

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/368364232>

# Machine Learning and Prediction of Infectious Diseases: A Systematic Review

Article · February 2023

DOI: 10.3390/make5010013

CITATIONS

0

READS

49

5 authors, including:



**Omar Enzo Santangelo**

Azienda Socio Sanitaria Territoriale di Lodi

106 PUBLICATIONS 470 CITATIONS

[SEE PROFILE](#)



**Vito Gentile**

Università degli Studi di Palermo

37 PUBLICATIONS 229 CITATIONS

[SEE PROFILE](#)



**Domiziana Giordano**

Università degli Studi di Palermo

24 PUBLICATIONS 113 CITATIONS

[SEE PROFILE](#)



**Fabrizio Cedrone**

Local Health Authority of Pescara

71 PUBLICATIONS 389 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



String theory applied to the prediction and epidemiology of infectious diseases [View project](#)



Promozione della salute nel Sito di Interesse Nazionale di Milazzo: participatory assessment della percezione del rischio tramite focus group [View project](#)



Systematic Review

# Machine Learning and Prediction of Infectious Diseases: A Systematic Review

Omar Enzo Santangelo <sup>1,\*</sup> , Vito Gentile <sup>2</sup>, Stefano Pizzo <sup>3</sup> , Domiziana Giordano <sup>4</sup> and Fabrizio Cedrone <sup>5</sup>

<sup>1</sup> Regional Health Care and Social Agency of Lodi, ASST Lodi, 26900 Lodi, Italy

<sup>2</sup> Synbrain Srl, 90128 Palermo, Italy

<sup>3</sup> Hospital Management, AOUP “G. Rodolico-San Marco” of Catania, 95123 Catania, Italy

<sup>4</sup> Hospital Hygiene Unit, AOUP “G. Rodolico-San Marco” of Catania, 95123 Catania, Italy

<sup>5</sup> Hospital Management, Local Health Unit of Pescara, 65122 Pescara, Italy

\* Correspondence: omarenzosantangelo@hotmail.it

**Abstract:** The aim of the study is to show whether it is possible to predict infectious disease outbreaks early, by using machine learning. This study was carried out following the guidelines of the Cochrane Collaboration and the meta-analysis of observational studies in epidemiology and the preferred reporting items for systematic reviews and meta-analyses. The suitable bibliography on PubMed/Medline and Scopus was searched by combining text, words, and titles on medical topics. At the end of the search, this systematic review contained 75 records. The studies analyzed in this systematic review demonstrate that it is possible to predict the incidence and trends of some infectious diseases; by combining several techniques and types of machine learning, it is possible to obtain accurate and plausible results.

**Keywords:** machine learning; epidemiology; medical informatics computing; vaccine-preventable diseases; infectious diseases; COVID-19



**Citation:** Santangelo, O.E.; Gentile, V.; Pizzo, S.; Giordano, D.; Cedrone, F. Machine Learning and Prediction of Infectious Diseases: A Systematic Review. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 175–198. <https://doi.org/10.3390/make5010013>

Academic Editors: Saeed Mian Qaisar, Humaira Nisar and Abdulhamit Subasi

Received: 7 December 2022

Revised: 20 January 2023

Accepted: 30 January 2023

Published: 1 February 2023



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

## 1. Introduction

### 1.1. Burden of Infectious Diseases

Infectious diseases are now an increasing part of the global health burden: every year, millions of deaths are due to infectious diseases [1]. To be able to develop protective measures against infectious diseases, it is crucial to identify the factors of disease diffusion in order to apply control and prevention measures [2]. Identifying the factors of disease diffusion could allow for making predictions that can be useful for policymakers to make appropriate decisions on, for example, the purchase of vaccines, public awareness campaigns and training programs for health professionals [3,4].

### 1.2. Machine Learning Applied to Infectious Diseases—Overview

Machine-learning algorithms can contribute to the control of infectious diseases by helping to both spatially and temporally predict the evolution and spread of infectious diseases [5]. Machine-learning algorithms are capable of analyzing large, complex data sets and identifying patterns and trends that may be difficult for humans to detect. This makes them well suited for the prediction of infectious diseases, which often involve multiple factors such as population demographics, environmental conditions, and individual behaviors. In recent years, many studies have applied machine-learning techniques to the prediction of infectious diseases, and the results have been promising. One of the key challenges in using machine learning for disease prediction is the availability of high-quality, comprehensive data. Infectious disease surveillance systems often collect data on a variety of factors, including the number of reported cases, the locations of outbreaks, and the demographics of infected individuals. However, these data are often incomplete, biased, or noisy, which can affect the performance of machine-learning models. Additionally, many infectious diseases

have long incubation periods, which means that data on past outbreaks may not accurately reflect current conditions. To overcome these challenges, researchers have employed a range of machine-learning algorithms, including decision trees, random forests, support vector machines, and deep-learning networks. These algorithms have been applied to a variety of data sets, including electronic health records, genomic data, and social media posts. In general, the results of these studies have shown that machine-learning algorithms can accurately predict the spread and onset of infectious diseases, with performance comparable with or better than traditional statistical methods. Machine-learning algorithms have shown great promise in predicting the spread and onset of infectious diseases. For example, some studies have used machine learning to forecast the number of cases of a particular disease in a given region on the basis of historical data and current conditions [6]. Others have used machine learning to identify the most likely sources of an outbreak on the basis of the genetic makeup of the pathogen and the patterns of infection [7,8]. Still others have used machine learning to predict the likelihood of an individual contracting an infectious disease on the basis of their personal characteristics and behaviors [9]. Overall, the use of machine learning in the prediction of infectious diseases is a promising area of research, with potential applications in public health, epidemiology, and clinical practice. However, there are also significant challenges and limitations to using machine learning in this context, including the need for high-quality data, the complexity of the underlying phenomena, and the potential for bias and overfitting [10]. Despite the promising results of these studies, there are also significant challenges and limitations to using machine learning in the prediction of infectious diseases. One of the main challenges is the need for high-quality data, which are often difficult to obtain owing to issues such as missing values, incomplete records, and varying data formats. Additionally, the complexity of the underlying phenomena, such as the transmission of infectious diseases, can make it difficult to develop accurate models. Finally, there is a potential for bias and overfitting in machine-learning models, which can lead to inaccurate predictions [10]. In this systematic review, we will explore the current state of the art in the use of machine learning for the prediction of infectious diseases. We will focus on recent, high-quality studies that have applied machine-learning techniques to real-world data and have evaluated the performance of these models in predicting the spread and onset of infectious diseases. We will also discuss the challenges and limitations of using machine learning in this context and provide insights into future directions for research in this area.

### 1.3. Aim

The aim of the study is to show whether it is possible to predict infectious disease outbreaks early, by using machine learning.

## 2. Materials and Methods

### 2.1. Search Strategy and Data Sources

The Cochrane Collaboration [11] and the meta-analysis of observational studies in epidemiology (MOOSE) guidelines [12] were followed in order to conduct the current systematic review. The preferred reporting items for systematic reviews and meta-analyses [13] guidelines [14] were used to report the process and the results. A bibliographic search was conducted on 9 November 2022, on the Scopus and PubMed/MEDLINE databases, combining keywords by using the Boolean operators “AND”, “OR” and “NOT”. The search strategy is reported in Supplementary Table S1. No time filter was used. Given the innovative nature of the study and its recent field of application halfway between medicine, epidemiology, and information technology, it was not always possible to apply all the items of the PRISMA checklist (more details are given in the study limitations section).

### 2.2. Inclusion and Exclusion Criteria

Studies had to meet the following criteria to be considered eligible: (i) language: written in English; (ii) population: human; (iii) interventions: machine learning; (iv) comparators/control:

infectious diseases; (v) outcomes: prediction/forecasting outbreaks infectious diseases; (vi) type of study: epidemiologic studies (case-control, cross-sectional, or cohort studies).

Exclusion criteria were as follows: (i) articles not published in English; (ii) not human; (iii) full text not available; (iv) interventions: not about machine learning; (v) comparators/control: not about infectious diseases; (vi) outcomes: not about prediction/forecasting outbreaks infectious diseases; (vii) type of study: review article, meta-analysis, trial, expert opinion, commentary, editorial, case report, letter to the editor, or book chapters. See Supplementary Table S2 where the detailed description of the inclusion/exclusion criteria is reported.

### 2.3. Selection Process and Data Extraction

Titles and abstracts of manuscripts found using the search strategy and those retrieved from additional sources were independently assessed by two reviewers (D.G. and C.F.). Subsequently, the same authors assessed the eligibility of the articles and independently reviewed the full downloaded text. When there was an unresolved disagreement between the two evaluators, the discussion was resolved by discussing the case with a senior reviewer (O.E.S.). Full texts were downloaded only for potentially eligible studies.

Data extraction was conducted only for those articles that met all the inclusion criteria, and it was performed using a predefined and prepiloted spreadsheet elaborated in Microsoft Excel for Windows. The extracted data included the author, publication year, study period, country where the study was conducted, disease, data source, model and/or techniques, aim, main results, accuracy/best model, space/time resolution, order of magnitude modeled populations, funds, and conflicts of interest.

### 2.4. Strategy for Data Synthesis

By following the PRISMA 2020 guidelines, a flowchart (Figure S1) was created showing the number of references at each stage of the review process [15]. Summary tables were created showing the qualitative results of the literature. A full report was produced; in this, there is a general overview of the main findings of the review.

### 2.5. Critical Appraisal

A critical evaluation of the articles using the Newcastle–Ottawa scale (NOS) was independently carried out by two authors (O.E.S. and D.G.) [16]; this was a bias-risk assessment tool for observational studies that can assign up to nine points for the lowest risk of bias in three domains: (i) study group selection; (ii) comparability; and (iii) assessment of exposure and outcomes for case-control and cohort studies, respectively.

An adapted version of the NOS was used to assess cross-sectional studies [17]. According to these criteria and on the standard cutoff used in the previous literature [18,19], studies were classified as being of high, moderate, or low quality when their NOS score was  $\geq 7$ , 4–6, and  $\leq 3$ , respectively.

## 3. Results

### 3.1. Literature Search

First, 375 and 333 records were found on Scopus and PubMed/MEDLINE, respectively, and the total was 708. Second, 89 records were eliminated because they were duplications. In the end, 619 records were evaluated for admissibility. By evaluating the title and abstract, 537 records were deleted because the topic was not related ( $n = 530$ ), the articles were not original ( $n = 3$ ), they were not written in English ( $n = 3$ ), and one was not about humans ( $n = 1$ ). The full text of the 82 records was downloaded, and 7 records were excluded with reasons following an in-depth assessment. At the end of the process, 75 records were included in our review [20–94]. Figure S1 shows the selection flowchart. There was a 0.7% disagreement among the authors during the first screening. Table 1 lists the characteristics of the included studies in alphabetical order by author. The articles appear from Absar N [20] to Zhong R [73], and the proceedings/conference papers appear from Ajith A [74] to the end.

**Table 1.** Qualitative characteristics of included studies, reported in alphabetical order. From the study of Ajith A [74] the proceedings/conference papers.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Absar N. [20]	2022	Bangladesh	March 2020–August 2021	COVID-19	Health division of the government of the Republic of Bangladesh	LSTM	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	LSTM shows a high accuracy for predicting cases, shows a medium–low accuracy especially in the long term for predicting death cases	17-month forecast divided daily	Thousands of cases	No	None
Adiga A. [21]	2021	US	August 2022–January 2021	COVID-19	Centers for Disease Control	LSTM, BMA, AR, ARIMA, EnKF, SEIR	Predicting the trend of COVID-19 pandemic	It is possible to make a more accurate prediction by integrating machine learning with other statistical and mathematical models	LSTM which is trained over short observation windows, is able to learn accurately the sharp rise or drop in cases relatively quickly, and start obtaining high weights in the subsequent weeks	5-month forecast divided weekly	Thousands of cases	Yes	None
Ahmad H.F. [22]	2021	KSA, Kuwait, Bahrain, UAE	March 2020–January 2021	COVID-19	Kaggle. Novel Corona Virus 2019 Data Set *	ARIMA, SIR, bi-LSTM, LRM, SVR	Predicting the trend of COVID-19 pandemic	Long-term outbreaks of infectious diseases cannot be predicted	Bi-LSTM is the best model with the highest accuracy	9-month forecast divided daily	Thousands of cases	No	None
Ak Ç. [23]	2020	Turkey	January 2004–December 2017	Crimean-Congo hemorrhagic fever (CCHF)	Ministry of Health of Turkey	Gaussian process regression (GPR)	Predicting Crimean-Congo hemorrhagic fever in 2016 and 2017	The model predicts annual cases in 2016 e 2017	Annual predictions for 2016 and 2017 are accurate, but the predictions for individual provinces are not as much accurate	24-month forecast divided monthly; identify the province where there will be more cases	Hundreds of cases	Yes	None
Ak Ç. [24]	2018	Turkey	January 2004–December 2015	Crimean-Congo hemorrhagic fever (CCHF)	Ministry of Health of Turkey	Gaussian process regression (GPR), BRT, RFR	Predicting Crimean-Congo hemorrhagic fever	Gaussian process formulation obtained better results than two frequently used standard machine-learning algorithms (i.e., random forests and boosted regression trees) under temporal, spatial, and spatiotemporal prediction scenarios	GPR algorithm did a better job than RFR and BRT algorithms by predicting CCHF case counts more accurately	24-month forecast divided monthly; identify the province where there will be more cases	Hundreds of cases	Yes	None
Alsmadi M.K. [25]	2022	Jordan	2020–2021	COVID-19	John Hopkins University **; Worldometer ***	Modified SEIR, DNN	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The DNN is more efficient and accurate compared with the traditional prediction methods, decision trees, and linear regression	3-month forecast divided daily	Thousands of cases	Yes	Not reported
Ardabili S.F. [26]	2020	Italy, German, China, US, Iran	January–March 2020	COVID-19	Worldmeter ***	MLP, GA, PSO, GWO, ANFIS, SIR, SEIR	Predicting the trend of COVID-19 pandemic	Outbreak prediction can be realized by integrating machine-learning and SEIR models	Integration of machine learning and SIR/SEIR models is suggested to enhance the existing standard epidemiological models in terms of accuracy and longer lead time	3-month forecast divided daily	Thousands of cases	Yes	None
Asfahan S. [27]	2020	South Korea	January–March 2020	COVID-19	Korea’s center for disease control	PROPHET (open-source automated machine learning)	Predicting the trend of COVID-19 pandemic	The difference between predicted and observed values ranged from 4% to 12% in naïve population and in short term	MAPE index of authors model for 1 week was 7.42%, which is indicative of a highly accurate forecasting model	1-month forecast divided daily	Thousands of cases	Not reported	None
Bagheri H. [28]	2020	Iran	2010–2018	Brucellosis	Health Surveillance System of Iran	RBF, MLP	Predicting human brucellosis cases	The model could be effectively used in predicting infectious disease	RBF is a more common type of neural network learning that responds to a limited section of the input space; it has a faster and more accurate and yet simpler network structure compared with other neural networks, while the MLP is more generalizable	18-month forecast divided monthly; identify the province where there will be more cases	Dozens of cases	Yes	None

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Balogh A. [29]	2022	European Union	2020–2022	COVID-19	Global COVID-19 Trends and Impact Survey, Johns Hopkins University **, Worldmeter ***, European Centre for Disease Prevention and Control, National Centers for Environmental Information, Eurostat	RF	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Not reported	16-month forecast divided daily	Thousands of cases	Yes	None
Benedum C.M. [30]	2020	Perù, Puerto Rico, Singapore	1990–2016	Dengue	Passive surveillance systems for Peru and Puerto Rico, Ministry of Health for Singapore	RF, RF-UFA, ARIMA, LASSO	Predicting weekly dengue cases	For near-term predictions of weekly case counts and when using surveillance data, ML models had 21% and 33% less error than regression and time series models, respectively	When using dengue surveillance, population, temporal, and weather data as model inputs, RF was more accurate than both Poisson regression and ARIMA models, for near-term predictions, while the ARIMA model performed best for long-term predictions; when predicting dengue outbreaks, RF-UFA outperformed both RF and logistic regression models when using only population, temporal, and weather data as model inputs	48-month forecast divided weekly	Thousands of cases	Yes	None
Chaurasia V. [31]	2022	World	January 2020–May 2020	COVID-19	WHO	ARIMA, Holt linear trend method, Holt's Winter method, naive method, simple average, moving average, single exponential smoothing	Predicting the trend of COVID-19 pandemic	The models were to be adjusted, often not adequate.	The naïve method was best suited as described, better than all other methods	3-month forecast divided daily	Milions of cases	Not reported	None
Chen J. [32]	2018	Japan, Taiwan, Thailandia, Singapore	2001–2014	Dengue, malaria, chickenpox, hand foot and mouth disease	National Institute of Infectious Diseases for Japan, Taiwan National Infectious Disease Statistics System, Ministry of Health of Singapore,	LASSO models	Predicting the trend of dengue, malaria, chickenpox, hand foot and mouth disease	The model could be effectively used in predicting infectious disease	For LASSO, the models used for prediction, including different sets of predictors, have varying effects in different situations; short-term predictions generally perform better than longer term predictions	Cases in the next 14 days; identify the province/states where there will be more cases	Thousands of cases	Yes	None
Chimmula V.K.R. [33]	2020	Canada	January–March 2020	COVID-19	Johns Hopkins University **, Canadian Health Authority	LSTM, deep learning	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The LSTM model has an accuracy of about 93%	3-month forecast divided daily	Thousands of cases	Yes	None
Dash S. [34]	2021	Brazil, US, India, France, UK, Russia	April 2020–April 2021	COVID-19	Johns Hopkins University **, WHO, CDC, COVID19 India °	ARIMA models	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The model achieved 85% accuracy for all the countries and all states of India	2-month forecast divided daily	Thousands of cases	Yes	Not reported

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Dixon S. [35]	2022	Australia, Israel, Norway, Sweden, US, Japan	2008–2019	Campylobacteriosis, Q-fever, typhoid	WHO, epi-archive.bsvgateway.org, gadm.org, NASA, diva-gis.org, United Nations, naturalearth-data.com	RF, XGB, MLP, ARIMA, ARIMAX, GLARMA, SARIMA	Predicting of campylobacteriosis, Q-fever, typhoid outbreaks	The model could be effectively used in predicting infectious disease	The XGB models performed the best for all diseases, and in general, tree-based ML models performed the best when looking at data splits	Divided daily, time interval varies depending on the disease	Hundreds of cases	Yes	None
González-Bandala D.A. [36]	2020	Mexico	2002–2019	Acute respiratory infection (ARI)	Mexican Health Ministry	SoS, FFNN	Predicting of ARI	The model could be effectively used in predicting infectious disease	The results show that the combination of different data analysis techniques (FFNN, SoS, and smoothed endemic channels) can provide an accurate prediction for ARI data 1 week in advance; the final model could be used, along with the endemic channels, to detect possible outbreaks	12-month forecast divided weekly	Thousands of cases	Yes	None
Goo T. [37]	2021	South Korea	January 2020–February 2021	COVID-19	Kaggle, Novel Corona Virus 2019 Data Set *, Ministry of Health of South Korea	GBM, LSTM, SEIR, local linear regression (LLR), negative binomial (NB), segment Poisson	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	LLR, GBM, SEIR, NB, and LSTM, separately, performed well in the forecasting of the pandemic situation of the country	2-month forecast divided daily	Thousands of cases	No	Declared
Guo Y. [38]	2020	China	July 2015–December 2017	Hepatitis E	Shandong Center for Disease Control and Prevention (SCDC)	ARIMA, SVM, LSTM	Predicting incidence of hepatitis E	The model could be effectively used in predicting infectious disease	LSTM is the most suitable for predicting hepatitis E monthly incidence and cases number	30-month forecast divided monthly; identify the province where there will be more cases	Dozens of cases	Yes	None
Haq I. [39]	2022	Bangladesh	April 2020–August 2021	COVID-19	Institute of Epidemiology Disease Control and Research (IEDCR) of Bangladesh, Worldometer **, Directorate General of Health Services of Bangladesh	SIR, PROPHET (an open-source automated machine learning)	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	No data on accuracy, but the authors say it appears that the PROPHET algorithm is appropriate for pandemic data with a growing trend	12-month forecast divided daily	Thousands of cases	Not reported	None
Kamana E. [40]	2022	China	January 2004–December 2016	Malaria	Chinese Center for Disease Control and Prevention	XGB, LSTM, LSTMSeq2Seq	Predicting incidence of Malaria	The LSTMSeq2Seq model significantly improved the prediction of malaria re-emergence according to the influence of climatic factors	The LSTMSeq2Seq model achieved an average prediction accuracy of 87.3%	24-month forecast divided monthly; identify the province where there will be more cases	Thousands of cases	Yes	None
Katragadda S. [41]	2022	US	March–December 2020	COVID-19	Corona Data Scraper open-source project	LR, SVR, K-nearest neighbor regression, MLP, RF, XGB	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The prediction accuracy improved by 33.78% for the whole duration of the pandemic in 2020 (March–December) when visitor mobility was used in the forecasting model; the prediction accuracy improved by 33.78% for the whole duration of the pandemic in 2020 (March–December) when visitor mobility was used in the forecasting model	Cases in the next 14 days; identify the province/states where there will be more cases	Thousands of cases	Yes	None



Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Ketu S. [42]	2021	World, China, India, Italy	December 2019–June 2020	COVID-19	WHO	Gaussian process regression, MTGP, LR, SVR, RF	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Enhanced multitask Gaussian process regression (MTGP) model on the COVID-19 outbreak forecasting is better than the traditional forecasting model	6-month forecast divided bi-weekly	Milions of cases	Not reported	Not reported
Kim J. [43]	2021	237 different Countries	January–December 2019	Multiple infectious diseases	Media articles, Medisys	SVM, SSL, DNN	Predicting incidence of infectious diseases	The model could be effectively used in predicting infectious diseases	SSL shows outstanding performance compared to other two models, SVM and DNN also show reasonable performance, showing average accuracy over 0.7	Cases in the next 3 months; identify the province/states where there will be more cases	Milions of cases	Yes	None
Kim J. [44]	2019	Hong Kong	January 2004–January 2018	Influenza-like illness (ILI)	Media articles, Centre for Health Protection (CHP) of Hong Kong	SVM	Predicting weekly incidence of ILI	The model could be effectively used in predicting infectious disease	The prediction result using news text data with SVM exhibited a mean accuracy of 86.7%	48-month forecast divided weekly	Thousands of cases	Yes	None
Kumar S.L. [45]	2021	US, India, Brazil, Russia	Until 25 November of 2020	COVID-19	World in Data by University of Oxford	LR, ARIMA, RF, LSTM	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	ARIMA outperformed linear regression and random forest in terms of accuracy prediction of test data; ARIMA and LSTMs were compared again with the death-forecasting task, in which LSTMs were able to provide very high accuracy in comparison	6-month forecast divided daily	Thousands of cases	Not reported	Not reported
Lmater M.A. [46]	2021	Belgium, Morocco, Netherlands, Russia	February–November 2020	COVID-19	Worldmeter ***	SIDR model	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Not reported	6-month forecast divided daily	Thousands of cases	No	None
Lu F.S. [47]	2018	US	September 2012–May 2017	Influenza-like illness (ILI)	Boston Public Health Commission, Google searches, Twitter posts, FNY mobile app reports, EHRs	ARGO	Predicting weekly incidence of ILI	It is possible to forecast influenza 1 week ahead of the current date	Ensemble-based methods incorporating information from diverse models that are based on multiple data sources, including ARGO, produced the most robust and accurate results; the observed Pearson correlations between our out-of-sample flu activity estimates and those historically reported by the BPHC were 0.98 in nowcasting influenza and 0.94 in forecasting influenza 1 week ahead of the current date	12-month forecast divided daily	Thousands of cases	Yes	Declared
Marra A.R. [48]	2020	US	2015–2017	<i>Clostridiodes difficile</i>	EHRs	LR, FR, naïve Bayes, K-nearest neighbor, MLP, Lib SVM, decision tree (J48), AdaBoost (M1), bagging, radial basis function classifier	Predicting incidence of <i>Clostridiodes difficile</i>	Multiple machine-learning models yielded only modest results in a real-world population	Logistic regression, random forest and naïve Bayes models yielded the highest performance: 0.6	Not applicable	Dozens of cases	No	None
Meng D. [49]	2021	China	January 2009–December 2017	Hand, foot, and mouth disease	The Dara-Center of China Public Health Science, China Meteorological Data Service Centre, Tencent Location Big Data, The 2018 China Statistical Yearbook	RF, XGB	Predicting the trend of hand, foot, and mouth disease	The model could be effectively used in predicting infectious disease	The prediction capability of XGBoost model was better than that of random forest model from the overall perspective	2-month forecast divided daily; identify the province where there will be more cases	Thousands of cases	Yes	None



Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Mohammad Masum A.K. [50]	2020	Bangladesh	May–June 2020	COVID-19	Institute of Epidemiology Disease Control & Research of Bangladesh	LSTM, ML, RFR, SVR	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	LSTM is perfect fitted of real-time analysis; LSTM is best mod	1-month forecast divided daily	Thousands of cases	Not reported	Not reported
Murphy C. [51]	2021	Spain	January 2020–December 2021	COVID-19	COVID-19 en España <a href="https://cneccovid.isciii.es">https://cneccovid.isciii.es</a> (accessed on 29 January 2023), Observatorio del Transporte y la Logística en España <a href="https://observatoriodeltransporte.mitma.gob.es/estudio-experimental">https://observatoriodeltransporte.mitma.gob.es/estudio-experimental</a> (accessed on 29 January 2023)	GNN, MLE	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	GNN provides more-accurate predictions than the MLE in general and across all degrees	14-month forecast divided daily	Thousands of cases	Yes	None
Nguyen V.H. [52]	2022	Vietnam	1997–2016	Dengue	National Institute of Hygiene and Epidemiology of Vietnam, Vietnam Institute of Meteorology, Hydrology and Environment	LSTM-ATT, LSTM, weather-based dengue fever forecasting, XGB, SARIMA, Poisson regression, SVR, SVR-L, CNN, TF	Predicting incidence of dengue	The model could be effectively used in predicting infectious disease	LSTM-ATT displayed the highest performance	Identify the province where there will be more cases	Hundreds of cases	Yes	None
Niraula P. [53]	2022	Spain	March 2020–February 2021	COVID-19	Open data portal of Castilla-Leon, Barcelona Supercomputing Center	LSTM, Poisson regression	Predicting the trend of COVID-19 pandemic	The results show that a Bayesian model informed by a neural network method is generally able to predict the number of cases of COVID-19 in both space and time	Not reported	3-month forecast divided daily; identify the province where there will be more cases	Thousands of cases	Yes	None
Nsoesie E.O. [54]	2021	Cameroon	2012–2018	Influenza-like illness (ILI)	Google searches for influenza symptoms	ARIMA, SVM, RF, multivariable linear regression	Predicting weekly incidence of ILI	The model could be effectively used in predicting infectious disease	RF and SVM had the highest average R2 (0.78 and 0.88, respectively) for predicting ILI per 100,000 persons at the country level	72-month forecast divided weekly; identify the province where there will be more cases	Thousands of cases	Yes	None
Patil S. [55]	2021	India	2009–2023	Dengue	Indian Meteorological Department, National Vector-Borne Disease Control Program	SVR, ARIMA, SARIMA, PROPHET, RF, ElasticNet regression, multiple linear regression, polynomial regression, decision tree regression	Predicting incidence of dengue	The model could be effectively used in predicting infectious disease	No model prevails over the other; the accuracy is quite high, but depending on the city, one model is better than the other	12-month forecast divided daily; identify the province where there will be more cases	Thousands of cases	Yes	None
Pourghasemi H.R. [56]	2020	Iran	March–June 2020	COVID-19	Iranian’s Ministry of Health and Medical Education	SVM, ARIMA	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The SVM model showing good forecast accuracy was used for mapping the outbreak risk of COVID-19; among the four kernel functions of SVM, RBF has been proved to generate high accuracy models; the generated model using SVM had a good predictive accuracy—0.786 and 0.799	4-month forecast divided daily; identify the province where there will be more cases	Thousands of cases	Yes	None

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Roster K. [57]	2022	Brazil	2014–2021	Dengue, zika, influenza, COVID-19	Notifiable Diseases Information System of Brazil	RF, TrAdaBoost, Neurol Network (NN)	Predicting incidence of dengue, zika, and COVID-19	The model could be effectively used in predicting infectious diseases	RF models performed better for zika forecasts, while NN models were more successful in predicting COVID-19; NN transfer models predicted the most similar target disease with greater accuracy, while TrAdaBoost and the direct RF model fared better for less similar target diseases	Not applicable	Thousands of cases	Yes	None
Saba T. [58]	2021	India, Iran, Greece, Bulgaria, China, Sweden, Netherlands, Iceland, Russia	January 2020–January 2021	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	RF, polynomial regression, SVR, GBR, KNN, decision tree, SARIMA, A'IMA, Holt's Winter	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	It is impossible to recommend a single approach to model and forecasting for all data sets in respect of obtained results; as the different data sets exhibited different trends, depending upon the size, nature, and type of the lockdown	14-month forecast divided weekly	Thousands of cases	Yes	None
Shaghaghi N. [59]	2021	US	Not reported	COVID-19	New York Times COVID-19 Data, Google Trends	LSTM, RNN	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	eVision has been able to achieve an accuracy of 89% for predicting the trend of the COVID-19 outbreak in the United States	Identify the province where there will be more cases	Thousands of cases	Not reported	Not reported
Shen C. [60]	2021	China	November 2019–March 2020	COVID-19	Weibo, Chinese Center for Disease Control and Prevention	Decision tree, extra trees, KNN, MLP, SVM, RF	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	SVM, RF, MLP, KNN, DT, and extra tree have accuracy better than 0.8	Identify the province where there will be more cases	Thousands of cases	Not reported	None
Shen L. [61]	2022	Europe	2008–2018	Brucellosis	European Center for Disease Control and prevention, World Animal Health Information System	Convolutional LSTM, LSTM, ARIMA	Predicting human brucellosis cases	The model could be effectively used in predicting infectious disease	The prediction results have shown that LSTM and ConvLSTM models have higher forecast precision	12-month forecast divided monthly; identify the province where there will be more cases	Dozens of cases	Yes	None
Shi Y. [62]	2016	Singapore	2001–2013	Dengue	Singapore's Ministry of Health	LASSO models	Predicting incidence of dengue	The model could be effectively used in predicting infectious disease	The model specifically optimizes predictive accuracy over a 3-month time horizon	12-month forecast divided weekly	Thousands of cases	Yes	None
Tiwari D. [63]	2022	World	January–May 2020	COVID-19	Kaggle *	Naïve Bayes, SVM, LR	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Naïve Bayes is more accurate than the other models in the study for predicting future COVID-19 trends	3-month forecast divided daily; identify the province where there will be more cases	Milions of cases	Not reported	None
Venkatramanan S. [64]	2021	US, Australia	2009–2018	Influenza	National Notifiable Disease Surveillance Syste of Australia Government, CDC FluView, The New Jersey Department of Health, Australian Bureau of Statistics, American Community Survey, Google,	Machine-learned anonymized mobility map	Predicting incidence of influenza	The model could be effectively used in predicting infectious disease	The model performs better during early weeks, especially before onset, but as the season progresses, its performance deteriorates in comparison to the other network models	3-month forecast divided weekly; identify the province where there will be more cases	Thousands of cases	Yes	None

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Verma H. [65]	2022	India	2020–2021	COVID-19	COVID-19India.org	Vanilla LSTM, stacked LSTM, ED_LSTM, Bi-LSTM, CNN, and hybrid CNN + LSTM	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The stacked-LSTM and hybrid CNN + LSTM models perform best relative to other models	3-week forecast divided daily	Thousands of cases	Yes	None
Wang H. [66]	2022	China	2010–2017	COVID-19	Influenza-Like Illness (ILI)	SEIRS, LSTM, spatial transmission network (STN)	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The results illustrated that the STN not only had good accuracy in forecasting performance but also indicated the spreading directions of infectious diseases among multiple regions, to a certain extent	5-month forecast divided weekly; identify the province where there will be more cases	Thousands of cases	Yes	None
Wang X. [67]	2022	US	April 2020–April 2021	COVID-19	Worldmeters <sup>***</sup> , <a href="https://www.google.com/covid19/mobility/">https://www.google.com/covid19/mobility/</a> (accessed on 29 January 2023)	GBM, ordinary differential equation (ODE) model	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The ODE model predicted the number of daily infected cases up to 35 days in the future, with an average mean absolute percentage error of 20.15%, which was further improved to 14.88% when combined with human mobility data	12-month forecast divided daily	Thousands of cases	Yes	None
Wang Y. [68]	2022	US, Brazil, India	2020–2021	COVID-19	WHO	PROPHET (an open-source automated machine learning), ARIMA, SARIMA	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The PROPHET model showed more accuracy in the daily COVID-19 new cases in the US; the ARIMA model is more suitable for predicting cases in Brazil and India	12-month forecast divided daily	Thousands of cases	Yes	None
Xu J. [69]	2020	China	January 2005–December 2018	Dengue	China National Notifiable Disease Surveillance System	LSTM	Predicting incidence of dengue	The model could be effectively used in predicting infectious disease	The LSTM model achieved superior performance in predicting dengue cases as compared with other previously published forecasting models	24-month forecast divided monthly	Thousands of cases	Yes	None
Xu Q. [70]	2017	Hong Kong	January 2014–December 2015	Influenza-like illness (ILI)	Hong Kong Centers for Health Protection, Google searches for influenza symptoms, Meteorological data are available from Hong Kong Observatory	GLM, LASSO, ARIMA, DL, FNN, Bayesian model averaging	Predicting incidence of ILI	The model could be effectively used in predicting infectious disease	DL with FNN remains the preferred method for predicting locations of influenza peaks	12-month forecast divided weekly	Thousands of cases	Yes	None
Yang Z. [71]	2020	China	Training on 2003 data of SARS	COVID-19	National Health Commission of China	Modified SEIR, LSTM	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The modified SEIR model is better than LSTM	2-month forecast divided daily; identify the province where there will be more cases	Thousands of cases	Yes	None
Zhang Y. [72]	2022	US	January–April 2020	COVID-19	Twitter	LR, KNN, SVM, Deep pyramid convolutional neural network, fine-tuning BERT	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Fine-tuning BERT as a tweet classification method and achieve a 0.98 and could to be an early warning on outbreak disease	3-month forecast divided daily; identify the province/States where there will be more cases	Thousands of cases	Yes	None

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Zhong R. [73]	2018	China	January 2017–September 2017	Hand, foot, and mouth disease	Shenzhen Health Information Center, Weather Underground app ( <a href="https://www.wunderground.com/">https://www.wunderground.com/</a> ) (accessed on 29 January 2023)	XGB	Predicting the trend of hand, foot, and mouth disease	The model could be effectively used in predicting infectious disease	Compared with the model only using the previous HFMD rate and temperature factors, the addition of the air-quality factors could make the model better by nearly 16.7%	6-month forecast divided daily	Thousands of cases	Yes	None
Ajith A. [74]	2020	Not reported	2014–2020	West Nile virus (WNV)	Kaggle *	RF, naïve Bayes classifier, adaptive boost	Predicting the trend of West Nile virus	The model could be effectively used in predicting infectious disease	Random forest is the best model for accurately predicting WNV cases	6-year forecast divided weekly	Thousands of cases	Not reported	Not reported
Andreas A. [75]	2020	World	2020	COVID-19	Worldmeters ***	Mathematical forecasting framework based on machine learning and the cloud computing system	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The proposed model has a R2 equal to 0.99	4-month forecast divided daily	Milions of cases	Yes	Not reported
Brock P.M. [76]	2019	Malaysia	2000–2014	Malaria	Case-control studies, global positioning system (GPS)	Boosted regression trees (BRT)	Predicting the outbreak of Malaria	The model could be effectively used in predicting infectious disease	The BRT models to predict <i>P. knowlesi</i> occurrence varied from an AUC of 0.55 (little better than a random model) to a maximum of 0.82	Identify the province where there will be more cases	Thousands of cases	Yes	None
Chumachenko D. [77]	2021	Ukraine, UK, Germany, Japan	2020–2021	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	LASSO	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The model shows more-accurate results when forecasting for 10 days or fewer	Not reported	Thousands of cases	Yes	Not reported
Chumachenko D. [78]	2021	Ukraine, UK, Germany, Japan	2020–2021	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	RF	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The RF method showed the most accurate result among K-nearest neighbor regression, linear regression, LASSO regression, ridge regression, and gradient boosting	Not reported	Thousands of cases	Yes	Not reported
Fan X.R. [79]	2022	China	January–February 2020	COVID-19	National Health Commission of the People's Republic of China	ISIR, MLR, XGBoost, LightGBM, StackCCPred (stacking based prediction of COVID-19 pandemic)	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The results positively proved that the proposed StackCCPred model outperformed the existing models for COVID-19 (R2 = 0.902)	1-week forecast divided daily	Thousands of cases	Yes	Not reported
Hasri H. [80]	2021	Malaysia	1–14 August 2021	COVID-19	Ministry of Health of Malaysia	LR, Holt's Winter	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Holt's Winter is more accurate than the LR model: LR model accuracy is 82%, Holt's Winter accuracy is 89%	2-week forecast divided daily	Thousands of cases	Yes	Not reported
Kolesnikov A.A. [81]	2019	Not reported	Not reported	Dengue	National Oceanic and Atmospheric Administration in the US Department of Commerce, Philippine Department of Health, CDC, International Society for Infectious Diseases, Landsat, Sentinel	RF, LSTM, XGB, SARIMA, LightGBM, LR, KNN, CatBoost, Keras	Predicting incidence of dengue	The model could be effectively used in predicting infectious disease	The most effective predictions were given by a mathematical model based on a combination of spatial analysis techniques (MGWR) and neural networks based on the LSTM architecture	Not reported	Thousands of cases	Not reported	Not reported

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Kumari P. [82]	2020	India	January–July 2020	COVID-19	Government of India	Artificial neural network	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The proposed model is found to be highly accurate in estimating the growth of COVID-19 related parameters	Not reported	Thousands of cases	Not reported	Not reported
Liu Z. [83]	2021	US, India, Brazil	January 2020–January 2021	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	Nonlinear autoregressive neural network (NAR), LR, ARIMA, SEIR	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	NAR dynamic neural network is better than the comparison model in the time prediction of the new crown epidemic, with the maximum error of 3.6% and the minimum error of −0.3%	2-month forecast divided daily	Thousands of cases	Yes	Not reported
Maaliw R.R. [84]	2021	Philippines, US, India, Brazil	March 2020–June 2021	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	ARIMA, S-LSTM	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The quantitative results show that the ensemble model outperforms stand-alone models of ARIMA and S-LSTM for a 15-day forecast accuracy of 93.50% (infected cases) and 87.97% (death cases)	1-week forecast divided daily	Thousands of cases	Yes	Not reported
Mahima Y. [85]	2020	World	Not reported	COVID-19	Kaggle *	LR, bagging regression, RF, KN	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The bagging regressor model gives the highest accuracy	3-month forecast divided daily; identify the province/states where there will be more cases	Thousands of cases	Not reported	Not reported
Mei W. [86]	2021	World	Not reported	COVID-19	Not reported	Time-variant relevance-based infected recovered extreme learning machine	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Authors say that the proposed method can achieve higher accuracy than existed methods	Not reported	Milions of cases	Yes	Not reported
Patayon U.B. [87]	2021	Philippines	March 2020–May 2021	COVID-19	Department of Health Zamboanga Peninsula–Center for Health Development	Vanilla LSTM, stacked LSTM, bidirectional LSTM, ConvLSTM	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	ConvLSTM trained using Adam and RMSProp delivers the best results as it closely adapted to the trend of actual data	1-month forecast divided daily	Thousands of cases	Yes	Not reported
Pickering L. [88]	2020	US	2018–2020	COVID-19	NASA Space Apps	SVR, multidimensional regression with interactions, stepwise regression method	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	All three methods predict a rise in cases similar to the actual rise in cases and are all able to predict to a certain degree the unexpected dip in cases on the 10th and 11th day of prediction	4-month forecast divided daily	Thousands of cases	Not reported	Not reported
Rohini M. [89]	2021	India	Not reported	COVID-19	Kaggle *	KNN, DT, SVM, RT	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The model developed using K-nearest neighbors (KNN) is effective, with a prediction accuracy of 98.34%	Not reported	Thousands of cases	Not reported	Not reported
Satu M.S. [90]	2021	World	2020	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	LR, PR, SVR, MLP, poly-MLP, PROPHET	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	PROPHET is the best model among those proposed: its R2 is equal to 1	Not reported	Thousands of cases	Not reported	Not reported
Sri S.B. [91]	2022	World	January 2020–November 2021	COVID-19	WHO	PROPHET	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	Through training, the model is able to predict the cases of the next 7 days	Not reported	Thousands of cases	Not reported	Not reported

Table 1. Cont.

Author [Ref.]	Publication Year	Country of Study	Study Period	Disease	Data Source	Model and/or Techniques	Aim	Main Results	Accuracy/Best Model	Space/Time Resolution	Order of Magnitude Modeled Populations	Funds	Conflicts of Interest
Wang H. [92]	2022	World	2020–2021	COVID-19	<a href="https://github.com/CSSEGISandData">https://github.com/CSSEGISandData</a> (accessed on 29 January 2023)	T-SIRGAN (susceptible–infected–recovered, generative adversarial networks), T-GAN, T-COVID, LSTM-SIRGAN, LSTM-GAN, LSTM-COVID, ARIMA, DT, SVM, KNN	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The authors' T-SIRGAN model has better accuracy than the others used as a comparison	4-month forecast divided daily; identify the states where there will be more cases	Thousands of cases	Yes	Not reported
Zhou Q. [93]	2020	World	January–July 2020	COVID-19	WHO, Baidu Baiken	Logistic model, ARIMA, SIR, SEIR	Predicting the trend of COVID-19 pandemic	The model could be effectively used in predicting infectious disease	The improved SEIR model is the best model and can more precisely predict the future development trend of the global epidemic	6-month forecast divided daily	Thousands of cases	Yes	Not reported
Zhang P. [94]	2022	China	January 2015–December 2019	Hand, foot, and mouth disease, hepatitis B	Xiamen Center for Disease Control and Prevention	Oriented attention model (OAM), AR, LSTM, gated recurrent unit, encoder-decoder, CNN, CNN-RNN, LSTM-attn, GRU-attn, ED-attn, CNN-attn and CNNRNN-attn	Predicting the trend of hand, foot, and mouth disease and hepatitis B	The model could be effectively used in predicting infectious disease	The self-attention significantly improves the predictive accuracy of all comparable methods; the MAE and RMSE values were decreased by 51.67% and 39.43% at most, respectively; the R2 is increasing by 52.99% at most	Not reported	Thousands of cases	Yes	Not reported

UK: United Kingdom, US: United States, KSA: Kingdom of Saudi Arabia, UAE: United Arab Emirates, LSTM: long short-term memory, Bi-LSTM: bidirectional LSTM, ED: encoder-decoder, LSTM-ATT: attention-enhanced LSTM, ARIMA: autoregressive integrated moving average, AR: autoregressive, EnKF: nonlinear ensemble Kalman Filter, SEIR: susceptible–exposed–infected–removed, SIR: susceptible–infected–recovered, SIDR: susceptible–infected–diagnosed–recovered, BMA: Bayesian model averaging, PR: polynomial regression, LRM: linear regression model, SVR: support vector regression, SVR-L: support vector regressor with linear kernel, SVM: support vector machine, DNN: deep neural network, MLP: multilayered perceptron, RF: random forest, ANFIS: adaptive neuro fuzzy inference system, GA: genetic algorithm, PSO: particle swarm optimization, GWO: gray wolf optimizer, RNN: recursive neural network, UFA: univariate flagging algorithm, SARIMA: seasonal ARIMA, SEIAR: susceptible–exposed–symptomatic infected–asymptomatic infected–removed, LASSO: least absolute shrinkage and selection operator algorithm, XGB: extreme gradient boosted, GLARMA generalized linear autoregressive moving averages, ARIMAX: ARIMA with exogenous variables, CNN: convolutional neural network, FFNN: feed forward neural network, SoS: sum of sines, DT: decision tree, GBM: tree-based gradient-boosting machine, MTGP: multitask Gaussian process regression, SSL: semi-supervised learning, TF: transformer, ARGO: autoregression with general online information, WHO: World Health Organization, N.A.: not available. \* Available online: <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset> (accessed on 29 January 2023); \*\* Available online: <https://coronavirus.jhu.edu/map.html> (accessed on 29 January 2023); \*\*\* Available online: <https://www.worldometers.info/coronavirus> (accessed on 29 January 2023); ° Available online: <https://www.covid19india.org/> (accessed on 29 January 2023).

### 3.2. Characteristics of Included Studies

Table 1 shows the characteristics of the included studies: the first part reports the studies published in international journals, the second part of Table 1 shows the studies extracted from the proceedings, the lines are highlighted in gray, and the authors are listed in alphabetical order. As for the studies excluding the proceedings, all studies were performed after 2016 [20–73]. Further, nine studies were carried out in China [38,40,49,60,66,69,71,73,94]; seven in the US [21, 41,47,48,59,67,71]; three in Bangladesh [20,39,50]; two each in Spain [51,53], Iran [28,56], Hong Kong [44,70], India [55,65], Turkey [23,24], and South Korea [27,37]; and the other studies in Jordan, Brazil, Cameroon, Canada, Mexico, Singapore, and Vietnam [25,33,36,52,54,57,62]. Finally, the remaining studies were multicentric.

Most of the studies were related to COVID-19 ( $n = 30$ ) [20–22,25–27,29,31,33,34,37,39, 41,42,45,46,50,51,53,56,58–60,63,65–68,71,72] and then influenza and influenza-like illness (ILI) [44,47,54,64,70]; for other details of the studies, see Table 1. No funds were reported by five studies [20,22,37,46,48], funds were not reported for nine studies [27,31,39,42,45,50,59,60,63], and funds were reported in the other studies ( $n = 40$ ). In two studies, the authors declared the conflicts of interest [37,47], and in six, the authors did not declare the conflicts of interest [25,34,42,45,50,59]. Furthermore, 21 proceedings/conference papers fulfilling the search criteria were included in this review. Of these, 17 reported on COVID-19 [75,77–80,82–93], one study West Nile virus [74], one malaria [76], one dengue [81], and one hand, foot, and mouth disease, hepatitis B [94]. Most of the studies were performed on worldwide data [75,85,86,90–93] and all were published after 2019 [74–94].

### 3.3. Quality Assessment

The quality of the 75 studies ranged from 6 to 9. The assessment revealed a medium–high quality level for cohort studies and cross-sectional ones. See Supplementary Table S3 for a complete overview that is based on the NOS checklist.

## 4. Discussion

Several infectious diseases are emerging and threatening the human health condition across the world. The burden of infectious diseases is certainly a planetary issue, annually causing millions of deaths [24]. Therefore, the study of infectious disease behavior has been a subject of scientific interest for many years; the early identification of emerging infectious disease outbreak patterns is critical and offers great advantages [36,43]. Indeed, as is evident from the studies under observation, accurate and reliable predictions of infectious diseases can be invaluable to public health organizations planning interventions to reduce or prevent disease transmission [38] and mitigate the negative impacts of diseases [35]. As seen by Ketu S. et al., the viral epidemic, in addition to exerting direct damage on people's lives, can affect a country's economy [42,43]. As reiterated by Roster K. et al., recent epidemic outbreaks, such as the COVID-19 pandemic and the Zika epidemic in Brazil, have demonstrated the importance and difficulty of accurately predicting new infectious diseases [57]. A lack of knowledge about new infectious diseases and their consequences, along with complicated social and governmental factors, may influence the spread of every newly emerging disease [42]. It is, in fact, essential to try to estimate the future movement and pattern of a new disease [39], so that preventive measures such as closing schools, shopping centers, and theaters; closing borders; suspending public services; and stopping travel can be quickly implemented [42]. However, because the epidemic spread of an infectious disease usually occurs sporadically and rapidly, it is not easy to predict whether an infectious disease will emerge and how. In addition, collecting data on a specific infectious disease is not always easy. Knowledge about the transmission paths of emerging diseases, the level and duration of immunity to reinfection, and other parameters needed to build realistic epidemiological models are often scarce. For these reasons, it is necessary to find appropriate and useful information sources and data and build up reliable prediction models with these [43,57]. Indeed, to develop increasingly effective control and prevention strategies, reliable computational tools that may help to understand disease dynamics



and predict future cases are needed. Policymakers can use these computational tools to make decisions that are more informed [24]. Several approaches have been proposed in the literature to produce accurate and timely predictions and potentially improve public health response [32]. Time series forecasting and machine learning, while less dependent on disease assumptions, require large amounts of data that may not be available in the early stages of an outbreak [57]. Modeling the spread of infectious diseases in space and time must account for complex dependencies and uncertainties. Machine-learning methods, especially neural networks, are useful for modeling these kinds of complex problems, even if they in some cases lack probabilistic interpretations [53]. Predicting the evolution of contagion dynamics is still an open problem, to which mechanical models offer only a partial answer. To remain mathematically or computationally tractable, these models must rely on simplifying assumptions, thus limiting the quantitative accuracy of their predictions and the complexity of the dynamics they can model [51]. Mathematical modeling is the most scientific technique for understanding the evolution of natural phenomena, including the spread of infectious diseases. Therefore, these modeling tools have been widely used in epidemiology to predict risks and inform decision-making [46]. While imperfect, these models offer an additional input for decision makers of infectious disease responses. These results could be useful insofar as informing decisions on planning, resource allocation, and social-distancing policies [37]. Deep learning offers a new and complementary perspective to build effective models of contagion dynamics on networks, as demonstrated by Murphy C. et al. [51]. The analysis found that models based on combining multiple machine-learning methods, incorporating information from different models that are based on multiple data sources, produced the most robust and most accurate results [47].

#### 4.1. Acute Respiratory Infection (ARI)

Acute respiratory infections (ARIs) are one of the main causes of morbidity and mortality in the world, particularly in children under 5 years and adults over 65 years [36]. González-Bandala et al. [36] propose a methodology that merges the predictions of a computational model with machine learning, a projection model, and a proposed smoothed endemic channel calculation. The predictions are made on weekly acute respiratory infection (ARI) data obtained from epidemiological reports in Mexico, along with the usage of key terms in the Google search engine. The results obtained with this methodology were compared with state-of-the-art techniques resulting in reduced root-mean-square percentage error (RMPSE) and maximum absolute percent error (MAPE) metrics, achieving a MAPE of 21.7%. The results show that the combination of different data analysis techniques (FFNN, SoS, and smoothed endemic channels) can provide an accurate prediction for ARI data 1 week in advance [36].

#### 4.2. Brucellosis

Brucellosis (Malta fever) is one of the most common zoonotic diseases and has long been one of the most important health concerns for humans and animals [61]. Two studies emerged in our review: one by Bagheri H. et al. [28], performed in Iran, and one by Shen L. et al. [61], performed in Europe. In both studies, the authors demonstrated that their models can predict brucellosis cases in humans; to demonstrate this, they used different models. Bagheri [28] compared RBF and MLP, stating that RBF is a more common type of neural network learning that responds to a limited section of the input space; it has a faster and more accurate and yet simpler network structure compared to other neural networks, while the MLP is more generalizable. Shen L. used LSTM, LSTM, and ARIMA as convolutional models, and the prediction results have shown that the LSTM and ConvLSTM models have higher forecast precision.

#### 4.3. Campylobacteriosis, Q-Fever, and Typhoid

Only one study addressed these diseases, and it was Dixon S. et al. [35], where multiple models were compared: RF, XGB, MLP, ARIMA, ARIMAX, GLARMA, and SARIMA. The

end result was that the XGB models performed the best for all diseases, and in general, tree-based ML models performed the best when looking at data splits [35]. According to the authors, this study demonstrated the power of ML approaches to incorporate a wide range of factors to more accurately forecast various diseases, regardless of location, than traditional statistical approaches [35].

#### 4.4. Chickenpox

Only one study addressed chickenpox; however, it was not dedicated only to this disease but rather took into consideration various infectious diseases with respiratory transmission [32]. The study by Chen et al. [32], using the LASSO model, showed that predictions made more than 4 weeks in advance were increasingly discrepant from the real scenario and that the prediction model was more accurate in capturing the epidemic but less sensitive to predicting the size of the epidemic, probably because the climatic variables have different levels of importance in the accuracy of the forecasts [32].

#### 4.5. *Clostridioides Difficile*

The study by Marra et al. [48] was the one that obtained the least accuracy. It was a study on *Clostridioides difficile*, and despite the large variability of the models used (LR, FR, naïve Bayes, K-nearest neighbor, MLP, Lib SVM, decision tree (J48), AdaBoost (M1), bagging, and radial basis function classifier), these machine-learning models have produced only modest results in a real-world population [32]. The logistic regression, random forest, and naïve Bayes models yielded the highest performance: 0.6 [32].

#### 4.6. Crimean-Congo Hemorrhagic Fever (CCHF)

Crimean-Congo hemorrhagic fever (CCHF) is a tick-borne viral infection usually transmitted by tick bites or through contact with tissues, blood, or other bodily fluids from infected people and animals [23,24]. In our systematic review, there were two studies conducted by the same author: Ak Ç [23,24]. The author demonstrated in both studies that Gaussian process formulation obtained better results than two frequently used standard machine-learning algorithms (i.e., random forests and boosted regression trees) under temporal, spatial, and spatiotemporal prediction scenarios. The Gaussian process was the best model to spatiotemporally show CCHF outbreaks [23,24].

#### 4.7. COVID-19

Most of the studies under observation concern the COVID-19 pandemic [20–22,25–27, 29,31,33,34,37,39,41,42,45,46,50,51,53,56,58–60,63,65–68,71,72]. Ever since the coronavirus pandemic (COVID-19) emerged in Wuhan, China, and was recognized as a global threat, several national and global studies have been conducted to try to predict the epidemic, with various levels of reliability and accuracy [34], as has been carried out in Korea [37] and India [65], among others. Academic researchers all around the world have proposed various predictive models to allow policymakers to make better decisions, to apply appropriate control measures [26,29,32,63], and to reduce the burden in hospitals [58]. Recently, new machine-learning approaches have been used to understand the dynamic trend of the COVID-19 spread, as was shown in the study by Verma H. et al., who used the temporal deep-learning architecture to predict COVID-19 cases in India [65]. In the study by Pourghasemi H.R. et al., in order to assess the risk of COVID-19 outbreak in Fars province, Iran, a machine-learning algorithm (MLA) based on a geographic information system (GIS) and a support vector machine (SVM) was used, all while daily observations of infected cases were analyzed with polynomial and autoregressive integrated moving average (ARIMA) models to examine the virus infestation patterns in the province and Iran [56]. The deep-learning model was used as a sustainable prognostic method of the COVID-19 outbreak in Bangladesh, by Mohammad Masum A.K. et al. [50]. Kumar S.L. et al. showed how, through statistical modeling on COVID-19 data, they performed linear regressions, random forests, ARIMA, and LSTM to estimate the empirical indication of COVID-19 infection and

intensity in four countries (the US, India, Brazil, and Russia) in order to arrive at better validation [45]. Wang Y. et al. proposed the ARIMA, SARIMA, and PROPHET models to predict daily new cases and cumulative confirmed cases in the US, Brazil, and India over the next 30 days on the basis of the data set on new confirmed cases and cumulative confirmed cases of COVID-19 published by the WHO [68]. Wang X. et al. created a method to predict the daily number of confirmed cases of infectious diseases by combining an ordinary differential equation mechanistic (ODE) model for infectious classes and a generalized machine-learning model (GBM) to predict how public health policies and mobility data influence the transmission rate in the ODE model [67]. In addition to monitoring general research and publication activities, the use of machine-learning approaches and a theoretical understanding of information-sharing behaviors is a productive approach to improve the effectiveness of infosurveillance [60], as demonstrated by the study conducted by Zhang Y. et al., whose experimental results suggest that it is feasible to use Twitter data to provide the surveillance and prediction of COVID-19 in the United States to support decision-making by health departments [72]. Shaghghi N. et al. demonstrated that the use of eVision, an epidemic prediction system that combines machine learning (ML) in the form of a recursive neural network (RNN) long short-term memory (LSTM) and search engine statistics in order to make accurate predictions of the weekly number of cases of highly communicable diseases, was able to achieve 89% accuracy in predicting the progress of the COVID-19 pandemic in the United States [59]. Among the proceedings concerning COVID-19 from an accuracy point of view, it is worth mentioning the study by Rohini et al. [89], carried out in India using K-nearest neighbors (KNN); the authors declared to have reached a predictive accuracy of 98.34%. Moreover, the studies on world data by Andreas et al. [75] and Satu et al. [90] should also be noted, which through a model created ad hoc by Andreas and the use of PROPHET by Satu, reached an  $R^2$  greater than or equal to 0.99.

#### 4.8. Dengue

Some of the included studies were conducted on dengue fever prediction models [55], as in the study conducted by Nguyen V.H. et al. It aimed to develop an accurate prediction model of dengue fever in Vietnam by using a wide range of weather factors as inputs to inform public health responses for outbreak prevention in the context of future climate change [52], by comparing convolutional neural network (CNN), transformer, long short-term memory (LSTM), and attention-enhanced LSTM (LSTM-ATT) models with traditional machine-learning models on weather-based DF forecasting. Interesting results were found by Shy et al., who demonstrated that statistical models built with machine-learning methods such as LASSO have the potential to greatly improve forecasting techniques for recurrent outbreaks of infectious diseases such as dengue [62], and by Xu J. et al., whose results, based on the use of LSTM, provide a more accurate dengue prediction model and could be used for other dengue-like infectious diseases [69]. As for the proceedings, Kolesnikov et al. [81] in their study indicate that the most effective predictions were given by a mathematical model based on a combination of spatial analysis techniques (MGWR) and neural networks based on the LSTM architecture.

#### 4.9. Epatitis B

The quality of the study is modest because it is a conference paper. The study by Zhang P. et al. [94] implemented various models: oriented attention model (OAM), AR, LSTM, gated recurrent unit, encoder-decoder, CNN, CNN-RNN, LSTM-attn, GRU-attn, ED-attn, CNN-attn, and CNNRNN-attn. The self-attention significantly improves the predictive accuracy of all comparable methods. The MAE and RMSE values were decreased by 51.67% and 39.43% at most, respectively. The  $R^2$  increased by 52.99% at most [94].

#### 4.10. Epatitis E

Guo Y. et al. investigated what might be the most appropriate model to predict the incidence of hepatitis E. By comparing ARIMA, SVM, and LSTM, they found that nonlinear models (SVM, LSTM) outperform linear models (ARIMA). LSTM obtained the best performance according to all three metrics: RSME, MAPE, and MAE. Hence, LSTM is the most suitable for predicting the hepatitis E monthly incidence and cases numbers [38].

#### 4.11. Hand, Foot, and Mouth Disease

Studies on a variety of other infectious diseases were included in the review, such as hand-foot-mouth disease (HFMD), an increasingly prominent public health problem that has caused an epidemic in China every year since 2008. Predicting the incidence of HFMD and analyzing the key factors that may play a role are of great importance for its prevention [49,63], which has been researched in the study conducted by Meng D. et al. They proposed two machine-learning algorithms, random forest and eXtreme gradient boosting (XGBoost), for the analysis and prediction of hand, foot, and mouth disease [49]. Among the proceedings was a study by Zhang et al. [94], where the authors stated that the self-attention methodology applied to various models increased the accuracy of their predictions.

#### 4.12. Influenza/Influenza-Like Illness (ILI)

Several studies included in the review concern influenza and influenza-like illness (ILI) [44,47,54,64,70]. Influenza epidemics are a major public health challenge worldwide and annually cause thousands of deaths, posing a serious threat to worldwide health [44,47]. In assessing the utility of influenza-like illness (ILI), surveillance systems and developing approaches for predicting future trends are important for pandemic preparedness [54]. Building forecasting models and accurate systems that track influenza activity at the city level are necessary to provide usable information for clinical, hospital, and community outbreak preparedness [44,47]. Such propositions find application in studies such as the one conducted by Venkatramanan S. et al., which can contribute to the development of the timely forecasting of infectious diseases on a global scale by using human mobility data, expanding its applications in the area of infectious disease epidemiology [64]; that of Xu Q. et al., who attempted to predict influenza in Hong Kong with Google search queries and statistical model fusion [70]; and that of Lu F.S. et al., who demonstrated how information from Internet-based data sources, when combined with using an informed, robust methodology, can be effectively used as early indicators of influenza activity at fine geographic resolutions [47].

#### 4.13. Malaria

Another application of machine learning was studied by Kamana E. et al. They demonstrated how the LSTMSeq2Seq model can be effectively applied in the prediction of malaria re-emergence [40]. The LSTMSeq2Seq model achieved an average prediction accuracy of 87.3% [40]. A lower accuracy was calculated by Brock et al. [76], who, using the BRT model, estimated an accuracy of between 55% and 82%.

#### 4.14. West Nile Virus

West Nile fever is a disease caused by the West Nile virus (WNV), which was discovered in Africa in 1979 and which quickly spread to Western Asia, Europe, Australia, and the US thanks to its natural reservoirs: birds and mosquitoes [74]. In this study, the author says that random forest was able to correctly predict the probability of WNV's presence with the highest accuracy; it also is possible to know not only the possibility of WNV's occurrence but also how it could spread. This could help policymakers to implement safety measures to prevent the deadly spread of WNV [74].

#### 4.15. Zika

This study by Roster K. et al. [57] took into consideration various diseases, including COVID-19 and Zika, which we will briefly discuss. Human infection with Zika virus is a viral disease transmitted by the bite of infected mosquitoes. If the subject is stung by a carrier mosquito and stung again by an uninfected mosquito, this can trigger a chain capable of giving rise to an endemic outbreak, in which human-to-human contagion, even if possible, is modest and unlikely [57]. The study was conducted in Brazil, and the models used were RF, TrAdaBoost, and neural network (NN) [57].

#### 4.16. Strengths and Limitations

As with any study, this systematic literature review has its limitations. First, we limited our search to articles published in English, and this might have reduced the total number of potential eligible studies, although English is the most commonly used language in the scientific community. The number of papers included in the systematic review was low, despite the fact that the initial search comprised more than 700 articles. Nevertheless, we believe this did not significantly affect our results; indeed, the small number of retrieved articles could be rather due to the fact that this is a relatively new area of research, proved by the fact that articles have been published only in the past few years: the 75 papers included in the study were published between 2016 and 2022. Second, this study reviewed only published journal articles and proceedings; however, normally important studies are disseminated mainly by scientific journal articles and not by editorials or commentaries. Another important limitation of the study is due to the fact that among the articles included are proceedings and conference papers, which could have diminished the quality of this systematic review. The study also bears limitations on the applicability of all items of the PRISMA checklist; in fact, items 12, 13, and 15 have not always been developed or completed adequately. For item 12, it was not possible to specify the effect measures for each result, because in most cases, the effect measures such as odds ratios were completely absent. This systematic review explored a new and objective field of the study, which was as specified in paragraph 1.3, to show whether it is possible to predict infectious disease outbreaks early by using machine learning. For item 13, the study was not composed of clinical trials, and a meta-analysis was not performed; therefore, most of the subitems were not applicable to this systematic review. For the applicable items, 13a to 13c, the synthesis methods were applied. In the assessment of certainty (item 15) as related to item 12 (measures of effect), the authors tried to evaluate the quality of the studies through NOS even if for this item it may not have been the appropriate tool.

Lastly, our paper also has some important strengths. First, as the title suggests, this article is a systematic literature review. Second, our search strategy was developed with several keywords. Third, the review was conducted in accordance with international guidelines.

### 5. Conclusions

The use of machine-learning models for the early or real-time verification of epidemics is a new and innovative field that today presents many study methodologies, and this makes it difficult to compare studies from a methodological point of view even if almost all of them come to the same conclusion, which is that the outbreak of major infectious diseases can be monitored.

Concerning public health and preventive strategies, this study suggests that machine learning is a new tool that can be widely used for public health practice. On the basis of the evidence collected so far, we can hypothesize that the research objectives explored so far and the results generated can be considered preliminary and that new research questions and new applications of this method can be developed in the future. In light of the recent outbreaks of infectious diseases that have occurred around the world over the past few decades, machine-learning models offer an opportunity to monitor the introduction and spread. Furthermore, their use in scientific research is expected to grow and become more and more used in the daily life of the healthcare world.



The results showed that with the association of multiple machine-learning models, it is possible to spatially and temporally predict the trends or the incidence of infectious diseases, and future research efforts will allow the construction of more-precise and more-plausible models. In conclusion, the combination of different data sets improves the results. The most probable ones are obtained when the data source is not unique, but if different data are used in this way, the results are more likely to be accurate. For future research, the integration of machine-learning models is suggested to improve existing standard epidemiological models in terms of accuracy.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/make5010013/s1>, Figure S1: PRISMA 2020 flow diagram; Table S1: Full search strategy; Table S2: Detailed description of inclusion/exclusion criteria, based on PICOS (population, intervention, comparison, outcomes and study design); Table S3: Quality assessment of the included studies, using the Newcastle–Ottawa scale (NOS).

**Author Contributions:** O.E.S. conceptualized the study, O.E.S. designed the study, and O.E.S. performed the literature search. O.E.S. performed resource analysis and data extraction. O.E.S., V.G., S.P., D.G. and F.C. wrote the first draft. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. World Health Organization. *Global Report on Infection Prevention and Control*; World Health Organization: Geneva, Switzerland, 2022; Licence: CC BY-NC-SA 3.0 IGO.
2. Provenzano, S.; Santangelo, O.E.; Giordano, D.; Alagna, E.; Piazza, D.; Genovese, D.; Calamusa, G.; Firenze, A. Predicting disease outbreaks: Evaluating measles infection with Wikipedia Trends. *Recenti. Prog. Med.* **2019**, *110*, 292–296. [[CrossRef](#)] [[PubMed](#)]
3. Gianfredi, V.; Santangelo, O.E.; Provenzano, S. Correlation between flu and Wikipedia's pages visualization. *Acta Biomed.* **2021**, *92*, e2021056. [[CrossRef](#)] [[PubMed](#)]
4. Santangelo, O.; Provenzano, S.; Piazza, D.; Giordano, D.; Calamusa, G.; Firenze, A. Digital epidemiology: Assessment of measles infection through Google Trends mechanism in Italy. *Ann Ig.* **2019**, *31*, 385–391.
5. World Health Organization. *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*; World Health Organization: Geneva, Switzerland, 2021; Licence: CC BY-NC-SA 3.0 IGO.
6. Palaniappan, S.; V, R.; David, B.; S, P.N. Prediction of Epidemic Disease Dynamics on the Infection Risk Using Machine Learning Algorithms. *SN. Comput. Sci.* **2022**, *3*, 47. [[CrossRef](#)]
7. Roy, S.; Biswas, P.; Ghosh, P. Spatiotemporal tracing of pandemic spread from infection data. *Sci. Rep.* **2021**, *11*, 17689. [[CrossRef](#)]
8. Ghannam, R.B.; Techtmann, S.M. Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1092–1107. [[CrossRef](#)]
9. Atkinson, A.; Ellenberger, B.; Piezzi, V.; Kaspar, T.; Salazar-Vizcaya, L.; Endrich, O.; Leichtle, A.B.; Marschall, J. Extending outbreak investigation with machine learning and graph theory: Benefits of new tools with application to a nosocomial outbreak of a multidrug-resistant organism. *Infect. Control. Hosp. Epidemiol.* **2022**, *16*, 1–7, Advance online publication. [[CrossRef](#)]
10. Roth, J.A.; Battagay, M.; Juchler, F.; Vogt, J.E.; Widmer, A.F. Introduction to Machine Learning in Digital Healthcare Epidemiology. *Infect. Control. Hosp. Epidemiol.* **2018**, *39*, 1457–1462. [[CrossRef](#)]
11. Higgins, J.P.T.; Altman, D.G.; Gøtzsche, P.C.; Jüni, P.; Moher, D.; Oxman, A.D.; Savović, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A.C.; et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **2011**, *343*, d5928. [[CrossRef](#)]
12. Stroup, D.F.; Berlin, J.A.; Morton, S.C.; Olkin, I.; Williamson, G.D.; Rennie, D.; Moher, D.; Becker, B.J.; Sipe, T.A.; Thacker, S.B. Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* **2000**, *283*, 2008–2012. [[CrossRef](#)]
13. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [[CrossRef](#)] [[PubMed](#)]
14. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.C.; Ioannidis, J.P.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *Ann. Intern. Med.* **2009**, *151*, W65–W94. [[CrossRef](#)] [[PubMed](#)]

15. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
16. Wells, G.A.; Shea, B.; O'Connell, D.; Pereson, J.; Welch, V.; Losos, M.; Tugwell, P. The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. Available online: [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp) (accessed on 1 November 2022).
17. Herzog, R.; Álvarez-Pasquin, M.J.; Díaz, C.; Del Barrio, J.L.; Estrada, J.M.; Gil, Á. Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. *BMC Public Health* **2013**, *13*, 154. [CrossRef]
18. Nucci, D.; Santangelo, O.E.; Provenzano, S.; Fatigoni, C.; Nardi, M.; Ferrara, P.; Gianfredi, V. Dietary Fiber Intake and Risk of Pancreatic Cancer: Systematic Review and Meta-Analysis of Observational Studies. *Int. J. Environ. Res. Public Health* **2021**, *18*, 11556. [CrossRef]
19. Nucci, D.; Santangelo, O.E.; Provenzano, S.; Nardi, M.; Firenze, A.; Gianfredi, V. Altered Food Behavior and Cancer: A Systematic Review of the Literature. *Int. J. Environ. Res. Public Health* **2022**, *19*, 10299. [CrossRef]
20. Absar, N.; Uddin, N.; Khandaker, M.U.; Ullah, H. The efficacy of deep learning based LSTM model in forecasting the outbreak of contagious diseases. *Infect. Dis. Model.* **2022**, *7*, 170–183. [CrossRef]
21. Adiga, A.; Wang, L.; Hurt, B.; Peddireddy, A.; Porebski, P.; Venkatramanan, S.; Lewis, B.L.; Marathe, M. All Models Are Useful: Bayesian Ensembling for Robust High Resolution COVID-19 Forecasting. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21), Singapore, 14–18 August 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 2505–2513. [CrossRef]
22. Ahmad, H.F.; Khaloofi, H.; Azhar, Z.; Algosaibi, A.; Hussain, J. An Improved COVID-19 Forecasting by Infectious Disease Modelling Using Machine Learning. *Appl. Sci.* **2021**, *11*, 11426. [CrossRef]
23. Ak, Ç.; Ergonul, O.; Gönen, M. A prospective prediction tool for understanding Crimean–Congo haemorrhagic fever dynamics in Turkey. *Clin. Microbiol. Infect.* **2020**, *26*, 123.e1–123.e7. [CrossRef]
24. Ak, C.; Ergonul, O.; Sencan, I.; Torunoglu, M.A.; Gonen, M. Spatiotemporal prediction of infectious diseases using structured Gaussian processes with application to Crimean–Congo hemorrhagic fever. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006737. [CrossRef]
25. Alsmadi, M.K. Modified SEIRand machine learning prediction of the trend of the epidemic of COVID-19 in Jordan under lockdowns impact. *IJECE* **2022**, *12*, 5455–5466. [CrossRef]
26. Ardabili, S.; Mosavi, A.; Ghamisi, P.; Ferdinand, F.; Varkonyi-Koczy, A.; Reuter, U.; Rabczuk, T.; Atkinson, P. COVID-19 Outbreak Prediction with Machine Learning. *Algorithms* **2020**, *13*, 249. [CrossRef]
27. Asfahan, S.; Gopalakrishnan, M.; Dutt, N.; Niwas, R.; Chawla, G.; Agarwal, M.; Garg, M.K. Using a Simple Open-Source Automated Machine Learning Algorithm to Forecast COVID-19 Spread: A Modelling Study. *Adv. Respir. Med.* **2020**, *88*, 400–405. [CrossRef] [PubMed]
28. Bagheri, H.; Tapak, L.; Karami, M.; Hosseinkhani, Z.; Najari, H.; Karimi, S.; Cheraghi, Z. Forecasting the monthly incidence rate of brucellosis in west of Iran using time series and data mining from 2010 to 2019. *PLoS ONE* **2020**, *15*, e0232910. [CrossRef] [PubMed]
29. Balogh, A.; Harman, A.; Kreuter, F. Real-Time Analysis of Predictors of COVID-19 Infection Spread in Countries in the European Union Through a New Tool. *Int. J. Public Health* **2022**, *67*, 1604974. [CrossRef] [PubMed]
30. Benedum, C.M.; Shea, K.M.; Jenkins, H.E.; Kim, L.Y.; Markuzon, N. Weekly dengue forecasts in Iquitos, Peru; San Juan, Puerto Rico; and Singapore. *PLoS Negl. Trop. Dis.* **2020**, *14*, e0008710. [CrossRef]
31. Chaurasia, V.; Pal, S. Application of machine learning time series analysis for prediction COVID-19 pandemic. *Res. Biomed. Eng.* **2020**, *38*, 35–47. [CrossRef]
32. Chen, Y.; Chu, C.W.; Chen, M.; Cook, A.R. The utility of LASSO-based models for real time forecasts of endemic infectious diseases: A cross country comparison. *J. Biomed. Inform.* **2018**, *81*, 16–30. [CrossRef]
33. Chimmula, V.K.R.; Zhang, L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons Fractals* **2020**, *135*, 109864. [CrossRef]
34. Dash, S.; Chakraborty, C.; Giri, S.K.; Pani, S.K.; Frnda, J. BIFM: Big-Data Driven Intelligent Forecasting Model for COVID-19. *IEEE Access* **2021**, *9*, 97505–97517. [CrossRef]
35. Dixon, S.; Keshavamurthy, R.; Farber, D.H.; Stevens, A.; Pazdernik, K.T.; Charles, L.E. A Comparison of Infectious Disease Forecasting Methods across Locations, Diseases, and Time. *Pathogens* **2022**, *11*, 185. [CrossRef]
36. González-Bandala, D.; Cuevas-Tello, J.; Noyola, D.; Comas-García, A.; García-Sepúlveda, C. Computational Forecasting Methodology for Acute Respiratory Infectious Disease Dynamics. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4540. [CrossRef] [PubMed]
37. Goo, T.; Apio, C.; Heo, G.; Lee, D.; Lee, J.H.; Lim, J.; Han, K.; Park, T. Forecasting of the COVID-19 pandemic situation of Korea. *Genom. Informatics* **2021**, *19*, e11. [CrossRef] [PubMed]
38. Guo, Y.; Feng, Y.; Qu, F.; Zhang, L.; Yan, B.; Lv, J. Prediction of hepatitis E using machine learning models. *PLoS ONE* **2020**, *15*, e0237750. [CrossRef] [PubMed]
39. Haq, I.; Hossain, I.; Saleheen, A.A.S.; Nayan, I.H.; Mila, M.S. Prediction of COVID-19 Pandemic in Bangladesh: Dual Application of Susceptible-Infective-Recovered (SIR) and Machine Learning Approach. *Interdiscip. Perspect. Infect. Dis.* **2022**, *2022*, 8570089. [CrossRef] [PubMed]
40. Kamana, E.; Zhao, J.; Bai, D. Predicting the impact of climate change on the re-emergence of malaria cases in China using LSTMSeq2Seq deep learning model: A modelling and prediction analysis study. *BMJ Open* **2022**, *12*, e053922. [CrossRef]



41. Katragadda, S.; Bhupatiraju, R.T.; Raghavan, V.; Ashkar, Z.; Gottumukkala, R. Examining the COVID-19 case growth rate due to visitor vs. local mobility in the United States using machine learning. *Sci. Rep.* **2022**, *12*, 12337. [\[CrossRef\]](#)
42. Ketu, S.; Mishra, P.K. Enhanced Gaussian process regression-based forecasting model for COVID-19 outbreak and significance of IoT for its detection. *Appl. Intell.* **2021**, *51*, 1492–1512. [\[CrossRef\]](#)
43. Kim, J.; Ahn, I. Infectious disease outbreak prediction using media articles with machine learning models. *Sci. Rep.* **2021**, *11*, 4413. [\[CrossRef\]](#)
44. Kim, J.; Ahn, I. Weekly ILI patient ratio change prediction using news articles with support vector machine. *BMC Bioinform.* **2019**, *20*, 259. [\[CrossRef\]](#)
45. Kumar, S.L.; Sarobin, M.V.R.; Anbarasi, L.J. Predictive Analytics of COVID-19 Pandemic: Statistical Modelling Perspective. *Walailak J. Sci. Technol. WJST* **2021**, *18*, 15583. [\[CrossRef\]](#)
46. Lmater, M.A.; Eddabbah, M.; Elmoussaoui, T.; Boussaa, S. Modelization of COVID-19 pandemic spreading: A machine learning forecasting with relaxation scenarios of countermeasures. *J. Infect. Public Health* **2021**, *14*, 468–473. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Lu, F.S.; Hou, S.; Baltrusaitis, K.; Shah, M.; Leskovec, J.; Sosic, R.; Hawkins, J.; Brownstein, J.; Conidi, G.; Gunn, J.; et al. Accurate Influenza Monitoring and Forecasting Using Novel Internet Data Streams: A Case Study in the Boston Metropolis. *JMIR Public Health Surveill.* **2018**, *4*, e4. [\[CrossRef\]](#)
48. Marra, A.R.; Alzunitan, M.; Abosi, O.; Edmond, M.B.; Street, W.N.; Cromwell, J.W.; Salinas, J.L. Modest *Clostridioides difficile* infection prediction using machine learning models in a tertiary care hospital. *Diagn. Microbiol. Infect. Dis.* **2020**, *98*, 115104. [\[CrossRef\]](#)
49. Meng, D.; Xu, J.; Zhao, J. Analysis and prediction of hand, foot and mouth disease incidence in China using Random Forest and XGBoost. *PLoS ONE* **2021**, *16*, e0261629. [\[CrossRef\]](#)
50. Masum, A.K.M.; Khushbu, S.A.; Keya, M.; Abujar, S.; Hossain, S.A. COVID-19 in Bangladesh: A Deeper Outlook into The Forecast with Prediction of Upcoming Per Day Cases Using Time Series. *Procedia Comput. Sci.* **2020**, *178*, 291–300. [\[CrossRef\]](#)
51. Murphy, C.; Laurence, E.; Allard, A. Deep learning of contagion dynamics on complex networks. *Nat. Commun.* **2021**, *12*, 4720. [\[CrossRef\]](#)
52. Nguyen, V.-H.; Tuyet-Hanh, T.T.; Mulhall, J.; Van Minh, H.; Duong, T.Q.; Van Chien, N.; Nhung, N.T.T.; Lan, V.H.; Cuong, D.; Bich, N.N.; et al. Deep learning models for forecasting dengue fever based on climate data in Vietnam. *PLoS Negl. Trop. Dis.* **2022**, *16*, e0010509. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Niraula, P.; Mateu, J.; Chaudhuri, S. A Bayesian machine learning approach for spatio-temporal prediction of COVID-19 cases. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 2265–2283. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Nsoesie, E.O.; Oladeji, O.; Abah, A.S.A.; Ndeffo-Mbah, M.L. Forecasting influenza-like illness trends in Cameroon using Google Search Data. *Sci. Rep.* **2021**, *11*, 6713. [\[CrossRef\]](#)
55. Patil, S.; Pandya, S. Forecasting Dengue Hotspots Associated with Variation in Meteorological Parameters Using Regression and Time Series Models. *Front. Public Health* **2021**, *9*, 798034. [\[CrossRef\]](#)
56. Pourghasemi, H.R.; Pouyan, S.; Farajzadeh, Z.; Sadhasivam, N.; Heidari, B.; Babaei, S.; Tiefenbacher, J.P. Assessment of the outbreak risk, mapping and infection behavior of COVID-19: Application of the autoregressive integrated-moving average (ARIMA) and polynomial models. *PLoS ONE* **2020**, *15*, e0236238. [\[CrossRef\]](#)
57. Roster, K.; Connaughton, C.; Rodrigues, F.A. Forecasting new diseases in low-data settings using transfer learning. *Chaos Solitons Fractals* **2022**, *161*, 112306. [\[CrossRef\]](#)
58. Saba, T.; Abunadi, I.; Shahzad, M.N.; Khan, A.R. Machine learning techniques to detect and forecast the daily total COVID-19 infected and deaths cases under different lockdown types. *Microsc. Res. Tech.* **2021**, *84*, 1462–1474. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Shaghaghi, N.; Calle, A.; Kouretas, G.; Mirchandani, J.; Castillo, M. eVision: Epidemic Forecasting on COVID-19. *Curr. Dir. Biomed. Eng.* **2021**, *7*, 839–842. [\[CrossRef\]](#)
60. Shen, C.; Chen, A.; Luo, C.; Zhang, J.; Feng, B.; Liao, W. Using Reports of Symptoms and Diagnoses on Social Media to Predict COVID-19 Case Counts in Mainland China: Observational Infection Study. *J. Med. Internet Res.* **2020**, *22*, e19421. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Shen, L.; Jiang, C.; Sun, M.; Qiu, X.; Qian, J.; Song, S.; Hu, Q.; Yelxiati, H.; Liu, K. Predicting the Spatial-Temporal Distribution of Human Brucellosis in Europe Based on Convolutional Long Short-Term Memory Network. *Can. J. Infect. Dis. Med. Microbiol.* **2022**, *2022*, 7658880. [\[CrossRef\]](#)
62. Shi, Y.; Liu, X.; Kok, S.-Y.; Rajarethinam, J.; Liang, S.; Yap, G.; Chong, C.-S.; Lee, K.-S.; Tan, S.S.; Chin, C.K.Y.; et al. Three-Month Real-Time Dengue Forecast Models: An Early Warning System for Outbreak Alerts and Policy Decision Support in Singapore. *Environ. Health Perspect.* **2016**, *124*, 1369–1375. [\[CrossRef\]](#)
63. Tiwari, D.; Bhati, B.S.; Al-Turjman, F.; Nagpal, B. Pandemic coronavirus disease (COVID-19): World effects analysis and prediction using machine-learning techniques. *Expert Syst.* **2022**, *39*, e12714. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Venkatramanan, S.; Sadilek, A.; Fadikar, A.; Barrett, C.L.; Biggerstaff, M.; Chen, J.; Dotiwalla, X.; Eastham, P.; Gipson, B.; Higdon, D.; et al. Forecasting influenza activity using machine-learned mobility map. *Nat. Commun.* **2021**, *12*, 726. [\[CrossRef\]](#)
65. Verma, H.; Mandal, S.; Gupta, A. Temporal deep learning architecture for prediction of COVID-19 cases in India. *Expert Syst. Appl.* **2022**, *195*, 116611. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Wang, H.; Qiu, J.; Li, C.; Wan, H.; Yang, C.; Zhang, T. Applying the Spatial Transmission Network to the Forecast of Infectious Diseases Across Multiple Regions. *Front. Public Health* **2022**, *10*, 774984. [\[CrossRef\]](#)
67. Wang, X.; Wang, H.; Ramazi, P.; Nah, K.; Lewis, M. From Policy to Prediction: Forecasting COVID-19 Dynamics Under Imperfect Vaccination. *Bull. Math. Biol.* **2022**, *84*, 90. [\[CrossRef\]](#)

68. Wang, Y.; Yan, Z.; Wang, D.; Yang, M.; Li, Z.; Gong, X.; Di Wu, D.; Zhai, L.; Zhang, W.; Wang, Y. Prediction and analysis of COVID-19 daily new cases and cumulative cases: Times series forecasting and machine learning models. *BMC Infect. Dis.* **2022**, *22*, 495. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Xu, J.; Xu, K.; Li, Z.; Meng, F.; Tu, T.; Xu, L.; Liu, Q. Forecast of Dengue Cases in 20 Chinese Cities Based on the Deep Learning Method. *Int. J. Environ. Res. Public Health* **2020**, *17*, 453. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Xu, Q.; Gel, Y.R.; Ramirez, L.L.R.; Nezafati, K.; Zhang, Q.; Tsui, K.-L. Forecasting influenza in Hong Kong with Google search queries and statistical model fusion. *PLoS ONE* **2017**, *12*, e0176690. [\[CrossRef\]](#)
71. Yang, Z.; Zeng, Z.; Wang, K.; Wong, S.-S.; Liang, W.; Zanin, M.; Liu, P.; Cao, X.; Gao, Z.; Mai, Z.; et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J. Thorac. Dis.* **2020**, *12*, 165–174. [\[CrossRef\]](#)
72. Zhang, Y.; Chen, K.; Weng, Y.; Chen, Z.; Zhang, J.; Hubbard, R. An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US. *Expert Syst. Appl.* **2022**, *198*, 116882. [\[CrossRef\]](#)
73. Zhong, R.; Wu, Y.; Cai, Y.; Wang, R.; Zheng, J.; Lin, D.; Wu, H.; Li, Y. Forecasting hand, foot, and mouth disease in Shenzhen based on daily level clinical data and multiple environmental factors. *Biosci. Trends* **2018**, *12*, 450–455. [\[CrossRef\]](#)
74. Ajith, A.; Manoj, K.; Kiran, H.; Pillai, P.J.; Nair, J.J. A Study on Prediction and Spreading of Epidemic Disease. In Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 28–30 July 2020; pp. 1265–1268. [\[CrossRef\]](#)
75. Andreas, A.; Mavromoustakis, C.X.; Mastorakis, G.; Mumtaz, S.; Batalla, J.M.; Pallis, E. Modified Machine Learning Technique for Curve Fitting on Regression Models for COVID-19 projections. In Proceedings of the 2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Pisa, Italy, 14–16 September 2020; pp. 1–6. [\[CrossRef\]](#)
76. Brock, P.M.; Fornace, K.M.; Grigg, M.J.; Anstey, N.M.; William, T.; Cox, J.; Drakeley, C.J.; Ferguson, H.; Kao, R.R. Predictive analysis across spatial scales links zoonotic malaria to deforestation. *Proc. Biol. Sci.* **2019**, *286*, 20182351. [\[CrossRef\]](#)
77. Chumachenko, D.; Chumachenko, T.; Meniailov, I.; Muradyan, O.; Zholtkevych, G. Forecasting of COVID-19 Epidemic Process by Lasso Regression. In Proceedings of the 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), Odesa, Ukraine, 29 November–3 December 2021; pp. 80–83. [\[CrossRef\]](#)
78. Chumachenko, D.; Meniailov, I.; Bazilevych, K.; Krivtsov, S. Forecasting of COVID-19 Epidemic Process by Random Forest Method. In Proceedings of the 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), Kharkiv, Ukraine, 5–7 October 2021; pp. 491–494. [\[CrossRef\]](#)
79. Fan, X.-R.; Zuo, J.; He, W.-T.; Liu, W. Stacking based prediction of COVID-19 Pandemic by integrating infectious disease dynamics model and traditional machine learning. In Proceedings of the 2022 5th International Conference on Big Data and Internet of Things (BDIOT '22), Chongqing, China, 12–14 August 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 20–26. [\[CrossRef\]](#)
80. Hasri, H.; Aris, S.A.M.; Ahmad, R. Linear Regression and Holt's Winter Algorithm in Forecasting Daily Coronavirus Disease 2019 Cases in Malaysia: Preliminary Study. In Proceedings of the 2021 IEEE National Biomedical Engineering Conference (NBEC), Kuala Lumpur, Malaysia, 9–10 November 2021; pp. 157–160. [\[CrossRef\]](#)
81. Kolesnikov, A.A.; Kikin, P.M.; Portnov, A.M. Diseases spread prediction in tropical areas by machine learning methods ensembling and spatial analysis techniques. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2019**, *8XLII-3/W*, 221–226. [\[CrossRef\]](#)
82. Kumari, P.; Toshniwal, D. Real-time estimation of COVID-19 cases using machine learning and mathematical models—The case of India. In Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), Rupnagar, India, 26–28 November 2020; pp. 369–374. [\[CrossRef\]](#)
83. Liu, Z.; Zuo, J.; Lv, R.; Liu, S.; Wang, W. Coronavirus Epidemic (COVID-19) Prediction and Trend Analysis Based on Time Series. In Proceedings of the 2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID), Guangzhou, China, 28–30 May 2021; pp. 35–38. [\[CrossRef\]](#)
84. Maaliw, R.R.; Ballera, M.A.; Mabunga, Z.P.; Mahusay, A.T.; Dejelo, D.A.; Seno, M.P. An Ensemble Machine Learning Approach For Time Series Forecasting of COVID-19 Cases. In Proceedings of the 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 27–30 October 2021; pp. 0633–0640. [\[CrossRef\]](#)
85. Mahima, Y.; Ginige, T. COVID-19 Spread prediction Based on Food Categories using Data Science. In Proceedings of the 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 6–8 November 2020; pp. 1–7. [\[CrossRef\]](#)
86. Mei, W.; Liu, Z.; Long, B.; Su, Y. Infectious Diseases Dynamic Transmissibility with Age Structure and Medical Resources. In Proceedings of the 2021 China Automation Congress (CAC), Beijing, China, 22–24 October 2021; pp. 1543–1548. [\[CrossRef\]](#)
87. Patayon, U.B. Time Series Analysis of Infected COVID-19 Cases in the Zamboanga Peninsula, Philippines using Long Short-Term Memory Neural Networks. In Proceedings of the 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), Depok, Indonesia, 14–15 September 2021; pp. 106–111. [\[CrossRef\]](#)
88. Pickering, L.; Viana, J.; Li, X.; Chhabra, A.; Patel, D.; Cohen, K. Identifying Factors in COVID—19 AI Case Predictions. In Proceedings of the 2020 7th International Conference on Soft Computing & Machine Intelligence (ISCMI), Stockholm, Sweden, 14–15 November 2020; pp. 192–196. [\[CrossRef\]](#)
89. Rohini, M.; Naveena, K.; Jothipriya, G.; Kameshwaran, S.; Jagadeeswari, M. A Comparative Approach to Predict Corona Virus Using Machine Learning. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 25–27 March 2021; pp. 331–337. [\[CrossRef\]](#)

90. Satu, S.; Rahman, K.; Alam Rony, M.; Shovon, A.R.; Alam Adnan, J.; Howlader, K.C.; Kaiser, M.S. COVID-19: Update, Forecast and Assistant—An Interactive Web Portal to Provide Real-Time Information and Forecast COVID-19 Cases in Bangladesh. In Proceedings of the 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), Dhaka, Bangladesh, 27–28 February 2021; pp. 456–460. [\[CrossRef\]](#)
91. Sri, S.; Nagarathinam, S.; Ishvarya, K.; Srinidhi, S. COVID-19 Prediction Using FbProphet. In Proceedings of the 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 25–26 March 2022; pp. 1–5. [\[CrossRef\]](#)
92. Wang, H.; Tao, G.; Ma, J.; Jia, S.; Chi, L.; Yang, H.; Zhao, Z.; Tao, J. Predicting the Epidemics Trend of COVID-19 Using Epidemiological-Based Generative Adversarial Networks. *IEEE J. Sel. Top. Signal Process.* **2022**, *16*, 276–288. [\[CrossRef\]](#)
93. Zhou, Q.; Tao, W.; Jiang, Y.; Cui, B. A Comparative Study on the Prediction Model of COVID-19. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020; pp. 1348–1352. [\[CrossRef\]](#)
94. Zhang, P.; Wang, Z.; Chao, G.; Huang, Y.; Yan, J. An Oriented Attention Model for Infectious Disease Cases Prediction. In *Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence. IEA/AIE 2022. Lecture Notes in Computer Science*; Fujita, H., Fournier-Viger, P., Ali, M., Wang, Y., Eds.; Springer: Cham, Switzerland, 2022; Volume 13343. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.