231654

OFFICE OF THE PRESIDENT
RECEIVED
Date: JUL 18 2023
Time: 6:20 pm
By: [illegible]

SLSU OFFICE OF RESEARCH SERVICES

Office of
Research Services

18 July 2023

Dr. DORACIE B. ZOLETA-NANTES
University President

Thru: **Dr. MARISSA C. ESPERAL**
Vice President for Research, Extension, Production, Development and Innovation

Dear Mesdames,

The CHED-DBM Joint Circular 3, s. 2023 (commonly known as the new instrument for Faculty Re-Classification) requires that Peer Reviewer engagement of faculty members in academic journals receive proper authorization from the President or the concerned Vice President. However, these guidelines were issued towards the end of the coverage period of the 1st Cycle of the Joint Circular (July 1, 2019–July 31, 2023).

As you may be aware, peer review requests from academic journals are normally directly communicated by editors to the peer reviewer and not through the institution where s/he may be affiliated. In consultation with the Institutional Evaluation Committee, I was informed that the CHED provides leeway for additional evidence for Peer Reviewer engagement – that a list of institutionally-recognized peer reviewer engagement would be enough as additional evidence for this cycle.

In this regard, I wish to respectfully seek your **approval in principle** of the participation of faculty members listed in the attached file. Rest assured that the ORS thoroughly screened these reported Peer Reviewer engagement of our faculty members to include only those done with reputable journal publications and book publishers.

We look forward to your usual support on this matter as this will contribute greatly to the career development of our dedicated faculty researchers.

Thank you very much!

Very truly yours,

NICANOR L. GUINTO, PhD
Director, Office of Research Services

Recommending Approval:

MARISSA C. ESPERAL, PhD
Vice President for Research, Extension, Production, Development and Innovation

APPROVED/ ~~DISAPPROVED~~

Doracie B. Zoleta-Nantes, PhD
University President
JUL 19 2023

SOUTHERN LUZON STATE UNIVERSITY
Office of Research Services

C E R T I F I C A T I O N

This is to certify that the **peer reviewer engagement** of the personnel named below are approved in principle as they have been invited to review journal articles and/or book proposals while being affiliated with the University. For having been directly contacted by Editors of reputable journals and book publishers, their recognized expertise and leadership in their respective areas of research specialization contributed significantly to building the good name of Southern Luzon State University in local and international academic circles.

Name	Academic Rank	College/ Campus	Area of Research Specialization	Journal Name/ Book Publisher that made the request	Coverage/ Readership	Indexed/ Published by:	Tentative Title of the Article/ Book Proposal reviewed	Date when the invitation is received:	Date when the review was sent back to the editor:
AGUDILLA, MARY ANN R.	ASSOCIATE PROFESSOR 4	CAG	BIODIVERSITY, INSECTS, ECOSYSTEM VALUATION	PHILIPPINE JOURNAL OF SCIENCE	International	Scopus	SETTING THE INITIAL CARBON TAX RATE FOR THE CARBON TAX POLICY IN THE PHILIPPINES THROUGH THE SOCIAL COSTS OF CARBON AND WILLINGNESS TO PAY METHODS, AND THE CORRESPONDING BENEFIT-COST ANALYSIS	12/11/2022	1/2/2023
AGUDILLA, MARY ANN R.	ASSOCIATE PROFESSOR 4	CAG	BIODIVERSITY, INSECTS, ECOSYSTEM VALUATION	ACADEMIA-BIOLOGY	International	Academia Publishing	TREE HEIGHT, CANOPY COVER AND LEAF LITTER PRODUCTION OF RHIZOPHORA APICULATA IN BAGANGA, DAVAO, ORIENTAL, PHILIPPINES	1/11/2023	1/27/2023
Alinea, Jess Mark L.	Assistant Professor I	Lucena Campus	TVET, Technical Teacher Education, Curriculum and Instruction	Journal of Technical Education and Training	International	Scopus	The Role of Al-Balqa Applied University in Developing Vocational Education in Jordan	10/26/2021	11/2/2021
Alinea, Jess Mark	Assistant Professor I	Lucena Campus	TVET, Technical Teacher Education,	Journal of Technical	International	Scopus	Training-based Assessment of Employees Performance: A Case Study of Bahir Dar	12/27/2021	1/5/2022

			Communication	Applied Linguistics					
Guinto, Nicanor L.	Associate Professor III	College of Arts and Sciences	Sociolinguistics, Discourse Analysis, Communication	rEFLections	International	Scopus/ King Mongkut's University, Thailand	Filipino Non-Native English-Speaking Teachers and the Bias in Their Own Backyard	07/10/2023	07/19/2023
Maaliw, Renato III R.	Associate Professor II	CEN	Computer Vision, Machine Learning, Data Analytics	Cogent Engineering	International	Scopus, Web of Science, ASEAN Citation Index	Integrating Video Feedback Into Architectural Design Education to Engage Diverse Learning Styles	3/27/2023	4/20/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Machine Learning, Computer Vision, Data Analytics	Healthcare Analytics (Elsevier)	International	Scopus, Web of Science, ASEAN Citation Index	Prediction of Systolic and Diastolic Blood Pressures Using Machine Learning	5/4/2023	5/16/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics	Engineering (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Using ARIMA to Predict the Growth in the Subscriber Data Usage	11/4/2022	11/14/2022
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Analytics	Sensors (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Missing Traffic Data Imputation with a Linear Model Based on Probabilistic Principal Component Analysis	12/2/2022	12/10/2022
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Sensors (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Using Machine Learning on V2X Communications Data for VRU's Collisions Predictions	12/23/2022	12/26/2022
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics	Applied Science (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Performance Predictions of Sci-Fi Films via Machine Learning	1/31/2023	2/5/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Sustainability (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Thermal Images Classifications of Solid Wastes with Deep Convolutional Neural Networks	2/15/2023	2/25/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Sustainability (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Static Evaluation of a Midimew Connected Torus Network for Next Generation Supercomputers	3/2/2023	3/13/2023
Maaliw,	Associate	College of	Computer Vision,	Journal of	International	Scopus, Web of	Machine-Learning-Based Composition	3/23/2023	4/1/2023

Renato III R.	Professor II	Engineering	Machine Learning, Data Analytics, Computer Engineering	Nuclear Engineering (MDPI)		Science, ASEAN Citation Index	Analysis of the Stability of V–Cr–Ti Alloys		
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Mathematics (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	A Federated Personal Mobility Service in Autonomous Transportation	5/19/2023	5/29/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	IJERPH (MDPI)	International	Scopus, Web of Science,	Machine Learning in Predicting Severe Acute Respiratory Infection	6/6/2023	6/11/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Journal of Theoretical and Applied Electronic Commerce Research	International	Scopus, Web of Science, ASEAN Citation Index	Unveiling the Power of ARIMA, Support Vector Machine and Random Forest Regressors for the Future of Dutch Employment Market	6/14/2023	6/23/2023
Mabunga, Zoren P.	Instructor 1	College of Engineering	Artificial Intelligence, Electronics and Communication Engineering, Internet of Things	2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA 2022)	International	Scopus	Semi Autonomous Detection of Bite Points for a Surgical Needle	2/24/2022	3/7/2022
Mabunga, Zoren P.	Instructor 1	Engineering	Artificial Intelligence, Electronics and Communication Engineering, Internet of Things	IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC-2021)	International	Scopus	1. A Survey of Vulnerability Management Using Machine Learning Techniques, 2. An Adaptive Algorithm based on Interference Aware Cooperative Energy Efficiency Maximization for 5G UltraDense Networks, 3. GRAMIN GENIE-A SMART KIOSK, 4. An Automated Deep Learning Model for Detecting Sarcastic Comments,	7/2/2021	8/12/2021

YAO, CLAIRE ANN M.	ASSISTANT PROFESSOR IV	CABHA MAIN	BUSINESS ENTREPRENEURSHIP, PRODUCT DEVELOPMENT, TOURISM, LEISURE, AND HOSPITALITY	PATHWAY TO REFEREED JOURNAL PUBLICATION IN THE FIELD OF BUSINESS	Local	INSTITUTIONAL	PROBLEMS ENCOUNTERED BY MSME'S IN TAGUIG CITY AND THE ACTION TO COUNTER THE POSSIBLE EFFECTS OF ASEAN INTEGRATION: A SITUATION ANALYSIS	3/24/2020	4/4/2020
YAO, CLAIRE ANN M.	ASSISTANT PROFESSOR IV	CABHA MAIN	BUSINESS ENTREPRENEURSHIP, PRODUCT DEVELOPMENT, TOURISM, LEISURE, AND HOSPITALITY	PATHWAYS TO REFEREED JOURNAL IN THE FIELD OF BUSINESS	Local	INSTITUTIONAL	MANYAMAN MANGAN QUENI (DELICIOUS TO EAT HERE):SUCCESS FACTORS OF SELECTED RESTAURANT ENTREPRENEURS IN PAMPANGA	4/16/2020	4/21/2020

Issued this 19th day of July 2023 at Southern Luzon State University, Lucban, Quezon.

NICANOR L. GUINTO, Ph.D.
Director, Office of Research Services

MARISSA C. ESPERAL, Ph.D.
Vice President, REPDI

DORACIE B. ZOLETA-NANTES, Ph.D.
University President

Renato Racelis Maaliw

(Maaliw III, Renato R. R.)
Southern Luzon State University
Web of Science ResearcherID: EXW-3524-2022

Published names	Maaliw, Renato R., III Maaliw, Renato R. Maaliw, Renato Racelis, III Maaliw Iii, R. R. Maaliw Iii, Renato R. R.
Published Organizations	Southern Luzon State University, Southern Luzon Univ
Subject Categories	Computer Science; Telecommunications; Engineering; Environmental Sciences & Ecology; Materials Science

Documents Peer Review

Verified peer reviews

4	Sensors
3	Sustainability
2	Applied Sciences
2	Eng
2	Journal of Theoretical and Applied Electronic Commerce Research
1	Cogent Engineering
1	International Journal of Environmental Research and Public Health
1	International Journal of Retail and Distribution Management
1	Journal of Nuclear Engineering
1	Mathematics

Metrics

Open dashboard

Profile summary

20	Total documents
15	Web of Science Core Collection publications
0	Preprints
18	Verified peer reviews
0	Verified editor records

Web of Science Core Collection metrics

3 H-Index	15 Publications in Web of Science
17 Sum of Times Cited	10 Citing Articles
0 Sum of Times Cited by Patents	0 Citing Patents

View citation report

Author Impact Beamplot Summary

This is a premium feature. Learn more about how to access all of Web of Science.

Author Position

This is a premium feature. Learn more about how to access all of Web of Science.

You might be interested in...

This is a premium feature. Learn more about how to access all of Web of Science.

Co-authors

[IJERPH] Manuscript ID: ijerph-2447934 - Review Request

External   Research Reviews (Journals) ×

IJERPH Editorial Office <ijerph@mdpi.com>   Tue, Jun 6, 10:28 AM
to me, IJERPH, Pepper

Dear Dr. Maaliw,

We have received the following paper, submitted to International Journal of Environmental Research and Public Health (https://www.mdpi.com/journal/ijerph/).

Type of manuscript: Article
Title: Machine Learning in Predicting Severe Acute Respiratory Infection Outbreaks

We kindly invite you to review this paper and evaluate its suitability for publication in IJERPH. The article abstract is available at the end of this message.

If you choose to accept this invitation, we would appreciate receiving your comments within 7 days. Please let us know if you are likely to need more time to complete your review.

Please click on the link below to let us know if you will be able to provide a review and access the full manuscript and review report form.

https://susy.mdpi.com/user/review/review/39736085/h2S7Vy1v

In recognition of the contribution of reviewers, for thorough and timely review reports we provide discount vouchers for Article Processing Charges (APCs) applicable for manuscripts accepted for publication after peer review in any MDPI journal. Advice for completing your review can be found at: https://www.mdpi.com/reviewers.

Please disclose any potential conflicts of interest you might have concerning the manuscript's contents or the authors.

If you are not able to review this manuscript, we kindly ask you to decline by clicking on the above link such that we can continue processing this submission. We would also appreciate any feedback you can provide at this time (i.e., your general impression regarding the quality of this manuscript) and any suggestions for alternative expert reviewers.

IJERPH is an open access journal of MDPI. Thank you very much for your consideration and we look forward to hearing from you.

Kind regards,
Ms. Pepper Cao
Section Managing Editor
E-Mail: pepper.cao@mdpi.com

MDPI Tianjin Office
Room 1804, Block A, Lujiazui Financial Plaza, Hongqiao District, 300131
Tianjin, China
--
MDPI IJERPH Editorial Office
St. Alban-Anlage 66, 4052 Basel, Switzerland
E-Mail: ijerph@mdpi.com
http://www.mdpi.com/journal/ijerph

Gmail

label:research-reviews--journals-

Until December 31

Compose

More

Labels

- Acceptance Notifications
- Certificates
- Citations
- Huawei
- ISA
- Journal Publications
- Licenses (Do not Delete)
- Research
- Research Reviews (Conf...
- **Research Reviews (Jo...**
- SLSU
- Support Requests

Mail
Chat
Spaces
Meet

[IJERPH] Manuscript ID: ijerph-2447934 - Review Request Accepted

External   Research Reviews (Journals) ×

ijerph@mdpi.com
to me, Pepper

Tue, Jun 6, 4:08 PM

Dear Dr. Maaliw,

Thank you very much for agreeing to review this manuscript:

Manuscript ID: ijerph-2447934
Type of manuscript: Article
Title: Machine Learning in Predicting Severe Acute Respiratory Infection Outbreaks
Authors: Amauri Duarte da Silva, Marcelo Ferreira da Costa Gomes, Tatiana Schäffer Gregianini, Leticia Garay Martins, Ana Beatriz Gortini da Veiga *
Submitted to section: Infectious Disease Epidemiology,
https://www.mdpi.com/journal/ijerph/sections/Infectious_Disease_Epidemiology

The review report form can be found here:
https://susy.mdpi.com/user/review/review/39736085/h2S7Vy1v

The review report due date is: 16 June 2023

To ensure your anonymity throughout the peer review process, please do not include any identifying information in your review report either in the comments or in the metadata of any files that you upload. Please check the Guidelines for Reviewers: https://www.mdpi.com/reviewers

We look forward to receiving your valuable comments.

Kind regards,
Ms. Pepper Cao
Section Managing Editor
E-Mail: pepper.cao@mdpi.com

MDPI Tianjin Office
Room 1804, Block A, Lujiazui Financial Plaza, Hongqiao District, 300131 Tianjin, China
--
MDPI IJERPH Editorial Office
St. Alban-Anlage 66, 4052 Basel, Switzerland
E-Mail: ijerph@mdpi.com
http://www.mdpi.com/journal/ijerph

*** This is an automatically generated email ***

Reply   Reply all   Forward

Support Requests

Gmail

label:research-reviews--journals-

Until December 31

Compose

More

Labels

- Acceptance Notifications
- Certificates
- Citations
- Huawei
- ISA
- Journal Publications
- Licenses (Do not Delete)
- Research
- Research Reviews (Conf...
- Research Reviews (Jo...

Mail

Chat

Spaces

Meet

13 of 71

[IJERPH] Manuscript ID: ijerph-2447934 - Acknowledgement - Review Received

External

Research Reviews (Journals) ×

ijerph@mdpi.com

to me, IJERPH, Pepper

Sun, Jun 11, 4:16 PM (12 days ago)

Dear Dr. Maaliw,

Thank you for submitting your review of the following manuscript:

Manuscript ID: ijerph-2447934
Title: Machine Learning in Predicting Severe Acute Respiratory Infection Outbreaks
Authors: Amauri Duarte da Silva, Marcelo Ferreira da Costa Gomes, Tatiana Schäffer Gregianini, Leticia Garay Martins, Ana Beatriz Gortini da Veiga *

Our Editorial Office and Academic Editors will contact you if they have any questions about your review report. We ask that you remain available, as far as possible, during the peer-review process in case of follow-up questions.
We are continuously working to improve the services we offer and would greatly appreciate receiving feedback about your experiences through the short survey below.
Click here to start the survey:
https://www.research.net/r/Reviewer_Survey_2023

We encourage you to register an account on our submission system and bind your ORCID account (https://susy.mdpi.com/user/edit). You are able to deposit the review activity to your ORCID account manually via the below link:
https://susy.mdpi.com/user/reviewer/status/finished

We also invite you to contribute to Encyclopedia (https://encyclopedia.pub), a scholarly platform providing accurate information about the latest research results. You can adapt parts of your paper to provide valuable reference information for others in the field.

Kind regards,
Ms. Pepper Cao
Section Managing Editor
E-Mail: pepper.cao@mdpi.com

MDPI Tianjin Office
Room 1804, Block A, Lujiazui Financial Plaza, Hongqiao District, 300131
Tianjin, China
--
MDPI IJERPH Editorial Office
St. Alban-Anlage 66, 4052 Basel, Switzerland
E-Mail: ijerph@mdpi.com
http://www.mdpi.com/journal/ijerph

Reply | Reply all | Forward

Review Report Form

Journal IJERPH (ISSN 1660-4601)
Manuscript ID ijerph-2447934
Type Article

Title Machine Learning in Predicting Severe Acute Respiratory Infection Outbreaks

Authors Amauri Duarte da Silva , Marcelo Ferreira da Costa Gomes , Tatiana Schäffer Gregianini , Leticia Garay Martins , Ana Beatriz Gortini da Veiga *

Section Infectious Disease Epidemiology

Abstract Severe Acute Respiratory Infection (SARI) outbreaks occur annually, with seasonal peaks varying between geographic regions. Case notification is important to prepare healthcare networks for patient attendance and hospitalization. Therefore, health managers need adequate resource planning tools for SARI seasons. This study aims to predict SARI outbreaks based on models generated with machine learning using SARI hospitalization notification data. In this study, data from the reporting of SARI hospitalization cases in Brazil from 2013 to 2020 were used, excluding SARI cases caused by COVID-19. These data were prepared to feed a neural network configured to generate predictive models for time series. The neural network was implemented with a pipeline tool. Models were generated for the five Brazilian regions and validated for different years of SARI outbreaks. With the use of neural networks, it was possible to generate predictive models for SARI peaks, volume of cases per season, and for the beginning of the pre-epidemic period, with good weekly incidence correlation (R^2 = 0.97 [0.95–0.98] for the season of 2019 in the Southeast Region of Brazil). The predictive models achieved a good prediction of the volume of reported cases of SARI; accordingly, 9,936 cases were observed in 2019 in southern Brazil, and the prediction made by the models showed a median of 9,405 [9,105–9,738]. The identification of the period of occurrence of a SARI outbreak is possible using predictive models generated with the use of neural networks and algorithms that apply time series.

Other reviewers' comments

Reviewer 1 Review Report (round1) (Reconsider after major revision (control missing in some experiments))

Reviewer 4 Review Report (round1) (Reconsider after major revision (control missing in some experiments))

Reviewer 3 Review Report (round1) (Accept after minor revision (corrections to minor methodological errors and text editing))

Review Report Form

Reviewer's Information (will not be revealed to authors)

Name Dr. Renato Maaliw
Email rmaaliw@slsu.edu.ph
Website https://renatomaaliw3.github.io/
Affiliation Southern Luzon State University, Lucban, Quezon, Philippines
Research Keywords data; machine learning; computer vision; deep learning

Report 1 Hide Report and Author Response [-]

	High	Average	Low	No Answer
Originality / Novelty	( )	(x)	( )	( )
Significance of Content	( )	(x)	( )	( )
Quality of Presentation	( )	(x)	( )	( )
Scientific Soundness	( )	( )	(x)	( )
Interest to the readers	( )	(x)	( )	( )
Overall Merit	( )	(x)	( )	( )

Overall Recommendation

( ) Accept in present form
( ) Accept after minor revision (corrections to minor methodological errors and text editing)
(x) Reconsider after major revision (control missing in some experiments)
( ) Reject (article has serious flaws, additional experiments needed, research not conducted correctly)

Quality of English Language

( ) I am not qualified to assess the quality of English in this paper
( ) English very difficult to understand/incomprehensible
( ) Extensive editing of English language required
(x) Moderate editing of English language required
( ) Minor editing of English language required
( ) English language fine. No issues detected

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	( )	(x)	( )	( )
Are all the cited references relevant to the research?	( )	(x)	( )	( )
Is the research design appropriate?	( )	( )	(x)	( )
Are the methods adequately described?	( )	( )	(x)	( )
Are the results clearly presented?	( )	( )	(x)	( )
Are the conclusions supported by the results?	( )	(x)	( )	( )

Comments and Suggestions for Authors

1. The methods section is missing important details, what specific kind of LSTM is used? What is the configuration?

2. How do you optimized the hyperparameters for both mathemarical model and RNN models?

3. Explain in detail why each selected variables contribute to the accuracy of predictions. This should be discussed further.

4. Please provide training/validation loss plots to identify the optimized number of epochs of the LSTM model/s. 5. Also include training and validation accuracy graphs to quantitatively measure convergence in order to identify if the LSTM is not over or under fitting. 6. If possible also, decompose the time series into Error-Trend-Seasonality before modeling to identify specific trends or seasonality for the ARIMA model. 7. Experiment on different configurations of LSTMs (Stacked LSTMs, Bidirectional LSTMs or combination of both and see if it will improve the accuracy more) Less...

Comments on the Quality of English Language

Moderate changes to grammar and sentence structure. Optimize sentences so that thoughts are not repeated. Proofread also.

REVIEW CONFIRMATION CERTIFICATE

We are pleased to confirm that

Renato Maaliw

has reviewed 8 papers for the following MDPI journals in 2023:

Journal of Theoretical and Applied Electronic Commerce Research, International Journal of Environmental Research and Public Health, Mathematics, Journal of Nuclear Engineering, Sustainability, Applied Sciences, Sensors

---

Shu-Kun Lin

Dr. Shu-Kun Lin, Publisher and President
Basel, 15 June 2023

Article

Machine Learning in Predicting Severe Acute Respiratory Infection Outbreaks

Amauri Duarte da Silva [1,†], Marcelo Ferreira da Costa Gomes [2,†], Tatiana Schäffer Gregianini [3], Leticia Garay Martins [4], Ana Beatriz Gorini da Veiga [1,†,‡,*]

[1] Programa de Pós-Graduação em Tecnologias da Informação e Gestão em Saúde, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA). Rua Sarmento Leite, 245 – Porto Alegre, RS – 90050-170, Brazil
[2] Fiocruz, Programa de Computação Científica, Grupo de Métodos Analíticos em Vigilância Epidemiológica (MAVE). Av Brasil, 4365. Rio de Janeiro, RJ – 21040-900, Brazil
[3] Laboratório Central de Saúde Pública, Centro Estadual de Vigilância em Saúde da Secretaria de Saúde do Estado do Rio Grande do Sul – LACEN/CEVS/SES-RS. Av. Ipiranga, 5400. Porto Alegre, RS, CEP 90450-190, Brazil
[4] Centro Estadual de Vigilância em Saúde da Secretaria de Saúde do Estado do Rio Grande do Sul – CEVS/SES-RS. Av. Ipiranga, 5400. Porto Alegre, RS, 90450-190, Brazil
* Correspondence: anabgv76@gmail.com; anabgv@ufcspa.edu.br
† These authors contributed equally to this work.
‡ Current address: Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA). Rua Sarmento Leite, 245. Porto Alegre, RS – 90050-170, Brazil.

Abstract: Severe Acute Respiratory Infection (SARI) outbreaks occur annually, with seasonal peaks varying between geographic regions. Case notification is important to prepare healthcare networks for patient attendance and hospitalization. Therefore, health managers need adequate resource planning tools for SARI seasons. This study aims to predict SARI outbreaks based on models generated with machine learning using SARI hospitalization notification data. In this study, data from the reporting of SARI hospitalization cases in Brazil from 2013 to 2020 were used, excluding SARI cases caused by COVID-19. These data were prepared to feed a neural network configured to generate predictive models for time series. The neural network was implemented with a pipeline tool. Models were generated for the five Brazilian regions and validated for different years of SARI outbreaks. With the use of neural networks, it was possible to generate predictive models for SARI peaks, volume of cases per season, and for the beginning of the pre-epidemic period, with good weekly incidence correlation (R^2 = 0.97 [0.95–0.98] for the season of 2019 in the Southeast Region of Brazil). The predictive models achieved a good prediction of the volume of reported cases of SARI; accordingly, 9,936 cases were observed in 2019 in southern Brazil, and the prediction made by the models showed a median of 9,405 [9,105–9,738]. The identification of the period of occurrence of a SARI outbreak is possible using predictive models generated with the use of neural networks and algorithms that apply time series.

Keywords: Severe Acute Respiratory Infection; Machine Learning; Predictive Models; Epidemiological Surveillance; Neural Networks.

Citation: Silva, A.; Gomes, M.; Gregianini, T.; Martins, L.; Veiga, A. Prediction of SARI outbreaks. *Int. J. Environ. Res. Public Health* **2023**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

1. Introduction

Viral respiratory infections are easily spread in the community, affecting millions of individuals annually worldwide, representing a public health problem with high morbidity and mortality, especially in children, the elderly, and in immunocompromised patients [1]. Most acute respiratory infections are caused by viruses, and symptoms, when present, range from mild (runny nose and cough), influenza-like illnesses (ILI), to Severe Acute Respiratory Infection (SARI).

Seasonality of SARI outbreaks varies between different geographic regions [2]. In Brazil, a continental-size country, a seasonal southward wave is seen, starting in late January in the north, and reaching the south in the middle of the year [3,4]. Notably, intensity and duration of outbreaks also vary by region [5].

Influenza A (IAV) and B (IBV) viruses cause human influenza, considered one of the most important infectious diseases of humanity. Seasonal flu epidemics occur every year and, eventually, new viral subtypes with pandemic potential emerge. In 1918–1919 the humanity was ravaged by the Spanish Flu, caused by IAV H1N1; in 2009, a new IAV subtype (H1N1pdm09) caused the first influenza pandemic of the 21st century [1]. In Brazil, of the 88,464 cases of SARI hospitalization reported in 2009, 50,482 were confirmed as IAV H1N1pdm09, with 2,060 deaths [6].

Besides influenza virus, other respiratory viruses are also associated with SARI epidemics. In this sense, the new coronavirus SARS-CoV-2 emerged in humans at the end of 2019, causing the COVID-19 pandemic [7]. Within 3 years, more than 700 million COVID-19 cases have been reported and approximately 7 million people have died. On April 28, 2023, the WHO declared the end of COVID-19 global health emergency; nonetheless SARS-CoV-2 is still circulating among humans, therefore constant surveillance is necessary to be prepared for outbreaks and epidemic situations [8].

Prevention and control of SARI outbreaks rely on constant epidemiological surveillance, and information of previous seasonal epidemics of respiratory viral infection, with accumulated data of cases, may be used to predict following epidemic seasons. In this regard, the development of prediction models that consider variables specific for each geographic region is important to prepare healthcare networks and to guide health authorities in decision-making for policies and planning [9,10]. In this study, we develop and analyze predictive models for SARI outbreaks using data from Brazil; results show that, using data from previous epidemics, it is possible to predict SARI outbreaks with good precision. The application of time series with the use of LSTM can be seen in a study that seeks to predict the trend and a possible stopping time for the Covid-19 pandemic in Canada [11]. The LSTM model of the study allowed a short-term accuracy of 93.4% and the long-term forecast with 92.67% accuracy.

The use of models for predicting influenza seasons has been encouraged and promoted by organizations such as the Centers for Disease Control and Prevention (CDC) through a challenge called "Predict the Influenza Season Challenge", which rewards researchers who develop the best models predictors for annual influenza epidemics, with the objective of better predicting the moment, peak and intensity of a given influenza season. Although the CDC has an influenza surveillance network in the United States, the challenge must be met with data from social media such as Twitter and internet search data. Thus, the predictions presented are based on a variety of methods and data sources [9].

2. Materials and Methods

2.1. Definition criteria

ILI and SARI definition in Brazil is in line with the World Health Organization (WHO), with minor differences. According to WHO, ILI is characterized by fever >38ºC, accompanied by cough within 10 days of infection, and SARI includes these symptoms and hospitalization of the patient [12]. For means of epidemiological surveillance in Brazil, until 2019 case definition for ILI included fever >38ºC, accompanied by cough or sore throat within 7 days of infection; and SARI included ILI symptoms accompanied by O_2 saturation <95%, dyspnea and increased respiratory rate, requiring hospitalization of the patient [13,14].

2.2. Data source

In Brazil, SARI became of universal notification in 2009 in response to the H1N1pdm09 pandemic [15]. Until 2018, notification was carried out in the SINAN system [16], and then migrated to the SIVEP-Gripe system [17], which collects data of SARI cases, including age, sex, symptoms, comorbidities, date of first symptoms, vaccination, PCR results for respiratory viruses, hospitalization date, date of discharge or death, health unit and geographic location, and other data, totaling 252 variables.

For this study, the dataset returned 689,797 records of SARI cases in Brazil between January 2009 and March 2021. For the predictive models tested, we restricted the dataset to the following variables: number of daily notifications (used as a dependent variable), date of first symptoms, sex (male, female), age group (young, adult, elderly), and positive diagnosis for IAV or IBV. Data that identify ILI and SARI (fever, sore throat, dyspnea, and cough) were not used, as this brings redundancy with the number of notifications, not giving positive effects on the generation of models.

This study excluded data of cases that did not meet all definition criteria for SARI, as well as duplicates. Due to lack of specificity regarding self-reported fever, especially for the elderly, the following criteria were adopted to filter notified cases: presence of cough or sore throat; accompanied by the presence of dyspnea or O_2 saturation <95% or difficult breathing; with hospitalization or death.

The dataset contains data from the five Brazilian regions (North, Northeast, Midwest, Southeast, South). The analyses were carried out preferably with data from the South region because data collection in this region is more consistent than in other regions and, also, the South has a better-defined period of SARI occurrence (winter months) in relation to the other regions. In terms of the volume of annual notifications of SARI hospitalization, there is a significant variation over the years. Taking, for example, the South region, the years 2014 and 2015 had low volumes of notifications (4,754 and 4,572, respectively) and, at the other extreme, the years 2016 and 2020 had large volumes of notifications (11,266 and 21,860, respectively).

In Brazil, a correlation between the occurrence of SARI and the rainy and cold periods has been observed especially in the South [4,5]. Considering that SIVEP-Gripe data does not include the variable temperature, and with the objective of evaluating the contribution of this variable in the generation of predictive models for SARI, the daily averages of minimum and maximum temperatures and thermal amplitude per region were added to the dataset, as well as the total average of temperatures per region. Temperature was calculated based on data from the INMET website [18]. The data retrieved bring values of maximum and minimum temperatures, at different times of the day in meteorological stations located in each state of the five regions.

2.3. Data processing

Data were treated considering the date of the first reported SARI symptoms. Due to the difference in protocols in the first years (2009 to 2012) and the insufficiency of data for 2021 (only three months), the predictive model was generated using data from 2013 to 2020.

After the previous step of filtering and analyzing the data, the dataset containing SARI notification data was grouped by date of first symptoms, in other words, SARI notifications of the same date are summed in a single line; the grouping result showed a further reduction in the amount of data (rows of dataset available) for the generation of models. For most variables, grouping was done by adding up the number of cases of hospitalization for SARI on each date of the first symptom that pointed to a certain category (sex, age, etc.). For sex, grouping was done in a similar way, adding up the total number of cases that pointed to the male or female categories on each date of first symptoms. Grouping by age was done for youth (0–19 years), adults (20–59 years) and elderly (>60 years).

To standardize notifications in the different periods, records referring to COVID-19 cases were removed. At the end of this first process of filtering and excluding records with problems, a total of 354,249 records were available for generating predictive models. Due to the specificities of each Brazilian region, the data were separated by region, and, during the evaluation process, different predictive models were generated for each region. At the end of data processing, temperature data was added. These data, grouped and divided by region, were used to generate predictive models. To reduce notification bias, the 7-day moving average was applied to the amount of daily notifications.

To suit the LSTM (Long Short Term Memories) algorithm, the date of first symptoms was decomposed into year, month and day. The values of all categories were normalized between 0 and 1. To simplify the generation of predictive models, the weeks of the year are shown in the analysis instead of the epidemiological weeks.

2.4. Predictive models

A range of machine learning algorithms were tested to model the observed seasons in each region to predict the next. Since we are dealing with multivariate time series [19] with seasonality, we opted to work with the Long Short Term Memory (LSTM) recurrent neural network (RNN) [20]. Data were separated into two sets: 80% for model training and 20% for model testing. The dataset used for training and test used data available from the years prior to the year in which the prediction is to be made, with 2013 as the beginning of the period. Test data (20%) corresponds to the most recent period of the data set.

2.5. Tools

The Knime [21], a tool that uses pipelines, was used for data preparation, generation of Machine Learning models and validation of these models, generating the results to be analyzed (Apendix A). In the analysis of results, the programming language R was used [22] with RStudio interface [23]. The Knime workflow, the R code of the web tool for visualization, and instructions for use can be obtained from Github (https://github.com/vigilanciaepidemiologica/sari_prediction).

3. Results

3.1. Generation of predictive models

The generation of predictive models was simulated using machine learning algorithms such as Random Forest, Naive Bayes, Tree Assemble, RPROP MLP, but the verified accuracies were very low. Given the nature of the problem to be solved, a time series approach was shown to be more appropriate. Therefore, we used LSTM to generate predictions, which can perform well with seasonal time series [20] Looking for models with better performance, combinations of different categories (independent variables) were used and, therefore, different metrics were obtained for the tests, with the categories sex (male and female), age group (young, adult, and elderly), and positive diagnosis for IAV and IBV giving the best performance in the prediction made in the test stage of the model, with R^2 = 0.99, MAE = 1.28, MSE = 2.89, RMSE = 1.70, MSD = 0.39 and MAPE = 0.06. Of note, when using temperature variables to generate the predictive models, no significant improvements in accuracy were observed, so this variable was not used in the models presented as results in this study. Metrics for other combinations of categories can be seen in Table 1.

Table 1. Metrics for the test set to predict notifications of hospitalizations due to SARI in the year 2019, considering a model generated for the South region of Brazil.

Attributes used	R^2	MAE	MSE	RMSE	MSD	MAPE
sex, age group, ICU*, death	0.99	1.57	3.93	1.98	0.95	0.09
sex, age group, ICU*, death, mean minimum temperature	0.97	2.25	8.01	2.83	2.05	0.12
sex, age group, positive diagnosis for IAV/IBV	**0.99**	**1.28**	**2.89**	**1.70**	**0.39**	**0.06**
sex, age group, ICU*, death, positive diagnosis for IAV/IBV	0.96	3.02	12.49	3.53	3.01	0.15
sex, age group, positive IAV/IBV, mean minimum temperature	0.99	1.39	3.79	1.95	0.64	0.06

Note. The set of categories used in the study that brought better performance to the model are shown in bold.
*Use of Intensive Care Unit.

3.2. Validation

For validation, models were generated to predict the reporting curves of SARI cases in years for which data are already available. The approximation of the curves was verified, considering the week of the beginning of the pre-epidemic period, volume of cases of notification in the season, and peak week. Accuracy metrics were used in the comparison (Table 2). To generate these metrics, 100 simulations were created for each scenario tested. We report metrics in terms of median and 95% CI. Due to the COVID-19 pandemic, which started in 2020 and remains active in 2022, most validations were made for the year 2019, which, in addition to not showing the pandemic bias, also has the largest amount of data for training models, as it accumulates data from previous years, starting in 2013.

Table 2. Values of prediction validation metrics, using the median and 95% CI, considering the time series of weekly cases for the South and Southeast regions for the years 2018 and 2019.

Target Year - Region	Pearson	R^2	RMSE	MAPE	Observed Cases	Predicted Cases
2019 - South	0.90 [0.90–0.91]	0.82 [0.81–0.83	52 [51–53]	11 [10–14]	9,936	9,405 [9,105–9,738]
2018 - South	0.84 [0.81–0.86]	0.71 [0.66–0.73]	87 [81–94]	19 [17–24]	8,918	6,906 [6,348–7,550]
2019 - Southeast	0.98 [0.97–0.99]	0.97 [0.95–0.98]	35 [30–46]	10 [7–13]	14,332	14,399 [13,763–15,022]
2018 - Southeast	0.80 [0.75–0.83]	0.63 [0.56–0.69]	147 [141–155]	28 [24–32]	13,390	10,581 [9,616–11,503]

Note. The last two columns show the annual (seasonal) volume of reported SARI cases, observed and predicted, respectively.
*Use of Intensive Care Unit.

3.3. Model application

The prediction models generated with neural networks, using time series algorithms, give good results in predicting SARI outbreak seasons. As shown in Table 2, the generated models had good performance in the prediction. For both the South and Southeast, the performance measures were better for the year 2019 due to the greater amount of data for training the models. This relationship between volume of data and performance of the models was observed in the simulations for previous years, as the farther away from the present time, the less data is available for training.

Figures 1 and 2 show prediction curves generated by the models. As already mentioned, curves generated for the year 2019 show greater accuracy with the observed curve. The previous year's curve (green) is shown to assess whether the simple application of data observed in the previous year would be a way of predicting the following year. The variability of the generated models, obtained from random seeds, is small, with lower and upper limits very close to the median. The metrics collected for this graph showed a good performance prediction, presenting a Pearson coefficient of 0.90 [0.90–0.91], R^2 = 0.82 [0.81–0.83], RMSE = 52 [51–53] and MAPE = 11 [10–14].

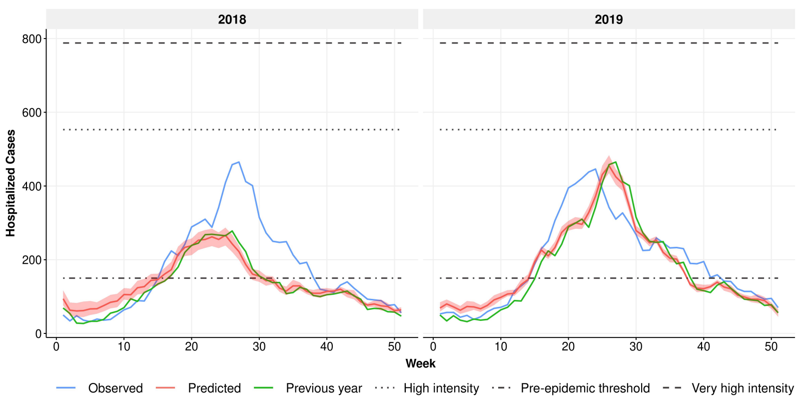

Figure 1. Prediction of the notification of hospitalization of SARI cases for the South region, in the years 2018 and 2019. The graphs show the curves of cases observed in the previous year (green), cases observed in the year (blue), and the median of the prediction of cases (red) with their variability. It also identifies the volumes of cases that characterize the pre-epidemic period, high intensity, and very high intensity..

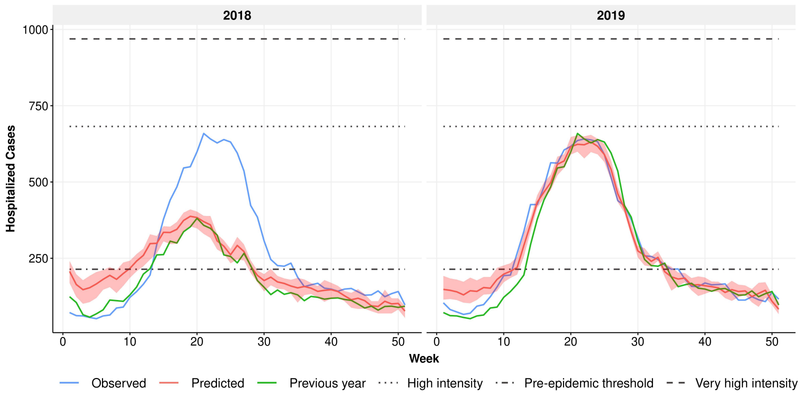

Figure 2. Prediction of the notification of hospitalization of SARI cases for the Southeast region, in the years 2018 and 2019. The graphs show the curves of cases observed in the previous year (green), cases observed in the year (blue), and the median of the prediction of cases (red) with their variability. It also identifies the volumes of cases that characterize the pre-epidemic period, high intensity and very high intensity.

There is an approximation of the predicted curve of a year (red lines in Figures 1 and 2) and the curve of cases in the previous year (green lines in Figures 1 and 2). Notwithstanding, this approximation is not always observed. In fact, as shown in Figure 3, the generated model shows a clear difference between the predicted curve for the year 2020 and the curve of the previous year (2019). Noteworthy, the volume of cases predicted for 2020 (18,078) offered a very good approximation to the volume of cases observed (21,860; 17% difference), despite the low number of cases notified in 2019 (9,901 cases, 55% difference). This is an interesting result considering that 2020 was an atypical year, during which the COVID-19 pandemic brought problems such as inconsistent notifications, especially in the initial peak. Hence, the simple use of the previous year's notification curve is not a good way to predict notification of SARI hospitalization cases for the following year.

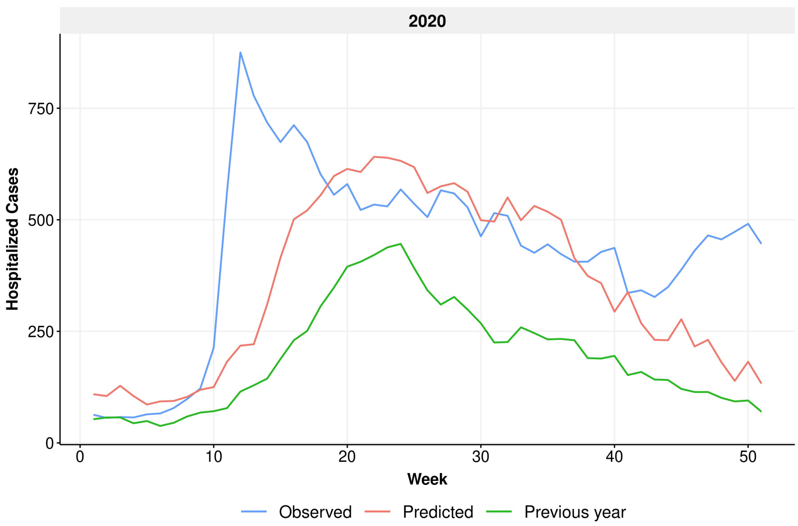

Figure 3. Prediction of the notification of hospitalization of SARI cases for the South region, in the year 2020. The graphs show the curve of SARI hospitalization cases reported in the previous year (green), curve of data observed in the year (blue), and prediction curve of number of cases (red).

Another comparative approach was to use the average notification of hospitalizations due to SARI in the years 2013 to 2018 for the South region. An average volume of 7,357 notifications was obtained against 9,936 observed and 9,327 notification cases predicted by the model (Figure 4). It is important to note that the peak of notifications of the predicted curve approximates the peak of the mean over previous seasons (week 26), and both are 2 weeks apart from the observed curve.

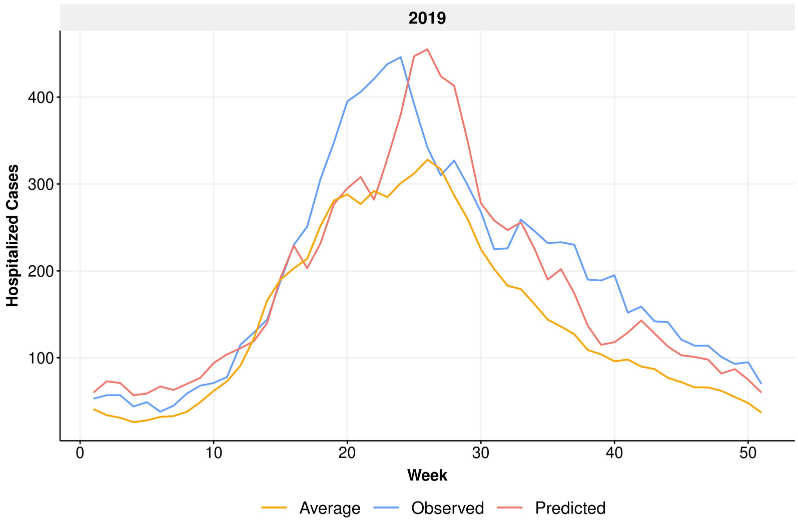

Figure 4. Prediction of the notification of hospitalization of SARI cases for the South region, in the year 2019. The graph shows the curve of the average number of hospitalized cases notified in the period from 2013 to 2018 (orange) and the curve of the data observed in the year 2019 (blue).

4. Discussion

In this work, we sought to find predictive models for SARI outbreaks using machine learning that present good performance (shown through metrics). These models can help guide managers in adopting adequate public health policies, in addition to allocating human and material resources in the necessary quantities and in a timely manner.

As results show, a prediction with good performance is viable, however exceptional events, such as emergence of new viruses and pandemics, can be unfavorable to the prediction of models. Nevertheless, these models provide important information for healthcare management.

Our study also shows that the quality and quantity of data influence the performance of predictive models. As for quality, it is noted that in the South and Southeast regions of Brazil, which have a more uniform and concise notification, the performance of predictive models is better than for other Brazilian regions. Table 3 shows the volume of SARI notifications by region. The total number of records corresponds to the total number of SARI notifications and the total number of treated records corresponds to notifications grouped by day. Aiming to emphasize the importance of the volume of data for the generation of SARI predictive models, Table 4 shows the test results for R^2. It is seen that, as the period decreases and, consequently, the amount of data for model generation also decreases, R^2 assumes smaller values, indicating that the model will not provide a good prediction when we have smaller time windows. The number of notifications is impacted by the region's population, but there is also the perception that the habit of making notifications in a more timely manner also impacts the volume of data available for the generation of predictive models. The South region is a typical case where SARI notifications are made more timely [24].

Table 3. Amount of SARI data for each Brazilian region.

Region	Total records (1)	Total records treated	Total records year 2020 (2)	Percentage of records (2)/(1)
North	22,994	2,361	13,129	57.10%
Northeast	60,908	2,837	36,335	59.66%
Midwest	29,712	2,653	12,660	42.61%
Southeast	164,354	2,897	83,763	50.96%
South	76,281	2,888	22,140	29.02%
Sum	**354,249**	**13,636**	**168,027**	**-**

Table 4. Comparison of the volume of data available in each region and in each period, with training results using the R^2 metric.

Region	Period 2013-2018		Period 2014-2018		Period 2015-2018		Period 2016-2018		Notifications 2013-2018
	Total records	R^2	Total records	R^2	Total records	R^2	Total records	R^2	
Southeast	2,173	0.96	1,811	0.73	1,453	0.44	1,094	0.05	98,166
South	2,164	0.98	1,810	0.97	1,452	0.90	1,093	0.80	48,942
Northeast	2,114	0.95	1,763	0.77	1,417	0.43	1,082	0.02	42,892
Midwest	1,929	0.67	1,618	0.45	1,329	0.41	1,046	0.10	17,604
North	71,644	0.58	1,403	0.33	1,176	0.6	1,000	0.00	16,696

Note. The evaluation metric of the predictive models is compared with the number of SARI notification records by each region of Brazil. For R^2, validation is being considered for the volume of data from 2013 to 2018.

It is important to note that, except for the South region, Brazil had a concentration of records in 2020 because of the COVID-19 pandemic. This factor influenced the performance of the models, in addition to the total number of records for each region. One explanation for this is the fact that respiratory infections are more common in regions with colder winters, such as the southern region of Brazil [3,4], but SARS-CoV-2 reached all geographic regions significantly, therefore the increase in case notification in relation to previous years was relatively greater in the other regions. In addition, it is possible that the pandemic alert situation led health institutions to adhere more strictly to SARI notification.

Obtaining a metric with great value (R^2 = 0.998) in the model trainingtest does not provide a prediction with the same performance when using real data, but it is an indicator that the prediction will be very close to the observed situation. As shown in Figure 1, for the year 2019 the model was able to predict the beginning of the pre-epidemic threshold with good accuracy. The volume of notifications in 2019 was 9,936 cases and the expected volume was 9,327 cases (-6%). If we had used the notification curve for the year 2018 to predict the year 2019, we would have 8,918 cases (-10%) and the pre-epidemic threshold would be off by one week.

In the decomposition of a historical series of SARI notifications, using moving averages, for the South region in 2019, we can observe that the seasonal factor has an impact on the notification curve and also the presence of noise shows significant variations, bringing difficulties to the approximation of a prediction curve. Even so, a predictive model based on neural networks and time series can absorb these difficulties (Figure 5).

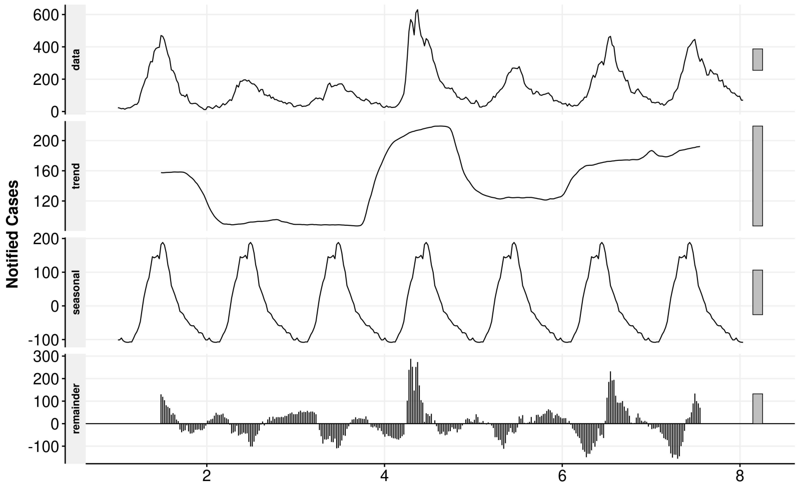

Figure 5. Decomposition of the time series for the period between 2013 and 2019 (sequence 1 to 8) for reported cases of SARI hospitalization in the South region.

The application of LSTM had better accuracy than the use of the Seasonal Autoregressive Integrated Moving Average (SARIMA) method, as can be seen in Table 5, the model generated with LSTM had a better metric in the model training step for almost all periods, except for the period with the least amount of data. The data used are from SARI notification for the southern region.

Table 5. Comparison of the SARIMA method and Neural Network models with LSTM, using the R^2 metric.

	2013-2018	2014-2018	2015-2018	2016-2018	2017-2018	2018-2018
Records	2,164	1,810	1,452	1,093	730	365
SARIMA (R^2)	0.66	0.68	0.68	0.65	0.54	0.46
Neural Network (R^2)	0.98	0.97	0.90	0.80	0.72	0.04

The main limitation of this study is related to the quality and quantity of data. With the need to standardize data collection, it was possible to use only data from years 2013 to 2020, given that previous years used other data collection and recording standards. Some regions of Brazil have fewer cases registered annually, therefore the amount of available records was not favorable to the generation of better models, especially for regions with fewer records. Another factor that caused some distortion in the models were the periods with sudden changes in the volume of notifications, such as the years 2016 – when high numbers of influenza cases were reported, due to a new H1N1pdm09 strain [25] – and 2020, with the COVID-19 pandemic. Even using the best models, it is not possible to predict rapid changes, such as the emergence of COVID-19, reliably [26]. Despite the difficulties reported here, the notification system in Brazil can be considered very good.

5. Conclusion

Identifying the period of occurrence of a SARI outbreak is possible using predictive models generated based on neural networks and algorithms that apply time series. The prediction offered by these models can help health managers define strategies to mitigate and combat outbreaks, contributing to the safety and quality of the population's health, and avoiding unnecessary expenditure on human and financial resources. For instance, estimates for the expected volume of SARI notifications, the beginning of the pre-epidemic threshold, and the peak week of notifications can be of great value in defining public health policies.

Abnormal events such as pandemics, and the reduced volume of available data are factors that make it difficult to generate predictive models. With accumulation of new data from years to come, the tendency is an increase in the quality of these models, allowing the generation of more assertive ones.

Even though it is not possible to generate prediction models that perfectly fit the observed curves, the performance metrics obtained in this study are very favorable and show the main points for decision making, such as peak week of an outbreak, case volume and pre-epidemic threshold.

Author Contributions: Conception of the study, A.D.S, M.F.C.G, T.S.G, L.G.M and A.B.G.V; design of the study, A.D.S, M.F.C.G and A.B.G.V; data collection, A.D.S, M.F.C.G, T.S.G and L.G.M; data analysis, A.D.S and M.F.C.G; data interpretation, A.D.S, M.F.C.G and A.B.G.V; drafting the article, A.D.S; critical revision of the article, M.F.C.G and A.B.G.V; final approval of the version to be submitted, A.D.S, M.F.C.G, T.S.G, L.G.M and A.B.G.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Health Department of Rio Grande do Sul, by the Ministry of Health of Brazil, and by Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS/MS/CNPq 08/2020–PPSUS, Grant process 21/2551-0000059-7). ABGV holds a Research Fellowship (PQ2) from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil – Grant process 306369/2019-2(2020-2023), 304476/2022-6(2023-2026)).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets supporting the conclusions of this article are available in https://opendatasus.saude.gov.br/dataset?tags=SRAG

Acknowledgments: Not applicable.

Conflicts of Interest: The authors have no conflicts of interest to declare.

Sample Availability: Not applicable.

Abbreviations

The following abbreviations are used in this manuscript:

IAV	Influenza A Virus
IBV	Influenza B Virus
ILI	Influenza-like illness
INMET	Instituto Nacional de Meteorologia
LSTM	Long short term memories
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MLP	Multi-layer perceptron
MSD	Mean squared deviation
MSE	Mean square error
PCR	Polymerase chain reaction
R^2	Determination coefficient
RMSE	Root mean square error
RNN	Recurrent neural network
RPROP	Resilient backpropagation
SARI	Severe acute respiratory infection
SINAN	Sistema de informação de agravos de notificação
SIVEP-Gripe	Sistema de Informação de Vigilância Epidemiológica da Gripe
WHO	World health organization

Appendix A

The Figures A1, A2 and A3 show the flow used to obtain the results of this work, going through data preparation, model training and finally the validation of the models generated with the prediction of SARI seasons. Within the dotted area are the steps performed with the Knime tool.

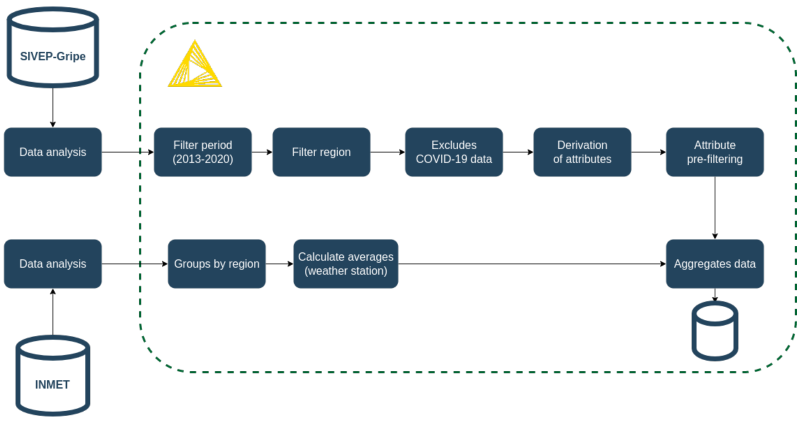

Figure A1. Steps required to prepare the data before training the models.

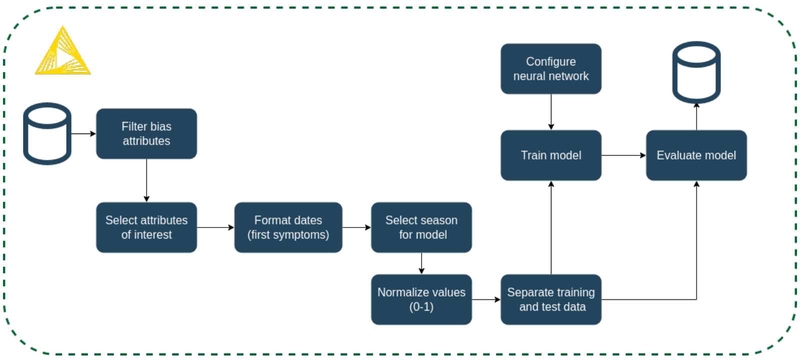

Figure A2. Steps required to train models from prepared data.

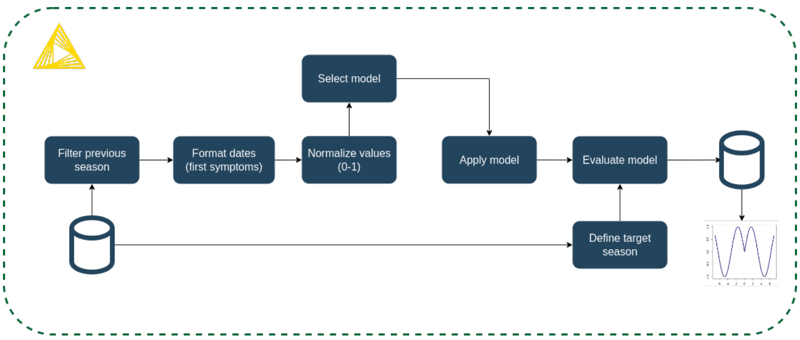

Figure A3. The necessary steps for validating the models and also for predicting a future SARI season.

References

1. World Health Organization (WHO). Global Influenza Strategy 2019–2030. World Health Organization. **2019**. Available online: https://apps.who.int/iris/handle/10665/311184 (accessed on 21 Apr 2023).
2. Li Y, Reeves RM, Wang X, Bassat Q, Brooks WA, Cohen C, et al. Global patterns in monthly activity of influenza virus, respiratory syncytial virus, parainfluenza virus, and metapneumovirus: a systematic analysis. *Lancet Glob Health* **2019**, *7*, e1031–45.
3. Alonso WJ, Viboud C, Simonsen L, Hirano EW, Daufenbach LZ, Miller MA. Seasonality of influenza in Brazil: a traveling wave from the Amazon to the subtropics. *Am J Epidemiol* **2007**, *165*, 1434–42.

4. Pscheidt VM, Gregianini TS, Martins LG, Veiga ABG. Epidemiology of human adenovirus associated with respiratory infection in southern Brazil. *Rev Med Virol*. **2021**, *31*, e2189.
5. Freitas FTM. Sentinel surveillance of influenza and other respiratory viruses, Brazil, 2000–2010. *Braz J Infect Dis* **2013**, *17*,62–8.
6. Felinto GM, Escosteguy CC, Medronho RA. Factors associated with death of severe cases of influenza A(H1N1)pdm09. *Cad Saúde Colet* **2019**,*27*, 11–9.
7. Wu D, Wu T, Liu Q, Yang Z. The SARS-CoV-2 outbreak: what we know. *Int J Infect Dis* **2020**, *94*,44–8.
8. Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. Available online: https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic (accessed on 13 Mai 2023)
9. Centers for Disease Control and Prevention (CDC). CDC Competition Encourages Use of Social Media to Predict Flu. Available online: https://www.cdc.gov/flu/news/predict-flu-challenge.htm (accessed on 09 Apr 2023)
10. Centers for Disease Control and Prevention (CDC). Flu Forecasting. Available online: https://www.cdc.gov/flu/weekly/flusight/how-flu-forecasting.htm (accessed on 09 Apr 2023).
11. Chimmula VKR, Zhang L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals* **2020**, *135*, 109864.
12. WHO surveillance case definitions for ILI and SARI. Available online: https://www.who.int/teams/global-influenza-programme/surveillance-and-monitoring/case-definitions-for-ili-and-sari (accessed on 09 Apr 2023).
13. Martins L, Menegolla I, Ranieri T, Bercini M, Ueda E, Porto M. A vigilância da influenza no Rio Grande do Sul. *Bol Epidemiológico* **2011**, *13*, 4–11.
14. Ministério da Saúde do Brasil. Vigilância sentinela de síndrome gripal (SG) no Brasil. Secretaria de *Vigilância em Saúde* **2017**.
15. Costa LMC, Merchan-Hamann E. Influenza pandemics and the structure of Brazilian health care system: brief history and characterization of the scenarios. *Rev Pan-Amazônica Saúde* **2016**, *7*, 11–25.
16. SINAN – Sistema de Informação de Agravos de Notificação. Ministério da Saúde, Brasil 2020. Available online: http://portalsinan.saude.gov.br (accessed on 09 Apr 2023).
17. SIVEP-Gripe – Sistema de Informação da Vigilância Epidemiológica da Gripe. Secretaria de Vigilância em Saúde, Brasil. Available online: https://sivepgripe.saude.gov.br/sivepgripe (accessed on 09 Apr 2023).
18. Instituto Nacional de Meteorologia: Dados Históricos Anuais. INMET. Available online: https://portal.inmet.gov.br/dadoshistoricos (accessed on 09 Apr 2023).
19. Karim F, Majumdar S, Darabi H, Harford S. Multivariate LSTM-FCNs for time series classification. *Neural Netw* **2019**, *116*, 237–45.
20. Bandara K, Bergmeir C; Hewamalage H. LSTM-MSNet: Leveraging forecasts on sets of related time series with multiple seasonal patterns. *IEEE Trans Neural Netw Learn Syst*. **2020**, *32*, 1586–99.
21. Open for Innovation Knime. Knime. Available online: https://www.knime.com/ (accessed on 21 Apr 2023).
22. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. R Core Team. Available online: https://www.R-project.org/ (accessed on 21 Apr 2023).
23. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. RStudio Team. Available online: http://www.rstudio.com/ (accessed on 21 Apr 2023).
24. Silva AD, Veiga ABG, Cruz OG, Bastos LS, Gomes MFC. Severe acute respiratory infection surveillance in Brazil: the role of public, private and philanthropic healthcare units. *Health Policy and Planning* **2022**, *37(9)*, 1075-1085.
25. InfoGripe. Fiocruz. Available online: http://info.gripe.fiocruz.br/ (accessed on 09 Apr 2023).
26. Centers for Disease Control and Prevention (CDC). Forecasts of Flu Hospitalizations. Available online: https://www.cdc.gov/flu/weekly/flusight/flu-forecasts.htm (accessed on 09 Apr 2023).