Research Article

The Iranian Journal of Health Insurance | Volume 6, Issue 2, Summer 2023 | Pages: 113-122

journal.ihio.gov.ir

Forecasting the Spread of COVID-19 Using Time Series in Mehriz City, Iran

Mohammad Hossein Karimizarchi[1] **MSc, Davood Shishebori**[1*] **PhD**

[1] Department of Industrial Engineering, Faculty of Technology and Engineering, Yazd University, Yazd

***Correspondence to:** Davood Shishebori, Email: shishebori@yazd.ac.ir

Received: April 5, 2023 **Revised:** July 20, 2023 **Accepted:** July 30, 2023 **Online Published:** August 27, 2023

Abstract

Introduction: Coronavirus disease 2019 or COVID-19, which is also called acute respiratory disease NCAV-2019 or commonly called corona, is a respiratory disease caused by acute respiratory syndrome coronavirus-2. Forecasting the number of new cases and deaths today can be a useful step in predicting the costs and facilities needed in the future. This study aims to model and predict new cases and deaths efficiently in the future.

Methods: In this article, 9 forecasting techniques were tested on the data of COVID-19 of Mehriz city, Iran as a case study from 2020/02/26 to 2021/12/19 and using the evaluation criteria of mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) of the models were compared.

Results: For cumulative cases of hospitalization, ARIMA, Exponential, Holt-Winters, and STL models performed better and autoregressive neural networks, Theta, and KNN regression showed poor performance. Also, for cumulative mortality cases, KNN regression, Exponential and Theta models have better performance in predicting cumulative mortality cases, and autoregressive neural networks, ARIMA, and cubic spline smoothing showed poor performance.

Conclusion: the best model according to the mentioned evaluation criteria for predicting cumulative cases of hospitalization of COVID-19 is STL model and for cumulative cases of death is the KNN regression model. Also, the autoregressive neural network model has the worst performance among other models, both for hospitalization and death cases. Also, the important point is that the data should be updated in real-time.

Keywords: COVID-19, SARS-CoV-2, Virus Disease, Pneumonia

Highlights

1. Due to the fact that many factors are involved in the spread of epidemic diseases such as Covid-19 (such as the emergence of new strains), the prediction of this disease has many complications and it is not possible to generalize a specific method for all regions. Because the behavior of this virus has different forms in different regions.
2. Also, by using traditional models or the so-called old models, better results can be obtained in practice, while the results obtained from newer models such as the autoregressive neural network model have worse results in predicting Covid-19 in a case study. Therefore, it can be concluded that the use of new models is not always effective in practice, and according to the type of data, old models may have better results and efficiency in forecasting.

Citation:
Karimizarchi MH, Shishebori D. Forecasting the spread of COVID-19 using time series in Mehriz city, Iran. Iran J Health Insur. 2023;6(2):113-22.

Copyright © 2023 The Author(s); Published by National Center for Health Insurance Research. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[DOI: 10.66224/ihio.6.2.113]

مقاله پژوهشی

فصلنامه علمی- پژوهشی بیمه سلامت ایران | دوره ۶، شماره ۲، تابستان ۱۴۰۲ | صفحات ۱۱۳ تا ۱۲۲

پیش‌بینی شیوع بیماری کووید-۱۹ با استفاده از سری‌های زمانی در شهرستان مهریز

محمدحسین کریمی‌زارچی[۱] MSc، داود شیشه‌بری[۲*] PhD

[۱] گروه مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه یزد، یزد

[*] **نویسنده مسئول:** داود شیشه‌بری، پست الکترونیک: shishebori@yazd.ac.ir

دریافت: ۱۴۰۲/۰۱/۱۶ تصحیح: ۱۴۰۲/۰۴/۲۹ پذیرش: ۱۴۰۲/۰۵/۰۸ انتشار آنلاین: ۱۴۰۲/۰۶/۰۵

چکیده

مقدمه: بیماری کووید-۱۹ یک بیماری تنفسی است که در اثر سندرم تنفسی حاد کرونا ویروس-۲ ایجاد می‌شود. پیش‌بینی تعداد موارد جدید و مرگ‌ومیر می‌تواند گام مفیدی در راستای پیش‌بینی هزینه‌ها و امکانات مورد نیاز در آینده باشد. هدف از این مطالعه مدل‌سازی، مقایسه عملکرد مدل‌ها و پیش‌بینی موارد جدید و مرگ‌ومیر در آینده نزدیک است.

روش‌بررسی: در این مقاله ۹ تکنیک پیش‌بینی بر روی داده‌های کووید-۱۹ شهرستان مهریز به‌عنوان یک مطالعه موردی از تاریخ ۱۳۹۸/۱۲/۰۷ الی ۱۴۰۰/۰۹/۲۸ تحت آزمایش قرار گرفت و با استفاده از معیارهای ارزیابی میانگین مربعات خطا (MSE)، جذر میانگین مربعات خطا (RMSE)، میانگین قدر مطلق خطا (MAE) و میانگین درصد قدر مطلق خطا (MAPE) در مدل‌ها با هم مقایسه شدند.

یافته‌ها: برای موارد تجمعی بستری مدل‌های ARIMA، نمایی، هولت-وینترز و STL عملکرد بهتر و شبکه‌های عصبی اتورگرسیو، تتا و رگرسیون KNN عملکرد نامناسبی را از خود نشان دادند. همچنین برای موارد تجمعی مرگ‌ومیر، مدل‌های رگرسیون KNN، نمایی و تتا دارای عملکرد بهتری در پیش‌بینی موارد تجمعی مرگ‌ومیر هستند و شبکه‌های عصبی اتورگرسیو، ARIMA و هموارسازی اسپلاین مکعبی عملکرد نامناسبی از خود نشان دادند.

نتیجه‌گیری: بهترین مدل با توجه به معیارهای ارزیابی مذکور برای پیش‌بینی موارد تجمعی بستری کووید-۱۹ مدل STL و برای موارد تجمعی فوت مدل رگرسیون KNN است. همچنین مدل شبکه‌های عصبی اتورگرسیو دارای بدترین عملکرد در میان دیگر مدل‌ها، برای موارد بستری و هم موارد فوت است. نکته حائز اهمیت این است که باید داده‌ها در زمان واقعی به‌روز شوند.

واژگان کلیدی: کووید-۱۹، کروناویروس سندرم حاد تنفسی-۲، ذات‌الریه، بیماری ویروسی

نکات ویژه

۱-به دلیل اینکه عوامل بسیاری در شیوع بیماری‌های همه‌گیر مثل کووید-۱۹ دخیل است (مانند ظهور سویه های جدید) پیش بینی این بیماری دارای پیچیدگی‌های زیادی می‌باشد و نمی‌توان روش خاصی را برای همه مناطق تعمیم داد. چرا که رفتاراین ویروس در مناطق مختلف دارای اشکال مختلفی می‌باشد.

۲- همچنین با استفاده از مدل‌های سنتی و یا به اصطلاح مدل‌های قدیمی می‌توان در عمل نتایج بهتری کسب نمود، درحالیکه نتایج به دست آمده از مدل‌های جدیدتر مثل مدل شبکه‌های عصبی اتورگرسیو نتایج بدتری در پیشبینی کووید-۱۹ در مطالعه موردی دارد. بنابراین می‌توان نتیجه گرفت که استفاده از مدل‌های جدید در عمل همیشه جوابگو نیست و با توجه به نوع داده‌ها ممکن است مدل‌های قدیمی نتیجه و کارایی بهتری در پیشبینی داشته باشند

مقدمه

بیماری کووید-۱۹یک بیماری تنفسی است که در اثر سندرم تنفسی حاد کرونا ویروس-۲ ایجاد می‌شود. این ویروس اولین بار در دسامبر ۲۰۱۹ شناسایی شد. در ۱۱ مارس سال ۲۰۲۰، سازمان بهداشت جهانی این بیماری را به‌عنوان بیماری ویروسی همه‌گیر جهانی اعلام کرد. از زمانی که کووید-۱۹ به‌عنوان بیماری همه‌گیر اعلام شد، بسیاری از کشورهای جهان به دلیل ابتلا به بیماری کووید-۱۹ به شدت تحت تأثیر قرار گرفتند و اقدامات پیشگیرانه متعددی از جمله قرنطینه، تست‌های سریع، ماسک زدن، خود قرنطینه کردن، رعایت فاصله اجتماعی برای جلوگیری از شیوع بیماری همهگیر کووید-۱۹ در حال اجراست. با وجود این اقدامات، کووید-۱۹ به دلایل مختلف از جمله تراکم جمعیت، سفرهای جهانی، ظهور سویههای جدید و... به سرعت در حال انتشار بوده و بر سلامت انسان و اقتصاد جهانی تأثیر گذاشته است. آمادهسازی و کنترل شیوع بیماری کووید-۱۹ نیاز به برنامه‌ریزی و سیاست‌های دقیق دارد. مدل‌سازی، برآورد

[DOI: 10.66224/ihio.6.2.113]

و پیش‌بینی شیوع ویروس‌ها و ویژگی‌های اپیدمیولوژیک از موارد مهم در تهیه تجهیزات مورد نیاز برای کنار آمدن با پیامدهای آن است. به‌منظور برخورداری از یک برنامه مناسب برای کووید-۱۹، پیش‌بینی موارد تأیید شده در آینده بسیار مهم و حیاتی است، بنابراین مدل‌سازی و پیش‌بینی تعداد موارد کووید-۱۹ تأیید شده و بهبودیافته و همچنین موارد مرگ‌ومیر، نقش مهمی در برنامه‌ریزی برای کنترل شیوع کووید-۱۹ در جهان دارد [۱].

مدل‌های آماری مختلفی برای پیش‌بینی تعداد موارد آینده و همچنین پیش‌بینی شیوع بیماری‌های عفونی در آینده نزدیک استفاده می‌شود. ژانگ و همکاران [۲] برای پیش‌بینی تب حصبه از مدل سری زمانی اتورگرسیو میانگین متحرک تلفیق شده فصلی (SARIMA) و سه مدل مختلف برگرفته از شبکه‌های عصبی شامل شبکه‌های عصبی پس انتشار (BPNN)، شبکه‌های عصبی تابع شعاعی (RBFNN) و شبکه‌های عصبی بازگشتی المان (ERNN) استفاده کردند. همچنین پژوهشگران دیگری به‌طور مشابه از مدل‌های اتورگرسیو میانگین متحرک تلفیق شده (ARIMA) برای پیش‌بینی بیماری‌های عفونی مانند سل [۳]، تب دانگ [۴] و تب مالت [۵] استفاده کرده‌اند.

سیلان [۶] از مدل‌های ARIMA برای پیش‌بینی روند اپیدمیولوژیک شیوع کووید-۱۹ در ایتالیا، اسپانیا و اروپا استفاده کرده است. لوتفی و بورکو [۷] برای پیش‌بینی تعداد موارد تأیید شده و مرگ‌ومیر کووید-۱۹ از مدل ARIMA در مرکز کنترل و پیشگیری اروپا (ECDC) استفاده کردند. تاندن و همکاران [۸] مدل ARIMA را برای پیش‌بینی موارد کووید-۱۹ در هند به‌کار بردند. این مطالعه نشان‌دهنده روند صعودی موارد کووید-۱۹ در روزهای آینده است. پرون [۹] در این مقاله یک مدلARIMA برای پیش‌بینی روند اپیدمی طی دوره زمانی بعد از ۴ آوریل ۲۰۲۰ (بیش از ۴۰روز)، با استفاده از داده‌های اپیدمیولوژیک ایتالیا در سطح ملی و منطقه‌ای استفاده کرده است.

کیبریا و همکاران [۱۰] از مدل ARIMA برای تخمین تعداد مورد انتظار روزانه موارد کووید-۱۹ در بنگلادش بر اساس داده‌های ۲۰ آوریل ۲۰۲۱ تا ۴ ژوئیه ۲۰۲۱ استفاده کرده‌اند. مدل ARIMA بهترین نتایج را در بین چند مدل اجرا شده نسبت به مدل اتورگرسیو (AR)، میانگین متحرک (MA)، اتورگرسیو میانگین متحرک (ARMA) نشان داد. خان و گوپتا [۱۱] از مدل سری زمانی تک‌متغیره برای پیش‌بینی تعداد موارد آلوده به کووید-۱۹ در هند، استفاده کردهاند. چایون و همکاران [۱۲]، مدل ARIMA را برای تحلیل جهانی کووید-۱۹ در بازه زمانی از ۲۲ ژانویه ۲۰۲۰ تا ۷ آوریل ۲۰۲۰ استفاده کردند. به منظور پیش‌بینی تعداد افراد مبتلا و فوت شده در آینده، مخایرز و آلاف [۱۳] از الگوریتم‌های پیش‌بینی کوتاه‌مدت برای پیش‌بینی تعداد مبتلایان و مرگ‌ومیرها در آینده کوتاه استفاده کردند. پیش‌بینی با استفاده از مدل ARIMA، ARIMA ترکیبی، هولت-وینترز و هموارسازی نمایی انجام شد. پس از ارزیابی الگوریتم‌های آزمایش‌شده با استفاده از MAPE، متوجه شدند که هموارسازی نمایی بهترین الگوریتم برای پیش‌بینی موارد بهبودیافته و تأیید شده و هولت-وینترز بهترین الگوریتم برای پیش‌بینی موارد مرگ‌ومیر است.

در تحقیقی دیگر که در سال ۲۰۲۲ انجام شد، لی و همکاران [۱۴] برای کشورهای بریتانیا، آفریقای جنوبی، برزیل و نیجریه از مدل ARIMA برای پیش‌بینی شیوع ویروس کرونا در بریتانیا، آفریقای جنوبی، برزیل و نیجریه استفاده کرده‌اند. مجموعه دادهها در این مطالعه از ۱ مارس ۲۰۲۰ تا ۳۱ ژانویه ۲۰۲۱ را شامل می‌شود و از وب سایت سازمان بهداشت جهانی به دست آمده است.

در مطالعه حاضر، قرار است موارد تجمعی کووید-۱۹ (موارد تأیید شده و مرگ‌ومیر) برای شهرستان مهریز به‌عنوان یک مطالعه موردی از تاریخ ۱۳۹۸/۱۲/۰۷ الی ۱۴۰۰/۰۹/۲۸ (دادهها از معاونت بهداشت و درمان استان یزد اخذ شده است) با استفاده از تکنیک‌های سری‌های زمانی و یادگیری ماشین ازجمله مدل‌های اتورگرسیو میانگین متحرک تلفیق شده (ARIMA)، هموارسازی نمایی (ETS)، هولت-وینترز (Holt-Winter)، رگرسیون KNN، تتا (Theta)، شبکه عصبی اتورگرسیو (ARNN)، هموارسازی اسپلاین مکعبی (CSS)، روش STL و روش BATS مدل‌سازی و براساس معیارهای میانگین خطای مطلق (MAE)، درصد میانگین خطای مطلق (MAPE)، خطای میانگین مربعات (MSE) و خطای جذر میانگین مربعات (RMSE) به‌عنوان معیار انتخاب بهترین مدل استفاده شود.

روش بررسی

با بررسی‌های انجام شده، ناشناخته بودن رفتار این ویروس

[DOI: 10.66224/ihio.6.2.113]

و پژوهش‌های صورت گرفته در پیش‌بینی کووید-۱۹ سعی بر این شد که از روش‌های مختلفی در زمینه‌های مختلف استفاده شـود تا به نوعی قابلیت اطمینان را بالا ببریم. در این مطالعه از مدل‌های آماری اتورگرسـیو میانگین متحرک تلفیق شـده (ARIMA)، هموارسـازی نمایـی (ETS)، هولت-وینترز (Holt-Winters)، رگرسیون KNN، تتا (Theta)، شبکه عصبی اتورگرسیو (ARNN)، هموارسازی اسپلاین مکعبی (CSS)، روش STL و روش BATS برای ایجاد یک پیش‌بینی ۱۴ روزه از موارد تجمعی کووید-۱۹ برای شهرسـتان مهریز استفاده کرده‌ایم. اغلب این مدل‌ها به دلیل سهولت تفسـیر نتایج، زمان محاسبه پایین و قدرت پیش‌بینی بالا مورد استفاده قرار می‌گیرند. برای داشتن یک پیش‌بینی معنادار آماری از داده‌های سری زمانی، حداقل حجم نمونه ۳۰ مشاهده مورد نیاز است [۱۵]. تعداد مشاهدات (یعنی حجم نمونه) مورد استفاده در کار حاضر بسیار بیشتر از حداقل اندازه مورد نیاز برای انجام پیش‌بینی سری‌های زمانی معنادار است، زیرا داده‌های جمع‌آوری شده برای مدت حدوداً ۲۲ مـاه اسـت (داده‌های کرونای مورد اسـتفاده در این مقاله ۶۶۲ داده از تاریـخ ۱۳۹۸/۱۲/۰۷ الی ۱۴۰۰/۰۹/۲۸ به‌صورت روزانه اسـت که از معاونت بهداشـت و درمان استان یزد اخذ شده است).

در این پژوهش ۱۴ دادهٔ مشاهده شدهٔ آخر را به‌عنوان داده‌های تست در نظر می‌گیریم و با الباقی داده‌ها و مدل‌هایی که تشریح می‌شـوند، برای ۱۴ روز آینده موارد تجمعی کووید-۱۹ (بستری و فوت) برای شهرستان مهریز پیش‌بینی می‌کنیم. سپس بهترین مدل پیشنهادی، به‌وسیله معیارهای ارزیابی را انتخاب می‌کنیم. لازم به ذکر است تاریخی را که اولین مورد، اعم از بستری یا فوت، گزارش شده به‌عنوان روز شروع سری زمانی در نظر گرفته‌ایم.

داده‌های سـری زمانی، دنباله‌ای از مقادیر عددی اسـت که دارای یک برچسب زمانی مرتبط با هر مقدار است [۱۶]. بنابراین سـری زمانی مجموعه‌ای از مشـاهدات است که بر حسب زمان (یا هر کمیت دیگر) مرتب شـده باشند. و معمولاً آن را به‌صورت $X_{t1},X_{t2},\ldots,X_{tn}$ نمایش می‌دهند.

برای توسـعه مدل‌های آماری براسـاس داده‌های سری‌های زمانی، فرضیات زیر در نظر گرفته شده‌اند:

داده‌های سری‌های زمانی حاوی ناهنجاری‌ها و داده‌های پرت نیسـتند. داده‌ها تک‌متغیری هستند. این مدل‌ها فرض می‌کنند داده‌ها ایسـتا هستند و نیاز به میانگین و واریانس در طول زمان ندارند. پارامترهای مدل‌ها و خطاها با توجه به گذشت زمان ثابت فرض می‌شوند.

بـرای محاسـبات و کدنویسـی از نرم‌افزار برنامه‌نویسـی R استفاده کرده‌ایم.

فرآیند اتورگرسیو میانگین متحرک تلفیق شده (ARIMA)

مـدل ARIMA(p,d,q) اولین بار توسـط باکس و جنکینز در سـال ۱۹۷۶ معرفی شد [۱۷] که می‌توان از آن برای پیش‌بینی داده‌های سـری زمانی ایستا غیرفصلی استفاده کرد. الگوی کلی ARIMA به‌صورت رابطه زیر خواهد بود

$$\varphi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t$$

کـه در آن عملگـر AR ایسـتا ($\varphi_p(B)=(1-\varphi_1 B-\ldots-\varphi_p B^p)$ و عملگر MA وارون‌پذیر

($\theta q(B)=(1-\theta_1 B-\ldots-\theta q Bq)$ است. پارامتر θ_0 برای d=۰ و d>۰ نقش‌های بسیار متفاوتی را بازی می‌کند.

فرآیند اتورگرسیو میانگین متحرک تلفیق شده فصلی (SARIMA)

الگـوی ARIMA فصلی جمعـی معروف باکـس و جنکینز به‌صورت روابط زیر تعریف می‌شود:

$$SARIMA: \phi_P(B^s)\varphi_p(B)(1-B)^d(1-B^s)^D Z_t$$
$$= \theta_q(B)\vartheta_Q(B^s)a_t$$
$$ARIMA: \varphi_p(B)(1-B)^d Z_t = \theta_q(B)a_t$$
$$\phi_P(B^s) = 1-\phi_1 B^s - \phi_2 B^{2s} - \ldots - \phi_P B^{Ps}$$
$$\vartheta_Q(B^s) = 1-\vartheta_1 B^s - \vartheta_2 B^{2s} - \ldots - \vartheta_Q B^{Qs}$$
$$\varphi_p(B) = 1-\varphi_1 B^1 - \varphi_2 B^2 - \ldots - \varphi_p B^p$$
$$\theta_q(B) = 1-\theta_1 B^1 - \theta_2 B^2 - \ldots - \theta_q B^q$$

at یـک فرآیند تصادفی محض با میانگین صفر اسـت. برای سهولت، $\varphi_p(B)$ و $\theta_q(B)$ را به ترتیب عوامل اتورگرسیو و میانگین متحـرک و $\vartheta_Q(B^s)$ و $\phi_p(B^s)$ را عوامـل اتورگرسـیو و میانگین متحرک فصلی می‌نامند و اغلب به‌صورت ARIMA(p,d,q)(P,D,Q) s نشان می‌دهند که در آن اندیس s دوره فصلی را نشان می‌دهد.

[DOI: 10.66224/ihio.6.2.113]

روش هموارسازی نمایی

یکی از روش‌های پیش‌بینی رایج و محبوب روش هموارسازی نمایی است. روش هموارسازی نمایی به جای اینکه یک میانگین ساده را حساب کند، میانگین وزنی مقادیر گذشته را می‌گیرد، به‌طوری که وزن‌ها به‌صورت نمایی متمایل به داده‌های گذشته کاهش می‌یابد، بدین‌معنا که به داده‌های اخیر اهمیت بیشتری نسبت به داده‌های قدیمی‌تر می‌دهد. روش هموارسازی نمایی برای پیش‌بینی سری‌هایی که فاقد روند و الگوی فصلی هستند، استفاده می‌شود. روش پیش‌بینی که روند هموارسازی در زمان t+1 را به دست می‌دهد، به‌صورت رابطه زیر است:

$$F_{(t+1)} = \alpha y_t + \alpha(1-\alpha)y_{(t-1)} + \alpha(1-\alpha)^2 y_{(t-2)} + \ldots$$

در رابطه بالا α ثابت هموارسازی نامیده می‌شود و یک مقدار ثابت بین صفر و یک است.

روش هولت-وینترز

روش پیش‌بینی هولت-وینترز توسعه‌ای از روش هموارسازی نمایی است. مدل هولت-وینترز، شامل سه بخش میانگین (مقدار ثابت)، بخش دوم روند (شیب خط) و بخش سوم تغییرات فصلی است. در چنین مدلی، پیش‌بینی مقدار آینده توسط ترکیب این سه مولفه که با α,β,γ نشان می‌دهند، انجام می‌شود. شکل ساده (بدون روند و تغییرات فصلی) به‌صورت رابطه زیر است

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)(b_{t-1})$$

که در آن y_t مشاهدات مربوط به زمان t و S_t مشاهده هموار شده در زمان t است. همان‌طور که در این مدل مشاهده می‌شود فقط پارامتر α حضور داشته و هر مشاهده به‌صورت ترکیب خطی از نقطه قبلی (St−1) دیده می‌شود. از طرفی I نیز شاخص تغییرات فصلی نامیده شده و L نیز طول دوره تغییرات فصلی است. در صورتی که مدل دارای روند هم باشد شکل مدل به‌صورت رابطه زیر خواهد بود:

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)(b_{t-1})$$

همان‌طور که مشخص است پارامتر گاما γ نیز در مدل اضافه شده است. در نهایت مدلی که به همراه تغییرات فصلی ارائه می‌شود به‌صورت زیر خواهد بود:

$$I_t = \beta\frac{y_t}{S_t} + (1-\beta)I_{t-L}$$

به منظور پیش‌بینی در مدل هولت-وینترز، از رابطه زیر استفاده می‌کنیم:

$$F_{t+m} = (S_t + mb_t)I_{t-L+m}$$

شبکه عصبی اتورگرسیو

در مدل شبکه عصبی اتورگرسیو پیش‌بینی در دو فاز صورت می‌گیرد. برای سری‌های زمانی مورد نظر، مرتبه مدل اتورگرسیو در فاز اول تعیین و در مرحله دوم، شبکه عصبی توسط مجموعه داده آموزشی با در نظر گرفتن مرتبه اتورگرسیو آموزش داده می‌شود. در این روش، مدل برازش شده با الگوی غیرفصلی از دو جزء p و k تشکیل شده که p نشان‌دهنده تعداد لگ‌های ورودی و k نشان‌دهنده تعداد نورون‌های پنهان است. بنابراین، این مدل به‌صورت ARNN(p,k) نشان داده می‌شود. همچنین مدل برازش شده با الگوی فصلی به‌صورت ARNN(p,P,k)(m) است که مشابه با ARIMA(p,0,0)(P,0,0)(m) با توابع غیرخطی است.

رگرسیون KNN

KNN مخفف عبارت K-Nearest Neighbors به معنی K نزدیک‌ترین همسایگی است که نام این الگوریتم اشاره به شیوه کار آن دارد. به طور خلاصه KNN مجموعه‌ای از نمونه‌های آموزشی را که با n ویژگی توصیف شده‌اند، ذخیره می‌کند. هر نمونه آموزشی نشان‌دهنده یک نقطه در فضای n بعدی است. با توجه به یک نمونه جدید، KNN نزدیکترین k نمونه‌های خود را در فضای n بعدی پیدا می‌کند، به این امید که اهداف آنها مشابه هدف ناشناخته‌اش باشد. حال می‌خواهیم ببینیم که چگونه KNN می‌تواند برای پیش‌بینی سری‌های زمانی اعمال شود. در این مورد، هدف مرتبط با یک نمونه آموزشی مجموعه‌ای از مقادیر سری زمانی است و ویژگی‌هایی که نمونه را توصیف می‌کنند، مقادیر عقب مانده هدف هستند، یعنی ما یک مدل اتورگرسیو داریم. برای مثال، با یک سری زمانی ماهانه شروع می‌کنیم که شامل ۱۳۲ مشاهده است، یعنی ۱۱ سال:

[DOI: 10.66224/ihio.6.2.113]

$t=\{x_{-1},x_{-2},x_{-3},x_{-4},x_{-5},x_{-6},x_{-7},\ldots,x_{132}\}$. همچنین فرض کنید می‌خواهیم ماه آینده را پیش‌بینی کنیم. از آنجا که قصد داریم تنها یک مقدار را پیش‌بینی کنیم، هدف یک نمونه آموزشی، مقدار سری زمانی است. همچنین فرض می‌کنیم ویژگی‌هایی که یک هدف را توصیف می‌کنند، ۱۲ مقدار اول تاخیر آن در سری‌های زمانی هستند. حال می‌خواهیم ببینیم کدام نمونه جدید برای پیش‌بینی مقدار آینده بعدی سری زمانی استفاده می‌شود. به دلیل اینکه از تاخیرهای ۱ تا ۱۲ به‌عنوان بردار ویژگی استفاده می‌کنیم، بردار ویژگی مرتبط با نقطه آینده بعدی، بردار $x_{121},x_{122},\ldots,x_{132}$ است که از ۱۲ مقدار آخر سری زمانی تشکیل شده است. اگر k برابر با ۲ باشد، ۲-نزدیکترین همسایگان نمونه جدید پیدا می‌شوند و اهداف آنها برای پیش‌بینی ماه آینده تجمیع می‌شود [۱۸].

مدل‌سازی تتا

این مدل بر اساس مفهوم اصلاح انحناهای محلی سری‌های زمانی است. این تغییر از ضریبی به نام Theta (به‌عنوان نماد از حرف یونانی Theta استفاده می‌شود) به دست می‌آید. اگر مقدار تتا به تدریج کاهش یابد، سری زمانی کاهش می‌یابد. هرچه مقدار ضریب تتا کوچک‌تر باشد، درجه قله‌ها کاهش پیدا می‌کند. در حالتی که $\theta=0$ است، سری زمانی به یک خط رگرسیون خطی تبدیل می‌شود. ضریب θ نیز می‌تواند مقادیر منفی داشته باشد. برعکس، اگر ضریب تتا افزایش یابد ($\theta>1$) آنگاه سری زمانی گشاد می‌شود. فرمول کلی روش تتا به شرح رابطه زیر است. سری زمانی اولیه به دو یا چند خط تتا تجزیه می‌شود. هر یک از خطوط تتا به طور جداگانه برون‌یابی شده و پیش‌بینی‌ها به سادگی ترکیب می‌شوند. به‌طور مثال $\theta=0$ و $\theta=2$

$$Data=\frac{1}{2}\left(L(\theta=0)+L(\theta=2)\right)$$

که در آن $L(\theta=0)$ مخفف خط تتا برای پارامتر θ برابر با صفر است. اولین خط تتا ($\theta=0$) خط رگرسیون خطی داده‌هاست و دومی دقیقاً دو برابر سری زمانی اولیه است. اولین جزء $L(\theta=0)$ سری زمانی را از طریق یک روند خطی توصیف می‌کند. مورد دوم، $L(\theta=2)$، انحناهای محلی را دو برابر کرده و رفتار کوتاه مدت را بزرگ‌نمایی می‌کند. اولین خط تتا به روش معمول برای یک روند خطی برون‌یابی می‌شود و دومی از طریق هموارسازی نمایی ساده برون‌یابی می‌شود [۱۹].

روش STL

STL یک روش فیلتر برای تجزیه یک سری زمانی به اجزای روند، فصلی و باقیمانده است. یعنی فرض کنید داده‌ها، جزء روند، مؤلفه فصلی و مؤلفه باقیمانده به ترتیب با S_v، T_v، Y_v و R_v برای $v=1,\ldots,N$ نشان داده می‌شوند. سپس: $Y_v=T_v+S_v+R_v$. STL متشکل از دنباله‌ای از عملیات هموارسازی است که از رگرسیون وزن محلی یا Loess استفاده می‌کند. STL چندین پارامتر دارد که باید توسط تحلیلگر داده انتخاب شود. STL ۶ پارامتر به شرح زیر دارد:

n(p): تعداد مشاهدات در هر چرخه فصلی/ n(i): تعداد عبور از حلقه داخلی/ n(o): تعداد تکرارهای استوارحلقه بیرونی/ n(l): پارامتر هموارسازی برای فیلتر پایین-گذر/ n(t): پارامتر هموارسازی برای مؤلفه روند/ n(s): پارامتر هموارسازی برای مؤلفه فصلی. اطلاعات بیشتر در مورد این روش به رفرنس [۲۰] مراجعه شود.

روش هموارسازی اسپلاین مکعبی

فرض کنید یک سری زمانی تک متغیره $y_t, t=1,\ldots,n$، با روند غیرخطی مشاهده می‌کنیم. ما علاقه‌مند به پیش‌بینی سری با برون‌یابی روند با استفاده از یک تابع خطی برآورد شده از سری زمانی مشاهده شده هستیم. برون‌یابی روند خطی بسیار مورد استفاده قرار می‌گیرد و در عمل عملکرد نسبتاً خوبی دارد. برای سری‌های زمانی با فواصل مساوی، یک هموارسازی اسپلاین مکعبی را می‌توان به‌عنوان تابع f(t) تعریف کرد که بر روی همه توابع دو بار قابل تمایز f در S که در آن $[R\ \ S]$ به $n,]$ حداقل می‌رسد. پارامتر هموارسازی λ نرخ تغییرات بین خطای باقیمانده و تغییرات محلی را کنترل می‌کند. مقادیر بزرگ λ تابع f(t) را نزدیک به یک خط صاف نشان می‌دهد در حالی که مقادیر کوچک λ یک تابع f(t) بسیار پر انحنا (پر پیچ و خم) نشان می‌دهد. برای اطلاعات بیشتر در مورد این روش به رفرنس [۲۱] مراجعه شود.

[DOI: 10.66224/ihio.6.2.113]

روش BATS

شناسه BATS مخفف ویژگی‌های کلیدی مدل است: تبدیل Box-Cox، خطاهای ARMA، روند و اجزای فصلی. برای نشان دادن پارامتر Box–Cox، پارامترهای تعدیل، پارامترهای ARMA یعنی p و q، و دوره‌های فصلی $(m_1,\ldots,m_T)$، با آرگومان‌هایی $(\omega, \varphi,p,q,m1,m_2,\ldots,m_T)$ تکمیل می‌شود.

مدل BATS واضح‌ترین تعمیم مدل‌های نوآوری فصلی سنتی است که برای دوره‌های فصلی متعدد امکان‌پذیر است. مولفه فصلی اولیه به تنهایی حاوی mT غیر صفر است. برای اطلاعات بیشتر در مورد این روش به رفرنس [۲۲] مراجعه شود.

معیارهای ارزیابی

میانگین مربعات خطا (MSE)

$$MSE(Y,\hat{Y}) = 1/n\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2 = 1/n\sum_{i=1}^{n}e_i^{\,2}$$

جذر میانگین مربعات خطا (RMSE)

$$RMSE(Y,\hat{Y}) = \sqrt{MSE(Y,\hat{Y})} = \sqrt{1/n\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2}$$

میانگین قدر مطلق خطا (MAE)

$$MAE(Y,\hat{Y}) = 1/n\sum_{i=1}^{n}|Y_i-\hat{Y}_i| = 1/n\sum_{i=1}^{n}|e_i|$$

میانگین درصد قدرمطلق خطا (MAPE)

$$MAPE(Y,\hat{Y}) = 100/n\sum_{i=1}^{n}\left|\frac{Y_i-\hat{Y}_i}{Y_i}\right| = 100/n\sum_{i=1}^{n}\left|\frac{e_i}{Y_i}\right|$$

که در آن y_i مقدار پیش‌بینی و Y_i مقدار واقعی است.

یافته‌ها

در این بخش نتایج برازش مدل و پیش‌بینی مقادیر تجمعی کرونا (مقادیر تجمعی بستری و مقادیر تجمعی فوت) برای شهرستان مهریز، با توجه به مدل‌های ذکر شده در بخش قبلی آورده شده است.

داده‌های کرونای مورد استفاده در این مقاله ۶۶۲ داده ، از تاریخ ۱۳۹۸/۱۲/۰۷ الی ۱۴۰۰/۰۹/۲۸ به‌صورت روزانه است که از معاونت بهداشت و درمان استان یزد اخذ شده است. داده‌ها در هر روز شامل تعداد افرادی است که تست کرونای آنها مثبت اعلام شده و افرادی که تست آنها مثبت شده و در نهایت فوت کرده‌اند. لازم به ذکر است آمار موجود برای افرادی است که در بیمارستان شهرستان مهریز پذیرش و بستری شده‌اند و شامل افرادی که در مراکز درمانی دیگر یا به اصطلاح مراکز درمان سرپایی مورد پذیرش قرار گرفته باشند، نمی‌شود.

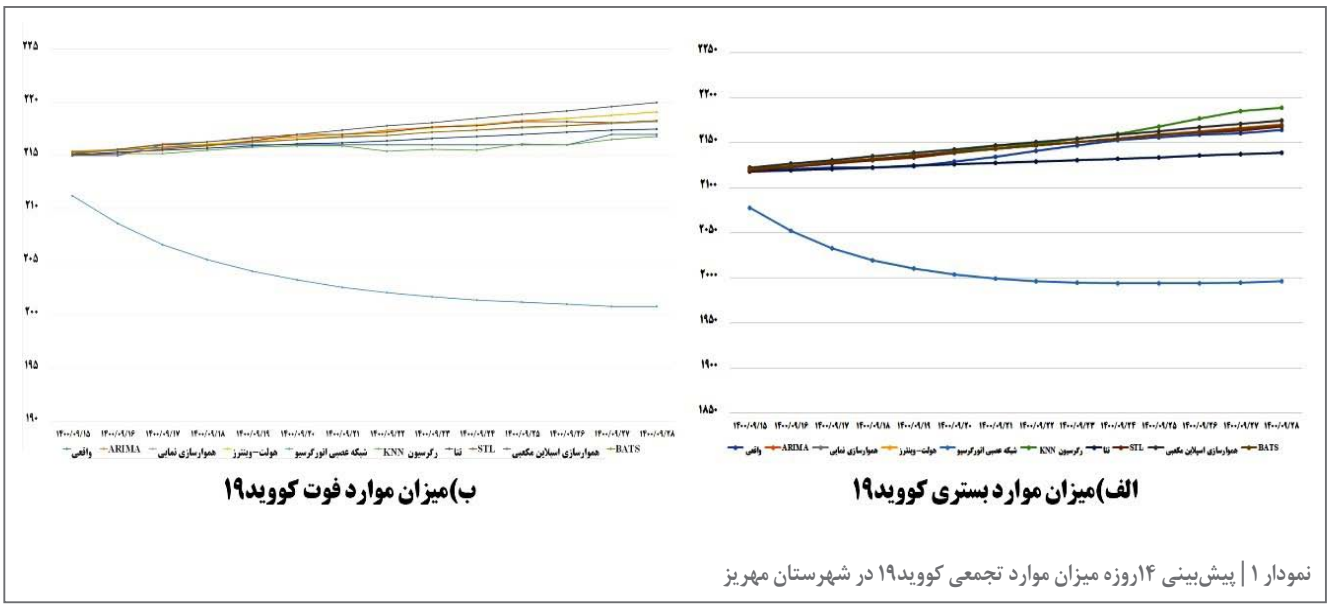
با توجه به تحلیل داده‌ها و نتایج به دست آمده مطابق جدول ۱، بهترین روش و مدل برای موارد تجمعی بستری کووید-۱۹ شهرستان مهریز با توجه به معیارهای ارزیابی یا همان خطاهای پیش‌بینی، مدل STL با معیارهای ارزیابی MSE=۳۱/۸۴، RMSE=۵/۶۴۲۷، MAE=۴/۷۱۳۹ و MAPE=۰/۲۲۰۷ است که برای آینده از این روش برای پیش‌بینی موارد تجمعی بستری کووید-۱۹ در شهرستان مهریز استفاده می‌کنیم. بهترین روش و مدل برای موارد تجمعی فوت کووید-۱۹ شهرستان مهریز با توجه به معیارهای ارزیابی یا همان خطاهای پیش‌بینی، مدل رگرسیون KNN با معیارهای ارزیابی MSE=۰/۱۲۸۷، RMSE=۰/۳۵۸۸، MAE=۰/۳۹۵۵ و MAPE=۰/۱۳۶۷ است که برای آینده از این روش برای پیش‌بینی موارد تجمعی فوت کووید-۱۹ در شهرستان مهریز استفاده می‌کنیم.

برای بررسی بهتر عملکرد این ۹ روش نمودارهای ۱و۲ ترسیم شده‌اند. بر اساس بررسی نمودارها، به نظر می‌رسد که مدل‌های ARIMA،STL، هموارسازی نمایی، هولت-وینترز و BATS برای پیش‌بینی موارد جدید بستری کووید-۱۹ دارای خطای کمتری هستند. این مدل‌ها را برای پیش‌بینی موارد تجمعی بستری شهرستان مهریز در یک سطح می‌توان در نظر گرفت که از این میان مدل STL برای پیش‌بینی موارد تجمعی بستری کووید-۱۹ دارای کمترین خطاست. همچنین برای موارد مرگ‌ومیر روش‌های تتا و روش رگرسیون KNN دارای خطاهای کمتری نسبت به بقیه روش‌ها هستند که روش رگرسیون KNN دارای کمترین خطا برای پیش‌بینی موارد تجمعی فوت است. همچنین مدل شبکه‌های عصبی اتورگرسیو دارای بدترین عملکرد در میان دیگر مدل‌ها، هم برای موارد بستری و هم موارد فوت، است.

[DOI: 10.66224/ihio.6.2.113]

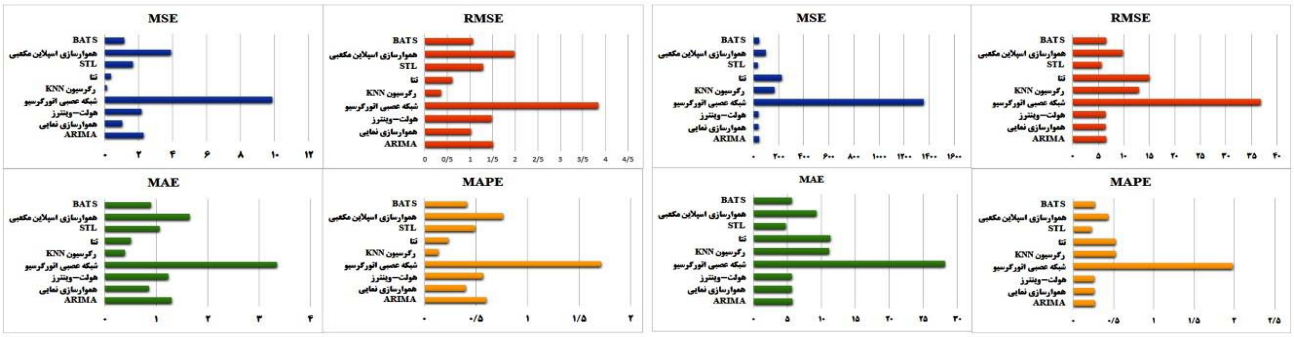
جدول ۱ | معیارهای ارزیابی مدل‌های مربوط به موارد تجمعی

موارد تجمعی بستری				
مدل	MSE	RMSE	MAE	MAPE
ARIMA	۴۳/۳۸۸۵	۶/۵۸۷۰	۵/۷۴۲۷	۲۶۸۸/۰
هموارسازی نمایی	۴۱/۷۶۲۶	۶/۴۶۲۴	۵/۵۷۴۶	۲۶۱۰/۰
هولت-وینترز	۴۱/۷۲۷۷	۶/۴۵۹۷	۵/۵۷۱۳	۲۶۰۸/۰
شبکه عصبی اتورگرسیو	۱۳۵۴/۸۱۱۹	۳۶/۸۱۲۶	۲۸/۱۴۸۸	۹۷۸۱/۱
رگرسیون KNN	۱۶۸/۰۴۱۹	۱۲/۹۶۳۱	۱۱/۱۴۱۵	۵۱۹۱/۰
تتا	۲۲۴/۴۶۳۳	۱۴/۹۸۲۱	۱۱/۳۲۵۰	۵۲۵۹/۰
STL	۳۱/۸۴۰۰	۵/۶۴۲۷	۴/۷۱۳۹	۲۲۰۷/۰
اسپلاین مکعبی	۹۵/۷۱۸۸	۹/۷۸۳۶	۹/۲۳۶۲	۴۳۱۸/۰
BATS	۴۲/۱۲۱۳	۶/۴۹۰۱	۵/۶۱۰۹	۲۶۲۷/۰
موارد تجمعی مرگ‌ومیر				
مدل	MSE	RMSE	MAE	MAPE
ARIMA	۲/۲۸۱۶	۱/۵۱۰۵	۱/۲۹۸۸	۰/۶۰۰۸
هموارسازی نمایی	۱/۰۲۴۲	۱/۰۲۲۱	۰/۸۶۰۷	۰/۳۹۸۲
هولت-وینترز	۲/۱۶۵۳	۱/۴۷۱۵	۱/۲۳۳۱	۰/۵۷۰۴
شبکه عصبی اتورگرسیو	۹/۸۲۷۰	۳/۸۳۸۵	۳/۳۴۱۱	۱/۷۱۰۱
رگرسیون KNN	۰/۱۲۸۷	۰/۳۵۸۸	۰/۳۹۵۵	۰/۱۳۶۷
تتا	۰/۳۶۳۴	۰/۶۰۲۹	۰/۵۰۲۹	۰/۲۳۲۷
STL	۱/۶۶۷۴	۱/۲۹۱۳	۱/۰۶۱۳	۰/۴۹۱۰
اسپلاین مکعبی	۳/۸۹۳۱	۱/۹۷۳۱	۱/۶۴۲۹	۰/۷۵۹۹
BATS	۱/۱۳۷۸	۱/۰۶۶۷	۰/۹۰۰۰	۰/۴۱۶۴



نمودار ۱ | پیش‌بینی ۱۴روزه میزان موارد تجمعی کووید۱۹ در شهرستان مهریز

[DOI: 10.66224/ihio.6.2.113]



نمودار ۲ | معیارهای ارزیابی موارد تجمعی کووید۱۹ در شهرستان مهریز

بحث

مهم‌ترین نکته‌ای کــه می‌توان از نتایج تحلیل برداشــت کرد، این است که نمی‌توان مدل یکتایی برای موارد بستری و فوت در نظر گرفت. همچنین با اســتفاده از مدل‌هایی که در ســال‌های خیلی قبل معرفی شده‌اند و یا به اصطلاح مدل‌های قدیمی، می‌توان در عمل نتایج بهتری کســب کرد، در حالی که در مطالعه موردی، عملکرد مدل‌های جدیدتر مثل مدل شــبکه‌های عصبی اتورگرســیو نتایج بدتــری دارد. بنابراین می‌توان نتیجه گرفت که اســتفاده از مدل‌های جدید در عمل همیشــه جوابگو نیســت و با توجه به نوع داده‌ها ممکن اســت مدل‌های قدیمی نتیجه و کارآیی بهتری در پیش‌بینی داشــته باشند. همچنین می‌توان از نتایج حاصله برای پیش‌بینی و کنترل بیماری‌های ویروسی دیگر که ممکن اســت در آینده به‌عنوان یــک پاندمی و یا حتی اپیدمی و اندمی ظهور کنند، استفاده کرد.

با توجه به ناشــناخته بودن رفتار این ویروس و پژوهش‌های صورت گرفتــه در پیش‌بینی کووید-۱۹، مهم‌ترین جنبه نــوآوری این مطالعه اســتفاده از ۹ روش متفــاوت در زمینه‌های مختلف بــرای موارد جدید بســتری و مرگ‌ومیر کرونا در شهرســتان مهریز، برخلاف پژوهش‌های انجام شده که از ۲ الی ۳ روش در زمینه خاصی بهره گرفته‌اند، است. در ســایر مطالعات صورت گرفته مقایسه کاملی بین روش‌ها انجام نشده و همان‌طور که در منابع مختلف اشاره شده، راجع به هر موضوعی در هر موقعیتی باید از روش‌های متعدد پیش‌بینی استفاده کرد تا بهترین روش با توجه به نوع داده‌ها انتخاب شــود. در این مطالعه سعی شده که کلیه روش‌های معتبر و مناسب برای پیش‌بینی کرونا مورد بحث قرار بگیرد.

در این میان نکته‌ای که حائز اهمیت است این است که باید داده‌ها در زمان واقعی به‌روز شوند، بدین‌معنا که با توجه به شرایط جدید جامعه (مثلاً واکسیناسیون، ظهور سویه‌های جدید) باید مدل‌ها به‌روز شوند.

مطالعــه حاضر محدودیت‌هایی را به همراه داشــت. داده‌ها به دلیل محدودیت فقط شامل افرادی اســت که در بیمارستان بستری شده‌اند و شــامل افرادی که در مراکز درمانی سرپایی پذیرش شدند، نمی‌شود. همچنین به دلیل اینکه عوامل بسیاری در شیوع بیماری‌های همه‌گیر، مثل کووید-۱۹، دخیل هستند، دارای پیچیدگی‌های زیادی است؛ مانند ظهور سویه‌های جدید که دومین محدودیت مطالعه حاضر است.

نتیجه‌گیری

در مطالعــه حاضر، از ۹ روش متفاوت برای پیش‌بینی تعداد موارد تجمعی بستری و مرگ‌ومیر کووید-۱۹ در شهرستان مهریز-یزد به‌عنوان یک مطالعه موردی، اســتفاده شــد. این روش‌ها با استفاده از معیارهای ارزیابی MAE، RMSE، MSE و MAPE با یکدیگر مقایسه شدند. بهترین مدل برای پیش‌بینی موارد تجمعی بســتری کووید-۱۹ در شهرســتان مهریــز، مدل STL اســت، در حالی که برای مــوارد تجمعی فوت، مدل رگرسیون KNN به‌عنوان بهترین مدل از نظر معیارهای ارزیابی برگزیده شد. در این میان نکته‌ای که حائز اهمیت است، این است که باید داده‌ها در زمان واقعی به‌روز شوند.

همچنین با توجــه به نتایج این مطالعه به محققان آتی پیشــنهاد میشود که پیش‌بینی کووید-۱۹ با در نظر گرفتن اثر استفاده از ماسک و رعایت فاصله اجتماعی بررسی شود، کووید-۱۹ را با ترکیب روش‌ها مثل روش تتا و STL پیش‌بینی کنند، از روش‌های اســتفاده شده برای دیگر بیماری‌های واگیردار یا پاندمیک برای پیش‌بینی استفاده شود و در آخر این روش‌ها را می‌توان در مکانی دیگر مدل‌سازی و نتیجه‌گیری کرد.

تأییدیه اخلاقی: این مقاله حاصل بخشی از پایان‌نامه مصوب به شماره ۱۶۴۳۴۱۰ و تأییدیه اخلاقی مصوب IR.YAZD.REC.1401.094 دانشگاه یزد است.
تعارض منافع: بین نویسندگان هیچ‌گونه تعارضی در منافع وجود ندارد.
سهم نویسندگان: هر دو نویسنده سهم یکسانی در نگارش این مقاله داشته‌اند.
منابع مالی: این مقاله از سوی هیچ نهادی مورد حمایت مالی قرار نگرفته است.

[DOI: 10.66224/ihio.6.2.113]

References

1. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed. 2020;91(1):157.
2. Zhang X, Liu Y, Yang M, Zhang T, Young AA, Li X. Comparative study of four time series methods in forecasting typhoid fever incidence in China. PloS One. 2013;8(5):e63116. doi: 10.1371/journal.pone.0063116.
3. Olanrewaju SO, Ojo EO, Oguntade ES. Time Series Analysis on Reported Cases of Tuberculosis in Minna Niger State Nigeria. Open J Stat. 2020;10(3):412-30. doi: 10.4236/ojs.2020.103027.
4. Nayak MS, Narayan KA. Forecasting dengue fever incidence using ARIMA analysis. Int J Collab Res Intern Med Public Health. 2019;11(3):924-32.
5. Wu W, An SY, Guan P, Huang DS, Zhou BS. Time series analysis of human brucellosis in mainland China by using Elman and Jordan recurrent neural networks. BMC Infect Dis. 2019;19(1):414. doi: 10.1186/s12879-019-4028-x.
6. Ceylan Z. Estimation of COVID-19 prevalence in Italy, Spain, and France. Sci Total Environ. 2020;729:138817. doi: https://doi.org/10.1016/j.scitotenv.2020.138817.
7. Bayyurt L, Bayyurt B. Forecasting of COVID-19 cases and deaths using ARIMA models. MedRxiv. 2020:2020-04. doi: 10.1101/2020.04.17.20069237.
8. Tandon H, Ranjan P, Chakraborty T, Suhag V. Coronavirus (COVID-19): ARIMA-based time-series analysis to forecast near future and the effect of school reopening in India. J Health Manag. 2022;24(3):373-88. doi: 10.1177/09720634221109087.
9. Perone G. An ARIMA model to forecast the spread and the final size of COVID-2019 epidemic in Italy. MedRxiv. 2020:2020-04. doi: 10.1101/2020.04.27.20081539.
10. Kibria HB, Jyoti O, Matin A. Forecasting the spread of the third wave of COVID-19 pandemic using time series analysis in Bangladesh. Inform Med Unlocked. 2022;28:100815. doi: 10.1016/j.imu.2021.100815.
11. Khan FM, Gupta R. ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. Journal of Safety Science and Resilience. 2020;1(1):12-8. doi: 10.1016/j.jnlssr.2020.06.007.
12. Chyon FA, Suman MN, Fahim MR, Ahmmed MS. Time series analysis and predicting COVID-19 affected patients by ARIMA model using machine learning. J Virol Methods. 2022;301:114433. doi: 10.1016/j.jviromet.2021.114433.
13. Mukhairez HHA, Alaff AJI. Short-term Forecasting of COVID-19. In: Kose, U., Watada, J., Deperlioglu, O., Marmolejo Saucedo, J.A. (eds) Computational Intelligence for COVID-19 and Future Pandemics. Disruptive Technologies and Digital Transformations for Society 5.0. Springer, Singapore, 2022. doi: 10.1007/978-981-16-3783-4_12.
14. Li C, Sampene AK, Agyeman FO, Robert B, Ayisi AL. Forecasting the severity of COVID-19 pandemic amidst the emerging SARS-CoV-2 variants: adoption of ARIMA model. Comput Math Methods Med. 2022;2022:3163854. doi: 10.1155/2022/3163854.
15. Yaffee RA, McGee M. An introduction to time series analysis and forecasting: with applications of SAS® and SPSS®. Elsevier; 2000 May 12.
16. Fanoodi B, Malmir B, Jahantigh FF. Reducing demand uncertainty in the platelet supply chain through artificial neural networks and ARIMA models. Comput Biol Med. 2019;113:103415. doi: 10.1016/j.compbiomed.2019.103415.
17. Kufel T. ARIMA-based forecasting of the dynamics of confirmed Covid-19 cases for selected European countries. Equilibrium. Quarterly Journal of Economics and Economic Policy. 2020;15(2):181-204. doi: 10.24136/eq.2020.009.
18. Martínez F, Frías MP, Charte F, Rivera AJ. Time Series Forecasting with KNN in R: the tsfknn Package. R J. 2019;11(2):229-42. doi: 10.32614/RJ-2019-004.
19. Assimakopoulos V, Nikolopoulos K. The theta model: a decomposition approach to forecasting. Int J Forecast. 2000;16(4):521-30. doi: 10.1016/S0169-2070(00)00066-2.
20. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: A seasonal-trend decomposition. J Off Stat. 1990;6(1):3-73.
21. Hyndman RJ, King ML, Pitrun I, Billah B. Local linear forecasts using cubic smoothing splines. Aust N Z J Stat. 2005;47(1):87-99. doi: 10.1111/j.1467-842X.2005.00374.x.
22. De Livera AM, Hyndman RJ, Snyder RD. Forecasting time series with complex seasonal patterns using exponential smoothing. J Am Stat Assoc. 2011;106(496):1513-27. doi: 10.1198/jasa.2011.tm09771.

[DOI: 10.66224/ihio.6.2.113]