

13 documents have cited:

An Ensemble Machine Learning Approach for Time Series Forecasting of COVID-19 Cases

Maaliw R.R., Ballera M.A., Mabunga Z.P., Mahusay A.T., Dejelo D.A., Seno M.P.

(2021) 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2021, , pp. 633-640.

Search within results... 

 Analyze search results

Show all abstracts Sort on: Date (newest)

Refine results

Limit to Exclude

Open Access

- All Open Access (4) >
- Gold (4) >
- Green (2) >

Learn more

Year

- 2023 (2) >
- 2022 (9) >
- 2021 (2) >

Author name

- Maaliw, R.R. (7) >
- Lagman, A.C. (4) >
- Alon, A.S. (3) >
- Garcia, M.B. (3) >
- Mabunga, Z.P. (2) >

View more

Subject area

- Computer Science (12) >
- Engineering (8) >
- Decision Sciences (4) >
- Medicine (3) >
- Physics and Astronomy (3) >

View more

Document type

Publication stage

Source title

Keyword

Affiliation

Funding sponsor

Country/territory

Source type

Language

Limit to Exclude

	Document title	Authors	Year	Source	Cited by
<input type="checkbox"/> 1	Machine Learning and Prediction of Infectious Diseases: A Systematic Review <i>Open Access</i>	Santangelo, O.E., Gentile, V., Pizzo, S., Giordano, D., Cedrone, F.	2023	Machine Learning and Knowledge Extraction 5(1), pp. 175-198	1
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 2	A Survey on the role of ML and AI in fighting Covid-19	Malhotra, D., Sodhi, G.K.	2023	2023 International Conference on Advancement in Computation and Computer Technologies, InCACCT 2023 pp. 27-32	0
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 3	Long-Term Water Quality Prediction Using Integrated Water Quality Indices and Advanced Deep Learning Models: A Case Study of Chaohu Lake, China, 2019–2022 <i>Open Access</i>	Yao, S., Zhang, Y., Wang, P., (...), Wang, Y., Zhang, Y.	2022	Applied Sciences (Switzerland) 12(22),11329	1
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 4	A Novel Approach on Deep Learning—Based Decision Support System Applying Multiple Output LSTM-Autoencoder: Focusing on Identifying Variations by PHSMs' Effect over COVID-19 Pandemic <i>Open Access</i>	Jang, Y.-J., Kim, M.-S., Lee, C.-H., (...), Lee, S.-H., Sung, T.-E.	2022	International Journal of Environmental Research and Public Health 19(11),6763	1
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 5	COVID-19 Spatio-Temporal Evolution Using Deep Learning at a European Level <i>Open Access</i>	Kavouras, I., Kaselimi, M., Protopapadakis, E., (...), Doulamis, N., Doulamis, A.	2022	Sensors 22(10),3658	4
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 6	Population Estimation Using Wi-Fi's Received Signal Strength Indicator Based on Artificial Neural Network	Bermudez, B.A.R., Cruz, C.R., Ramos, J.D., (...), Maaliw, R.R., Ballado, A.H.	2022	Proceeding - ELTICOM 2022: 6th International Conference on Electrical, Telecommunication and Computer Engineering 2022 pp. 160-165	0
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 7	A Multistage Transfer Learning Approach for Acute Lymphoblastic Leukemia Classification	Maaliw, R.R., Alon, A.S., Lagman, A.C., (...), Fernando-Raguro, M.C., Hernandez, A.A.	2022	2022 IEEE 13th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2022 pp. 488-495	7
	View abstract View at Publisher Related documents				
<input type="checkbox"/> 8	Cataract Detection and Grading Using Ensemble Neural Networks and Transfer Learning	Maaliw, R.R., Alon, A.S., Lagman, A.C., (...), Tan, J.B., Maano, R.A.	2022	2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2022	8
	View abstract View at Publisher Related documents				



< Back to results | 1 of 13 Next >

 [Download](#)
 [Print](#)
 [Save to PDF](#)
 [Add to List](#)
 [Create bibliography](#)

Document type
Review • Gold Open Access

Source type
Journal

ISSN
25044990

DOI
10.3390/make5010013

[View more ▾](#)

Machine Learning and Knowledge Extraction • Open Access • Volume 5, Issue 1, Pages 175 - 198 • March 2023

Machine Learning and Prediction of Infectious Diseases: A Systematic Review

Santangelo, Omar Enzo^a ; Gentile, Vito^b; Pizzo, Stefano^c;
Giordano, Domiziana^d; Cedrone, Fabrizio^e

^a Regional Health Care and Social Agency of Lodi, ASST Lodi, Lodi, 26900, Italy

^b Synbrain Srl, Palermo, 90128, Italy

^c Hospital Management, AOUP "G. Rodolico-San Marco" of Catania, Catania, 95123, Italy

^d Hospital Hygiene Unit, AOUP "G. Rodolico-San Marco" of Catania, Catania, 95123, Italy

[View additional affiliations ▾](#)

1 85th percentile
Citation in Scopus

1.2
FWCI

[View all metrics >](#)

[Full text options ▾](#)

[Export ▾](#)

Cited by 1 document

Impact on online research on celebrities' uncommon diseases: the curious case of Justin Bieber and Ramsay Hunt syndrome

Santangelo, O.E., Gianfredi, V., Provenzano, S.

(2023) *Journal of Public Health (Germany)*

[View details of this citation](#)

Inform me when this document is cited in Scopus:

[Set citation alert >](#)

Related documents

A Review on Statistical and Machine Learning Approaches to Forecasting the Occurrence of Covid-19 Positive Cases

Jaulip, V., Alfred, R.
(2022) *Lecture Notes in Electrical Engineering*

Infectious Diseases Dynamic Transmissibility with Age Structure and Medical Resources

Mei, W., Liu, Z., Long, B.
(2021) *Proceeding - 2021 China Automation Congress, CAC 2021*

Machine learning in drug supply chain management during disease outbreaks: a systematic review

Emmanuel, G., Ramadhan, A., Zarlis, M.
(2023) *International Journal of Electrical and Computer Engineering*

[View all related documents based on references](#)

Find more related documents in Scopus based on:

[Authors >](#) [Keywords >](#)

Abstract

Author keywords

Sustainable Development Goals
2023

SciVal Topics

SciVal Topics

Metrics

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. © 2023 by the authors.

Author keywords

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

Sustainable Development Goals 2023

SciVal Topics

Metrics

References (94)

[View in search results format >](#)

All [Export](#) [Print](#) [E-mail](#) [Save to PDF](#) [Create bibliography](#)

- 1 (2022) *Global Report on Infection Prevention and Control*. Cited 18 times.
World Health Organization, Geneva, Switzerland, Licence: CC BY-NC-SA 3.0 IGO

- 2 Provenzano, S., Santangelo, O.E., Giordano, D., Alagna, E., Piazza, D., Genovese, D., Calamusso, G., (...), Firenze, A.

Predicting disease outbreaks: Evaluating measles infection with Wikipedia Trends

(2019) *Recenti Progressi in Medicina*, 110 (6), pp. 292-296. Cited 12 times.

http://www.recentiprogressi.it/r.php?v=3182&a=316108&l=337580&f=allegati/03182_2019_06/fulltext/07_Articolo%20originale%20Provenzano.pdf

Search for Articles:

Title / Keyword

Author / Affiliation / Email

Machine Learning and...

All Article Types

Search

Advanced

Journals / MAKE / Volume 5 / Issue 1 / 10.3390/make5010013

machine learning &
knowledge extraction

Submit to this Journal

Review for this Journal

Edit a Special Issue

Article Menu

Academic Editors



Saeed Mian Qaisar



Humaira Nisar



Abdulhamit Subasi

Subscribe SciFeed

Related Info Link

More by Authors Links

Article Views 4034

Citations 2

Table of Contents

- Abstract
- Introduction
- Materials and Methods
- Results
- Discussion
- Conclusions
- Supplementary Materials
- Author Contributions
- Funding
- Institutional Review Board Statement
- Informed Consent Statement
- Data Availability Statement
- Conflicts of Interest
- References

K

Order Article Reprints



Open Access Systematic Review

Machine Learning and Prediction of Infectious Diseases: A Systematic Review

by Omar Enzo Santangelo 1,* Vito Gentile 2, Stefano Pizzo 3 Domiziana Giordano 4 and Fabrizio Cedrone 5

1 Regional Health Care and Social Agency of Lodi, ASST Lodi, 26900 Lodi, Italy

2 Synbrain Srl, 90128 Palermo, Italy

3 Hospital Management, AOUP "G. Rodolico-San Marco" of Catania, 95123 Catania, Italy

4 Hospital Hygiene Unit, AOUP "G. Rodolico-San Marco" of Catania, 95123 Catania, Italy

5 Local Health Unit of Pescara, 65122 Pescara, Italy

* hom correspondence should be addressed.

Mach. Learn. Knowl. Extr. **2023**, *5*(1), 175–198; <https://doi.org/10.3390/make5010013>

Received: 7 December 2022 / Revised: 20 January 2023 / Accepted: 30 January 2023 /

Published: 1 February 2023

(This article belongs to the Special Issue Machine Learning for Biomedical Data Processing)

Download

Versions Notes

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

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



Share

Help

Cite

Discuss in
SciProfiles

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 piloted 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 ≥ 7 , 4–6, and ≤ 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 [20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94]. **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] to the proceedings/conference papers.

Author(s)	Title	Study Type	Population	Intervention	Outcomes	Design
Absar N	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Ajith A	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals	Case-control	Asymptomatic individuals	Machine learning	Prediction of COVID-19 infection	CASE-CONTROL
Al-Deek M	Machine learning for prediction of COVID-19 infection in asymptomatic individuals					

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 [20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73]. 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 ($n = 30$) [20,21,22,25,26,27,29,31,33,34,37,39,41,42,45,46,50,51,53,56,58,59,60,63,65,66,67,68,71,72] 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 [75,77,78,79,80,82,83,84,85,86,87,88,89,90,91,92,93], 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 [75,85,86,90,91,92,93] and all were published after 2019 [74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94].

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]. Gómez-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 (RMSE) 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

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. *Clostridiodes difficile*

The study by Marra et al. [48] was the one that obtained the least accuracy. It was a study on *Clostridiodes 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)*

Crimeane-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 [20,21,22,25,26,27,29,31,33,34,37,39,41,42,45,46,50,51,53,56,58,59,60,63,65,66,67,68,71,72]. 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]. Shaghaghi 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 R² 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, LSIM, 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 R² increased by 52.99% at most [94].

4.10. *Epiditis 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 Venkatraman 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

[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. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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 (CAMA), Pisa, Italy, 14–16 September 2020; pp. 1–6. [Google Scholar] [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. [Google Scholar] [CrossRef]
77. Chumachenko, D.; Chumachenko, T.; Menialov, 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. [Google Scholar] [CrossRef]
78. Chumachenko, D.; Menialov, 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. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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 (AID), Guangzhou, China, 28–30 May 2021; pp. 35–38. [Google Scholar] [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. [Google Scholar] [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), Bangalore, India, 6–8 November 2020; pp. 1–7. [Google Scholar] [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. [Google Scholar] [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. [Google Scholar] [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 (ISCM), Stockholm, Sweden, 14–15 November 2020; pp. 192–196. [Google Scholar] [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. [Google Scholar] [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 (IICICT4SD), Dhaka, Bangladesh, 27–28 February 2021; pp. 456–460. [Google Scholar] [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. [Google Scholar] [CrossRef]
92. Wang, H.; Tan, G.; Ma, L.; Liu, S.; Chi, L.; Yang, H.; Zhao, Z.; Tan, L. Predicting the Epidemic Trend of