>=10	>=5
1	Gupta,M.	CU	Mohali,India	4	18	5.25	60	15	0	2	0	0
2	Ittamalla, R.	IIT	Hyderabad,India	4	13	4	92	23	0	2	0	2
3	Jain, R.	BPIT	New Delhi, India	4	19	5.5	60	15	0	2	0	0
4	Al-Turjman,F.	YDU	Nicosia, Cyprus	3	14	5.3	5	1.66	0	0	0	0
5	Battacharyya, D.	KLDU	Vaddedwaram, India	3	11	4.3	1	0.33	0	0	0	0

The fourth highest publication institute is Vellore Institute of Technology; Vellore, India published 9 documents with the highest citation of 242 and a C/P of 30.25. One of the document of VIT is above 200 citations[4]. The fifth highest publication institute is Siksha O Anusandhan Deemed University, Bhubaneswar, India published 8 documents with citations of 89 one of the paper above or equal to 50 citations. Sixth highest publication institute Christ University, Bangalore, India has published 8 documents with 80 citations having one of the paper above or equal to 50 citations.

Table No.4: Most frequent authors' affiliating institutions in COVID-19

Rank	Institute	City, country	P	C	C/P	>=100	>=50	>=20	>=10	>=5
1	AU	Noda, India	20	86	4.3	0	0	2	1	1
2	DU	New Delhi, India	11	77	7.7	0	0	2	1	1
3	DTU	New Delhi	8	10	1.25	0	0	0	0	0
4	VIT	Vellore, India	9	242	30.25	1	0	1	0	2
5	SOADU	Bhubaneswar, India	8	89	11.12	0	1	0	0	1
6	CU	Bengaluru,India	8	80	10	0	1	0	0	0

4. NETWORK VISUALISATION

Around 15100 references were cited by 658 documents from 4867 different sources. In this section, network visualisation is performed using various types of analysis i.e. co-authorship of authors, co-occurrences of keywords, bibliographic coupling, and co-citation of cited sources. Network Visualisation is performed using VOSviewer open source software.

4.1.CO-AUTHORSHIP OF AUTHORS

Co-authorship network analysis used authors as a unit of analysis. For analysis, consider the minimum number of documents of an author as 1 and the minimum number of citations of an author as 5. Using this criterion, a number of selected authors is 633 out of 1968. A total of 20 clusters, 1572 links, and a total link strength of 1632 have been generated. Largest node of the network analysis is (Kumar S.) which belongs to cluster 5(pink) with 33 links. Author has published 15 documents with 373 citations and has highest total link strength of 40.

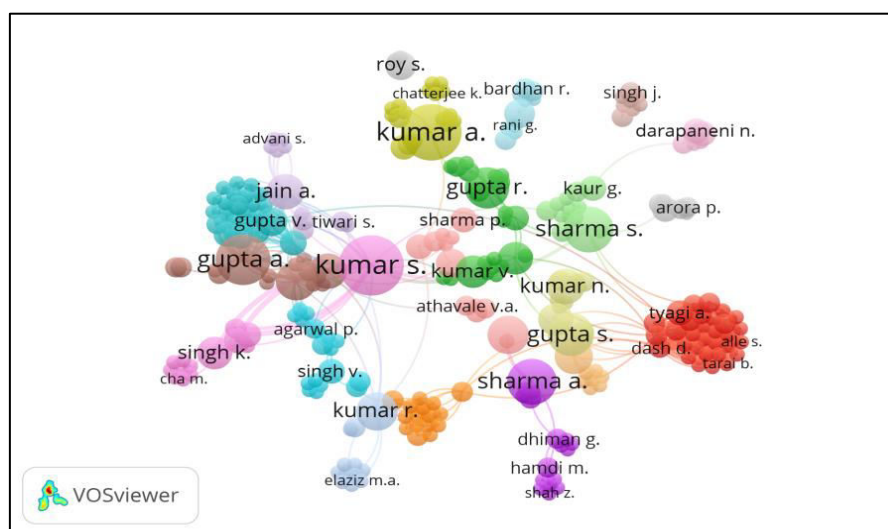


Figure 2. Co-authorship of authors

4.2.CO-CITATION OF CITED SOURCES

Cited sources are used as a unit of analysis, minimum number of citations of a source is considered as 20. On the basis of the threshold only 52 sources have qualified out of 9149 sources. A total of 4 clusters have been identified consisting of 801 links and 8700 total link strength. PLOS one with the highest total link strength tops the list having largest node with 51 links belonging to cluster 1. The second largest source is IEEE access (cluster: 1) with 202 citations and total link strength of 1155 having 51 links. The third largest source node is Lancet with 147 citations and 1045 total link strength having 46 links. The fourth largest source node is Chaos Salitons Fractals with 139 citations and 895 total link strength having 37 links. Other larger nodes sources are medrxiv(cluster:1, citations: 159, Total link strength: 798, links: 49), and nature(cluster:2, citations:82, total link strength: 703, links: 48) have contributed great in field of COVID-19.

4.3.CO-OCCURENCES OF KEYWORDS

For co-occurrence of keywords analysis, the minimum criteria for number of occurrences of a keyword should be 10. On this basis, of the 4239 keywords only 132 keywords meet the threshold. For each 132 keywords, the total strength of the co-occurrence link with other keywords will be calculated. 5 clusters have been identified with 4499 links and 20606 total link strength. The most wide keyword occurrence in many of the sources was found to be

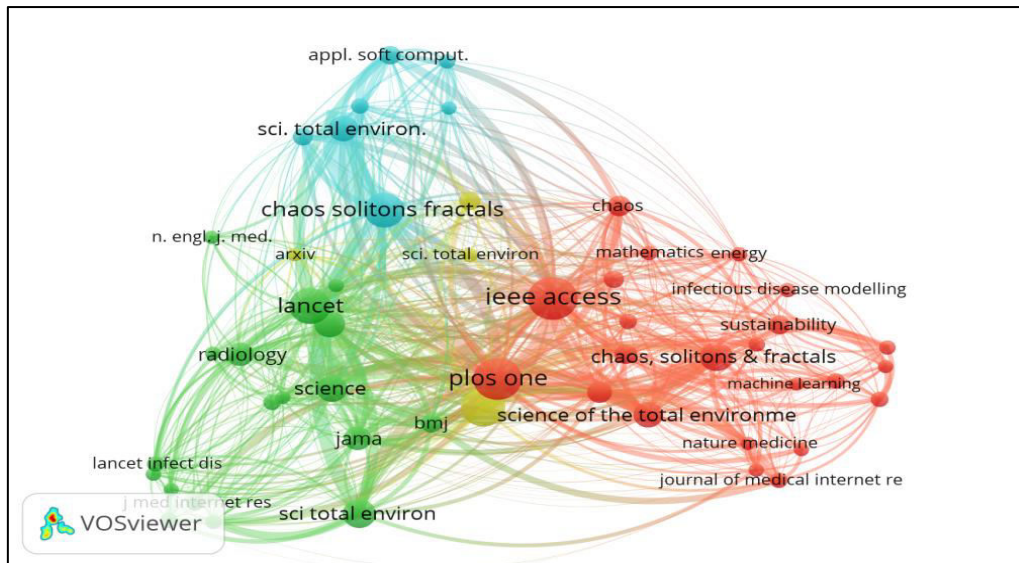


Figure 3. Co-citation of cited sources

“Covid-19”(Cluster:1) which occurred 405 times and generated 131 links with 2992 total link strength. The Second highest occurrence of word is “machine learning” (Cluster:1) which occurred 282 times and generated 131 links with 2063 total link strength. The third highest occurrence of the keyword is “India”(Cluster: 2) occurred 120 times and generated 135 links with 1559 total link strength.

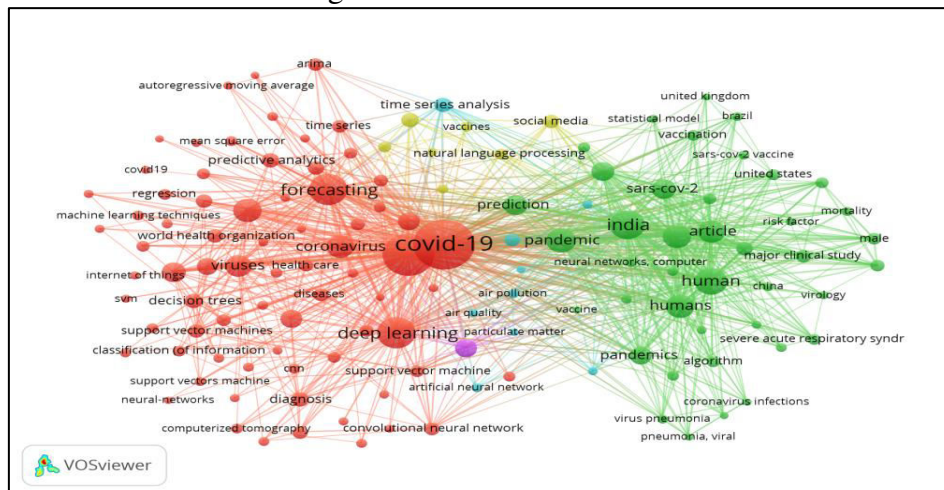


Figure 4. Co-occurrences of keywords

5. CONCLUSION

The research on COVID-19 prediction trends in India, employing bibliometric analysis, provides valuable insights crucial for enhancing pandemic preparedness. Firstly, it aids in identifying emerging trends by comprehending the methodologies and technologies employed in COVID-19 prediction studies, guiding future research endeavors. The identification of key researchers and institutions fosters collaboration, paving the way for a more cohesive and effective response to future pandemics. Highlighting highly cited papers offers a roadmap for policymakers and healthcare professionals to implement successful strategies, while recognizing regional variations informs the development of tailored approaches for diverse areas. The evaluation of prediction models contributes to the refinement of future infectious disease models, and acknowledging institutional excellence plays a pivotal role in shaping preparedness plans and facilitating information dissemination. Additionally, keyword analysis ensures

policymakers stay abreast of prevalent themes, and network visualizations provide a comprehensive overview, identifying key nodes for strategic resource allocation. Lastly, learning from the success of highly cited documents offers valuable insights for future planning efforts.

1. 383 citations publications(CP) and 5651 citations(C) is identified.
2. 27 publications have above 50 citations(C).
3. Article with 227 citations having a citation density of 75.66 published in 2020.
4. Meenu Gupta affiliated with Chandigarh University; Mohali, India has topped the most frequent authors.
5. Amity University, Noida, India published the highest 20 documents with citations of 86 and the citations per publication is 4.3.

The future scope of this research involves leveraging the gained insights to enhance global preparedness for pandemics. Building on identified trends and successful strategies, future research can focus on refining prediction models, incorporating advancements in technologies, and fostering international collaboration. Policymakers can utilize the regional insights to develop adaptable and targeted response plans, considering the diverse landscapes of different areas. Institutions recognized for excellence can become hubs for knowledge exchange and training programs, further strengthening the global network for pandemic preparedness. Additionally, the keyword analysis and network visualizations offer a foundation for continuous monitoring of emerging themes and key players in the field, ensuring that preparedness efforts remain dynamic and proactive. Continuous evaluation and adaptation of strategies based on the evolving landscape of infectious diseases will be essential for a resilient and effective global response to future pandemics.

ACKNOWLEDGEMENTS

This preprint of the paper has been submitted on SSRN. <http://dx.doi.org/10.2139/ssrn.4754962>

REFERENCES

- [1] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, “How to conduct a bibliometric analysis: An overview and guidelines,” *J. Bus. Res.*, vol. 133, no. March, pp. 285–296, 2021, doi: 10.1016/j.jbusres.2021.04.070.
- [2] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, and V. Singh, “A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images,” *Chaos, Solitons and Fractals*, vol. 140, p. 110190, 2020, doi: 10.1016/j.chaos.2020.110190.
- [3] K. Nikolopoulos, S. Punia, A. Schäfers, C. Tsinopoulos, and C. Vasilakis, “Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions,” *Eur. J. Oper. Res.*, vol. 290, no. 1, pp. 99–115, 2021, doi: 10.1016/j.ejor.2020.08.001.
- [4] R. S. Jyotir and M. Chatterjee, “A machine learning forecasting model for COVID-19 pandemic in India,” *Stoch. Environ. Res. Risk Assess.*, vol. 34, no. 7, pp. 959–972, 2020, doi: 10.1007/s00477-020-01827-8.
- [5] P. Arora, H. Kumar, and B. K. Panigrahi, “Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India,” *Chaos, Solitons and Fractals*, vol. 139, p. 110017, 2020, doi: 10.1016/j.chaos.2020.110017.
- [6] P. Wang, X. Zheng, J. Li, and B. Zhu, “Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics,” *Chaos, Solitons and Fractals*, vol. 139, p. 110058, 2020, doi: 10.1016/j.chaos.2020.110058.
- [7] G. Marques, D. Agarwal, and I. de la Torre Díez, “Automated medical diagnosis of

- COVID-19 through EfficientNet convolutional neural network,” *Appl. Soft Comput. J.*, vol. 96, p. 106691, 2020, doi: 10.1016/j.asoc.2020.106691.
- [8] S. Gupta, G. S. Raghuvanshi, and A. Chanda, “Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020,” *Sci. Total Environ.*, vol. 728, p. 138860, 2020, doi: 10.1016/j.scitotenv.2020.138860.
- [9] S. Shastri, K. Singh, S. Kumar, P. Kour, and V. Mansotra, “Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study,” *Chaos, Solitons and Fractals*, vol. 140, p. 110227, 2020, doi: 10.1016/j.chaos.2020.110227.
- [10] R. Salgotra, M. Gandomi, and A. H. Gandomi, “Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming,” *Chaos, Solitons and Fractals*, vol. 138, 2020, doi: 10.1016/j.chaos.2020.109945.
- [11] R. Wang, J. Chen, K. Gao, and G. W. Wei, “Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries,” *Genomics*, vol. 113, no. 4, pp. 2158–2170, 2021, doi: 10.1016/j.ygeno.2021.05.006.
- [12] M. Gupta, R. Jain, S. Taneja, G. Chaudhary, and M. Khari, “Real-time measurement of the uncertain epidemiological appearances of COVID-19 infections,” *Appl. Soft Comput. J.*, vol. 101, p. 107039, 2021, doi: 10.1016/j.asoc.2020.107039.
- [13] M. Gupta *et al.*, “AI-enabled COVID-19 outbreak analysis and prediction: Indian states vs. union territories,” *Comput. Mater. Contin.*, vol. 67, no. 1, pp. 933–950, 2021, doi: 10.32604/cmc.2021.014221.