ORIGINAL ARTICLE

Predicting Fatality and Injuries of Traffic Accidents by Conventional Time Series and Neural Network Modeling in Iran

Mohsen Arami Sham Asbi[1]
Jamshid Yazdani Cherati[2]
Reza Ali Mohammadpour[2]

[1] Master's degree, Biostatistics, Mazandaran University of Medical Sciences, Sari, Iran
[2] Professor, Department of Biostatistics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

(Received May 17, 2023; Accepted November 6, 2023)

Abstract

Background and purpose: Traffic accidents are one of the most important causes of death and disability in the world, which cause much damage to countries and millions of people every year. Many researchers have focused on the time series caused by different aspects of these accidents. Box-Jenkins is one of the most common and widely used time series forecasting models, which is used under the condition of model stationarity. With the advancement of technology and the development of neural networks in time series, the discussion of comparing the predictive power of these models with traditional time series models has been raised. In this study, we investigated the traffic accidents leading to death or injury in iran during 2011 to 2021.

Materials and methods: Statistical analyses related to the SARIMA time series model have been performed by using Minitab and EViews and the LSTM neural network model by Python and Visual Studio Code.

Results: During period of the study, The trend of the number of deaths has always been decreasing, but the number of injured has been increasing before the coronavirus epidemic and decreasing after that. The highest number of dead and injured has always been in Shahrivar (August and September) every year and the months with 9 days off had significantly less deaths and more injuries. 21.38% and 27.56% of dead and injured were women. The ages of 20 to 29 had the highest number of deaths and injuries, but children and the elderly were more vulnerable and their share in the deaths was more than the injuries. According to the results, as the weather worsens, the share of accidents leading to death increases. In comparison of SARIMA and LSTM the results indicated that LSTM has been better in estimating data trends. The forecast shows the increasing trend of fatality and injuries in the comind years.

Conclusion: Compared to the SARIMA model, the LSTM model showed better performance in predicting trend and time series components.

Keywords: Box-Jenkins, LSTM, Neural Network, SARIMA, Time Series, Traffic Accidents

J Mazandaran Univ Med Sci 2023; 33 (Supple 1): 205-215 (Persian).

Corresponding Author: Jamshid Yazdani-Charati - Antimicrobial Resistance Research Center, Communicable Diseases Institute, Mazandaran University of Medical Sciences, Sari, Iran. (E-mail: jamshid.charati@gmail.com)

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

پژوهشی

مجله دانشگاه علوم پزشکی مازندران
دوره سی و سوم شماره ویژه نامه 1 آبان سال 1402 (215-205)

مقایسه مدل‌های SARIMA و LSTM در برآورد سری زمانی فوتی‌ها و مصدومان حوادث ترافیکی سال‌های 1390 تا 1400

محسن آرامی شام اسبی[1]
جمشید یزدانی چراتی[2]
رضاعلی محمدپور[2]

چکیده

سابقه و هدف: حوادث ترافیکی یکی از مهم‌ترین دلایل مرگ و میر و به وجود آمدن ناتوانی در جهان است که سالانه آسیب‌های فراوانی را به کشورها و میلیون‌ها انسان تحمیل می‌کند. سری‌های زمانی ناشی از ابعاد مختلف این حوادث مورد توجه پژوهشگران بسیاری قرار می‌گیرد. یکی از سری‌های زمانی متداول و پرکاربرد در مدل‌بندی و پیش‌بینی سری‌های زمانی، مدل‌های باکس- نکینز هستند که شرط استفاده از آن‌ها ایستایی مدل است. با پیشرفت فناوری و ظهور مدل‌های شبکه عصبی و گسترش کاربرد انواع آن‌ها در سری‌های زمانی، بحث مقایسه قدرت پیش‌بینی این مدل‌ها با مدل‌های سنتی سری زمانی پیش آمده است. در این مطالعه، سری زمانی تعداد فوتی‌ها و مجروحین حوادث ترافیکی در ایران در سال‌های 1390 تا 1400 بررسی شده است.

مواد و روش‌ها: تحلیل‌های آماری مربوط به مدل سری زمانی SARIMA با استفاده از نرم‌افزار Minitab و EViews و مدل شبکه عصبی LSTM با استفاده از نرم‌افزارهای Python و Visual Studio Code انجام شده است.

یافته‌ها: در بازه زمانی مطالعه، روند تعداد فوتی‌ها همواره کاهشی ولی مصدومان تا پیش از همه‌گیری کرونا افزایشی و پس از آن کاهشی شده و همواره بیش‌ترین تعداد فوتی و مصدوم در شهریورماه هر سال بوده و ماه‌های با تعطیلی نه روز به شکل معنی‌داری فوتی کم‌تر و مصدوم بیش‌تری داشتند. به ترتیب 21/38 درصد و 27/56 درصد از فوتی‌ها و مصدومان زنان بودند. سنین 20 تا 29 سال بیش‌ترین تعداد فوتی و مصدوم را داشتند اما کودکان و افراد مسن آسیب‌پذیرتر بودند و سهمشان در فوتی‌ها بیش‌تر از مصدومان است. بنابه نتایج، هرچه هوا بدتر شده است سهم حوادث منجر به فوت بیش‌تر شده است. در مقایسه SARIMA و LSTM یافته‌ها حاکی از برتری مدل LSTM در برآورد بهتر روند داده‌ها بودند. پیش‌بینی صورت گرفته نشان از روند افزایشی تعداد فوتی‌ها و مصدومان طی سال‌های آتی دارد.

استنتاج: مدل LSTM در مقایسه با مدل SARIMA عملکرد بهتری را در پیش‌بینی روند و تغییرات موجود در سری زمانی از خود نشان داد.

واژه های کلیدی: سری زمانی، باکس-جنکینز، SARIMA، شبکه عصبی، LSTM، حوادث ترافیکی

مقدمه

حوادث ترافیکی یکی از مهم‌ترین دلایل مرگ‌ومیر و به وجود آمدن ناتوانی در جهان است. استفاده‌کنندگانِ جاده در کشورهای با درآمدهای کم و متوسط در معرضِ آسیب بیش‌تری هستند(1).

مولف مسئول: جمشید یزدانی چراتی- ساری، کیلومتر 17 جاده خزرآباد، مجتمع دانشگاهی پیامبر اعظم(ص)، دانشکده بهداشت E-mail: jamshid.charati@gmail.com
۱. کارشناسی ارشد آمار زیستی، دانشگاه علوم پزشکی مازندران، ساری، ایران
2. استاد گروه آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران، ساری، ایران
تاریخ دریافت: 1402/3/27 تاریخ ارجاع جهت اصلاحات: 1402/4/3 تاریخ تصویب: 1402/8/15

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

حوادث ترافیکی یکی از مهم‌ترین دلایل مرگ‌ومیر و به وجود آمدن ناتوانی در جهان است. استفاده‌کنندگانِ جاده در کشورهای با درآمدهای کم و متوسط در معرض آسیب بیش‌تری هستند(1). طبق آخرین گزارش سازمان جهانی بهداشت (WHO) در سال 2018، تصادفات به عنوان هشتمین دلیل مرگ‌ومیر انسان‌ها، روزانه جان 3700 نفر و سالانه بیش از 1/35 میلیون انسان را می‌گیرد. بیش‌ترین درصد این نوع مرگ‌ومیر، مربوط به گروه سنی 29-5 سال است. هرساله بیش از 50 میلیون نفر دیگر نیز مجروح یا معلول شده و اثرات ماندگاری در زندگی‌شان باقی می‌ماند(2). برآورد سازمان جهانی بهداشت در مورد تعداد مرگ‌ومیر سالیانه ناشی از تصادفات در ایران در سال 2020 حدود 25 هزار نفر تخمین زده شده است و پیش‌بینی می‌شود حدود زیان حاصل از این موضوع نیز شش درصد از تولید ناخالص ملی (GDP) باشد(3). با توجه به اهمیت و ماهیت داده‌های حوادث ترافیکی، محققان مختلفی از مدل‌های باکس-جنکینز برای سری زمانی ناشی از تلفات حوادث ترافیکی استفاده می‌کنند(4, 5). اما با پیشرفت مدل‌های مبتنی بر شبکه عصبی مصنوعی مقایسه‌های زیادی در کاربرد این مدل‌ها در پیش‌بینی سری‌های زمانی و مقایسه کیفیت برآوردهای به دست آمده با سری‌های زمانی کلاسیک صورت گرفته است که نتایج حاصله نشان از برتری این مدل‌ها بر روش‌های سنتی دارد(6-8). باتوجه به اهمیتی که حوادث ترافیکی بر کاهش شاخص‌های سلامت جامعه دارند و هم‌چنین، لطمه‌های اجتماعی، اقتصادی و به‌ویژه انسانی ناشی از آن‌ها، پژوهش حاضر به بررسی دقیق این حوادث می‌پردازد تا با شناخت بهتر شرایط موجود، دقیق‌ترین پیش‌بینی‌های ممکن با ضریب اعتماد مناسب از آینده وضعیت موجود ارائه شود. نتایج این پژوهش می‌تواند به امکان ریشه‌یابی، مهار و چاره‌اندیشی برای یافتن راهکارهای علمی‌تر و کارآمدتر، به منظور کاهش آسیب‌ها و هزینه‌های جامعه از حوادث ترافیکی را در اختیار مدیریت‌های حوزه ترافیکی و بهداشتی جامعه قرار دهد. از آن‌جایی که داده‌های مطالعه مورد بررسی، به‌صورت پیاپی می‌باشند و جنبه دنباله‌ای بر حسب زمان دارند، از روش روش‌های مبتنی بر سری زمانی متداول و نیز با وجود تعداد مناسب رکوردهای زمانی (تعداد رخداد روزانه، هفتگی و ماهانه در طی سال‌های 1398-1387)، می‌توان از شبکه‌های عصبی بازگشتی (که دارای ورودی به‌صورت دنباله‌ای هستند)، استفاده کرد. در حالت خاص، تمرکز اصلی در سری‌های زمانی، بر روی مدل بسیار پرکاربرد باکس- جنکینز و در شبکه‌های عصبی بازگشتی، بر روی شبکه LSTM که دقت و عملکرد مناسب آن‌ها در حوزه چنین تحلیل‌هایی در مطالعات گوناگون مورد تأیید قرار گرفته است، خواهد بود. هم‌چنین، تلاش خواهد شد همه روش‌های توصیفی و تحلیلی بکار رفته بر روی داده‌ها، به شکلی بهینه و گویا ارائه شوند.

مواد و روش‌ها

منابع داده‌ها

این پژوهش با مصوبه اخلاق شناسه IR.MAZUMS.REC.1398.1370 انجام شده است و اطلاعات مورد نیاز به‌صورت ثبتی بوده و داده‌های ثبتی مربوط به تصادفات جرحی و منجر به فوت سال‌های 1390 تا 1400 در ایران از مرکز تحقیقات کاربردی پلیس راهنمایی و رانندگی گرفته شد و پس از آماده‌سازی و تطبیق با گزارش‌های سایت پزشکی قانونی(9) جهت انجام تحلیل‌ها مورد استفاده قرار گرفتند.

آماده‌سازی داده‌ها

داده‌های اولیه دارای نواقص و رکوردهای گمشده بودند به همین دلیل پس از تطبیق با اطلاعات مرجع پزشکی قانونی، موارد مورد نیاز برآورد شده و با مبنای

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

ماهانه به صورت دو سری زمانی با نام‌های Fatal برای مرگ‌های ناشی از حوادث ترافیکی و Injury برای مصدومان این حوادث مرتب شدند. برای انجام آزمون و مقایسه مدل‌ها، از 106 ماه (80 درصد) ابتدایی جهت مرحله آموزش (برازش مدل) و از 26 ماه (20 درصد) آخر برای آزمون (بررسی عملکرد مدل) استفاده شد.

مدل سری زمانی باکس جنکینز (Box-Jenkins)

سری زمانی برای مجموعه مشاهدات دنباله‌ای که بر حسب زمان مرتب شده باشد با هدف توصیف و تشریح گذشته، پیش‌بینی یا شبیه‌سازی و کنترل و بدون نیاز به برازش مدل ریاضی، به کار می‌رود. سری زمانی با تعداد محدودی داده، تغییرات موجود در روند مشاهده داده‌ها را شناسایی نموده و برای پیش‌بینی آینده نیز همان تغییرات را اعمال می‌کند. توجه سری‌های زمانی به وجود خودهمبستگی داده‌ها در طول زمان و روند و تغییرات فصلی موجب می‌شود از سایر روش‌های نظارتی متمایز باشد. یکی از معروف‌ترین مدل‌های آماری در سری‌های زمانی، استفاده از روش باکس جنکینز است که روشی بر مبنای شناسایی مدل با استفاده از نمودارهای خودهمبستگی(Auto correlation)، برآورد پارامترهای مدل و از آزمون پایایی مدل است(10). سه نوع فرآیند داده بر سری‌های زمانی اعمال می‌شود:

1.روندزدایی(Detrending)

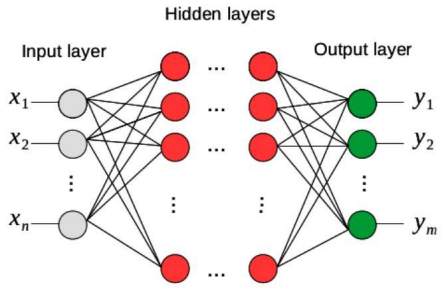
2.غیرفصلی‌سازی (Depersonalization)

3. روندزدایی و غیرفصلی‌سازی به صورت همزمان.

فرآیند روندزدایی، یک روند زمانی خطی بر داده‌ها برازش می‌دهد و سپس روند تخمین زده شده را از سری کم می‌کند(11).

مدل LSTM(Long Short-Term Memory)

شبکه عصبی مصنوعی یک روش یادگیری ماشین و پردازش اطلاعات است که دارای ویژگی‌های خاص عملکردی مشابه شبکه عصبی زیستی دارد و به مثابه یک توسعه مدل‌های ریاضی از شناخت انسان یا شبکه عصبی زیستی گسترش پیدا کرده است که پردازش اطلاعات درون واحدهای ساده زیادی به نام نرون انجام می‌پذیرد. سیگنال‌ها توسط لینک‌های ارتباطی بین نرون‌ها منتقل می‌شوند و هرکدام از این لینک‌های ارتباطی وزن بخصوصی دارند که روی سیگنال منتقل شده اعمال می‌شوند(12). اساس کار شبکه‌های عصبی بر پایه وجود یک سری مثال آموزشی است که آن‌ها را دریافت کرده و سپس نظامی ایجاد می‌کند که می‌تواند آن‌ها را یاد بگیرد. به عبارتی شبکه عصبی، با استفاده از مثال‌ها قوانینی را برای تشخیص خروجی استنباط می‌کند. علاوه براین افزایش تعداد مثال‌های آموزشی شبکه (مرحله تمرین) موجب به دست آمدن اطلاعات بیش‌تر و در نتیجه بهبود دقت می‌شود (تصویر شماره 1)(13).

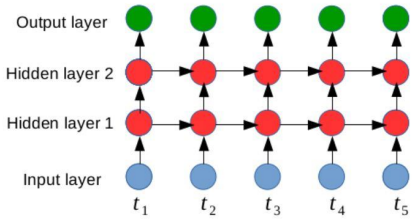


تصویر شماره 1. لایه‌های شبکه عصبی مصنوعی(14)

شبکه‌های عصبی بازگشتی(Recurrent neural network) یک مدل قوی برای داده‌های دنباله‌ای هستند. RNNها به‌صورت ذاتی در زمان عمیق هستند، زیرا لایه پنهان آن‌ها تابعی از تمام لایه‌های پنهان قبلی است. به‌جای ترکیب RNNها می‌توان آن‌ها را به‌صورت end-to-end مورد استفاده قرار داد. آن‌ها هم‌چنین از مزیت عمق در فضا نیز سود می‌برند، بدین‌صورت که چندین لایه پنهان می‌توانند به‌صورت پشت سرهم که منبع تغذیه یکدیگر باشند قرار بگیرند(15). در شبکه‌های عصبی بازگشتی مرسوم، فقط ارتباط کامل بین لایه‌های مجاور

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

وجود دارد اما هیچ ارتباطی بین گره‌ها در یک لایه وجود ندارد. این نوع از شبکه ممکن است درون وادادگی گرفتار شوند زیرا زمانی‌که با داده‌های زمانی فضایی سروکار داریم همیشه اثر متقابلی بین گره‌ها وجود دارد (تصویر شماره 2)(16).

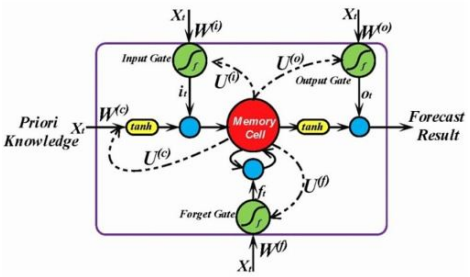


تصویر شماره 2. عملکرد شبکه عصبی بازگشتی(14)

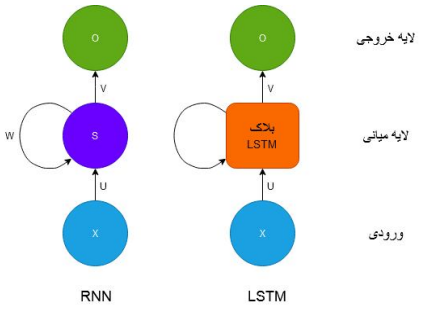
LSTM یکی از انواع شبکه‌های عصبی بازگشتی است و توسط Hochester و Schmid Huber در جهت کمک به گمشدگی‌ها و جهش در شیب RNNهای سنتی طراحی شده است و دارای بلوک‌های حافظه است که هر بلوک شامل سلول‌های حافظه و دروازه است. این سلول‌های حافظه قادر به یادآوری وضعیت زمانی شبکه با ارتباطات درون خود هستند و دروازه وظیفه کنترل جریان اطلاعات را دارند. هر بلوک حافظه شامل یک دروازه ورودی برای کنترل جریان فعال‌سازی ورودی به سلول حافظه و یک دروازه خروجی برای کنترل جریان خروجی فعال‌سازی در بقیه شبکه و یک دروازه فراموشی است(14). در شبکه LSTM لایه‌ای پنهان مانند یک واحد حافظه عمل می‌کند که این شبکه را قادر می‌سازد که همبستگی‌های درون سری زمانی را در هم به صورت کوتاه‌مدت و هم بلندمدت مدل‌سازی نماید (تصویر شماره 3)(16).

یک واحد از شبکه LSTM به صورت زیر نمایش داده می‌شود:

سلول حافظه، در مرکز واحد است که در شکل 3، توسط دایره قرمز مشخص شده است. i_t و f_t و o_t خروجی‌های دروازه‌های مختلف، S_t وضعیت جدید واحد حافظه، S_t وضعیت نهایی واحد حافظه و O_t خروجی نهایی واحد حافظه هستند. $W^{(i)}$، $W^{(f)}$، $W^{(o)}$، $W^{(c)}$، $U^{(i)}$، $U^{(f)}$، $U^{(o)}$، $U^{(c)}$ ماتریس ضرایب هستند که در تصویر شماره 3 مشخص شده‌اند. ازطریق تعیین تابع‌های مختلف برای دروازه‌ها، واحد حافظه LSTM می‌تواند همبستگی‌های پیچیده ویژگی‌های(Feature) را در سری‌های زمانی، هم به‌صورت کوتاه‌مدت و هم بلندمدت ثبت کند که در مقایسه با RNN، یک پیشرفت چشمگیر است (تصویر شماره 4)(16).



تصویر شماره 3. طرح واحد حافظه LSTM(16)



تصویر شماره 4. مقایسه یک شبکه RNN عادی با LSTM(14)

دو روش باکس - جنکینز و شبکه عصبی به کار رفته در این پژوهش به وسیله نرم‌افزارهای Excel، Minitab و Python بررسی شده است. ازآنجایی‌که واحد زمانی به‌صورت ماهانه در نظر گرفته شده است و تعداد 132 مشاهده برای 11 سال درنظر گرفته شده‌است، هر دو مدل از این جهت مناسب استفاده بر روی داده‌ها بودند. برای مقایسه اینکه دقت کدام مدل در برآورد روند داده‌ها بهتر بوده است، برای هر مدل میانگین

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

مربعات خطا (MSE) را محاسبه نموده و نتایج حاصل نیز نشان داده که قدرت پیش‌بینی مدل LSTM از مدل SARIMA به مقدار قابل توجهی بهتر است. هم‌چنین این مدل برای پیش‌بینی آینده مورد استفاده قرار می‌گیرد در این پژوهش برای وضعیت سه سال آینده از این مدل استفاده شد. جهت پیاده‌سازی مدل LSTM چند نکته دارای اهمیت است که به ترتیب عبارت‌اند از: تمیز کردن داده‌ها، تعریف درست داده‌ها در نرم‌افزار، استانداردسازی متغیرهای پاسخ، برازش مدل و در نهایت پیش‌بینی آینده با استفاده از مدل برازش شده؛ که این مدل و پیش‌بینی صورت گرفته با استفاده از Visual Studio Code و کتابخانه‌های karas و TensorFlow انجام شده است. برای سنجش قابلیت پیش‌بینی مدل، داده‌ها به سه بخش 70، 10 و 20 درصد (یادگیری، اعتبارسنجی و آزمون) تقسیم شده و در پایان برآورد صورت گرفته برای 26 ماه پایانی داده‌ها (1398 تا 1400) در کنار مقادیر مشاهده شده برای مرگ (Fatal) و مصدومیت (Injury) ناشی از حوادث ترافیکی در نمودار شماره 4 به نمایش درآمده‌اند. معرفی ماتریس برای انجام این آزمون بدین‌صورت است که نخست مشاهده اول تا دهم به عنوان مقادیر متغیر پیشگو و مشاهده یازدهم به عنوان متغیر پاسخ در یک سطر قرار می‌گیرند و در سطر بعدی مشاهده دوم تا یازدهم به عنوان مقدار متغیر پیشگو و مشاهده یازدهم به عنوان متغیر پیشگو و تا آخرین مشاهده به همین طریق معرفی می‌شوند. برای انجام آزمون از تابع زیان میانگین مربعات خطا و بهینه‌ساز «Adam» استفاده شده است. هم‌چنین تعداد نرون‌های LSTM0، نرون‌های Dense120 و تعداد تکرار 120 درنظر گرفته شد.

یافته‌ها

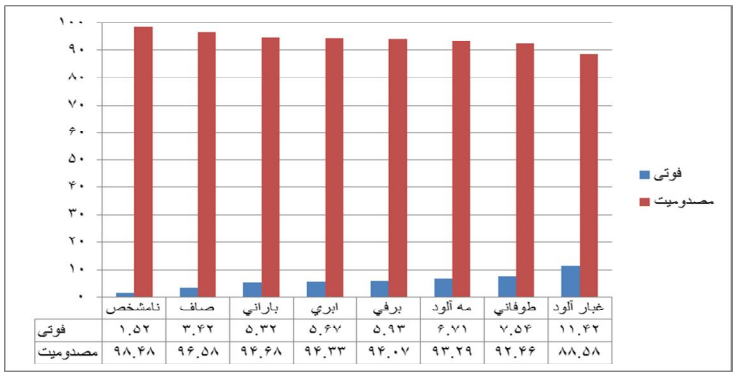
طی سال‌های 1390 تا 1400 بیش‌ترین تعداد فوتی در سال 1390 با 20068 نفر و کم‌ترین تعداد در سال 1399 با 15396 نفر اتفاق افتاده است. هم‌چنین بیش‌ترین تعداد مجروحین در سال 1397 با 367451 مصدوم و کم‌ترین مجروح مربوط به سال 1399 با 276771 مصدوم است. 21/38 درصد از کل فوتی‌ها و 27/56 درصد مصدومان را زنان تشکیل داده‌اند. به ترتیب، برخورد دو وسیله نقلیه، برخورد وسیله نقلیه با موتورسیکلت، برخورد با عابر پیاده، واژگونی، برخورد با جسم ثابت و تصادفات زنجیره‌ای انواع حوادث ثبت شده بودند. به ترتیب بازه سنی 20 تا 29 سال به ترتیب با 23/79 درصد و 0/81 درصد 30 درصد از کل فوتی‌ها و مصدومان بیش‌ترین سهم را در حوادث ترافیکی و بازه سنی زیر 10 سال با 8/24 درصد و 5/11 درصد کم‌ترین تعداد فوتی‌ها و مصدومان را داشتند. نتایج حاصل نشان می‌دهد جوانان با داشتن سهم بیش‌تر از مصدومیت‌ها، سهم کم‌تری از فوتی‌ها داشته و برعکس کودکان و افراد سالخورده سهمشان در فوتی‌ها بیش‌تر از مصدومان است. عمده افراد فوت‌شده و مصدوم به ترتیب دیپلمه، بی‌سواد و کم‌سواد بودند. با در نظر گرفتن پنجشنبه به عنوان روز تعطیل، ماه‌های با نه روز تعطیلی که بیش‌تر در نیمه دوم سال‌های مورد مطالعه قرار می‌گیرند؛ به شکل معنی‌داری دارای فوتی کم‌تر و مصدومان بیش‌تر بوده‌اند. درحالی‌که 90/83 درصد از کل حوادث ترافیکی در هوای صاف اتفاق افتاده است اما طبق آنچه در نمودار شماره 1 آمده، هرچه شرایط جوی بدتر شود نسبت تعداد فوتی به مصدوم افزایش می‌یابد. در مقایسه نمودارهای مرگ و مصدومیت ناشی از حوادث ترافیکی، یک روند تقریباً یکنواخت نزولی در مرگ‌ومیر مشاهده می‌شود (نمودار شماره 2 متغیر Fatal سمت چپ) درحالی‌که مصدومیت ناشی از حوادث رانندگی تا پیش از اواخر سال 1398 حالت صعودی و پس از آن (مصادف با همه‌گیری کووید-19) (Covid-19) به شکل نزولی درآمده است (نمودار شماره 2 متغیر Injury سمت راست)؛ هم‌چنین واریانس تعداد مصدومان در ماه‌های مختلف سال نیز به نسبت سال‌های قبل افزایش یافته است.

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

در مجموع سال‌های مورد مطالعه بیش‌ترین فوتی و مصدوم به ترتیب به تعداد 20285 و 359107 نفر مربوط به شهریورماه و کم‌ترین فوتی با 11683 نفر در بهمن‌ماه بوده ولی کم‌ترین تعداد مصدوم مربوط به اسفندماه با 213996 نفر است. طبق آنچه در جدول شماره 1 آورده شده است، در این مدت متوسط کشوری شاخص نرخ بروز فوت ناشی از حوادث ترافیکی 22 نفر و مصدومان 402 در هر یک هزار نفر از جمعیت کشور است. (نرخ سالانه فوت طی سال 1400 برای کل کشور حدود 20 نفر و نرخ مصدومیت 377 نفر به ازای هر 100 هزار نفر بوده که نشان از کاهش این نرخ‌ها به نسبت میانگین زمان مورد مطالعه است).

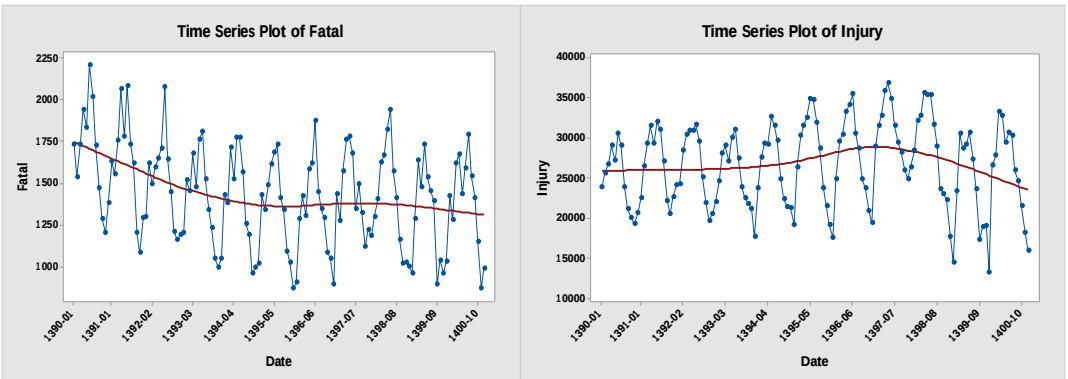
جهت برازش مدل باکس-جنکینز با دقت در نمودار شماره 3 به وضوح وجود تغییرات فصلی در داده‌ها مشخص است. در نمودار خودهمبستگی نیز این فرض را تأیید می‌کند.

درپایان و پس از مقایسه مدل‌های ایستای ممکن، بهترین مدل برای تعداد فوتی‌ها SARIMA(0,1,1)(0,1,1) با AIC=8/015 و برای مصدومان SARIMA(1,0,0)(0,1,4) با AIC=17/547 است که برآورد پارامترهای آن‌ها در جداول شماره 2 و 3 آورده شده‌اند.



	نامشخص	صاف	بارانی	ابری	برفی	مه آلود	طوفانی	غبار آلود
فوتی	۱.۵۲	۳.۴۲	۵.۳۲	۵.۶۷	۵.۹۳	۶.۷۱	۷.۵۴	۱۱.۴۲
مصدومیت	۹۸.۴۸	۹۶.۵۸	۹۴.۶۸	۹۴.۳۳	۹۴.۰۷	۹۳.۲۹	۹۲.۴۶	۸۸.۵۸

نمودار شماره 1: درصد تصادفات منجر به فوت و مصدومیت در شرایط آب و هوایی مختلف

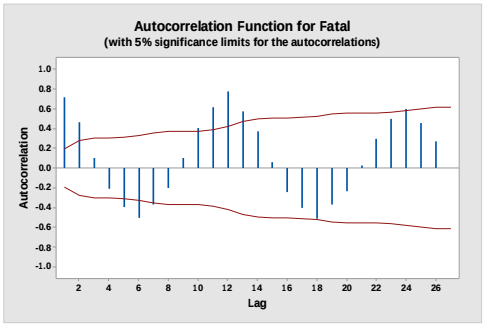


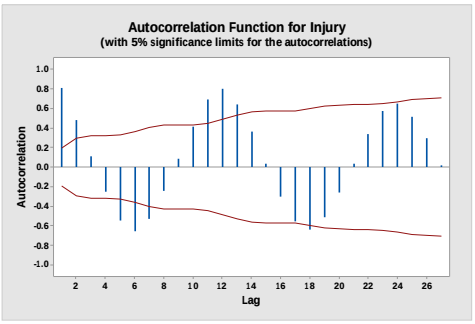
نمودار شماره 2: روند تعداد فوتی‌ها و مصدومان حوادث ترافیکی در ایران طی سال‌های 1390-1400

جدول شماره 1: نرخ بروز فوت ناشی از حوادث ترافیکی در ایران طی سال‌های 1390 تا 1400

سال	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400
نرخ متوسط فوت	27	25	23	22	21	20	20	21	21	18	20
نرخ متوسط مصدومیت	396	419	410	391	396	417	416	450	421	333	377

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]





نمودار شماره 3: خودهمبستگی تعداد مرگ و مصدومان ناشی از حوادث ترافیکی در ایران طی سال‌های 1390-1398

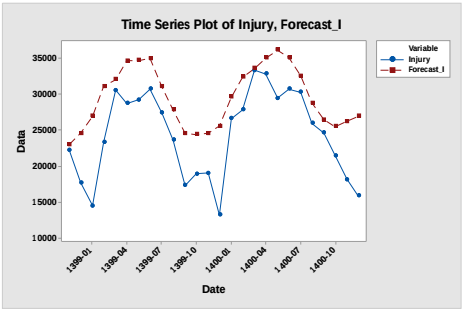
جدول شماره 2- برآورد مدل SARIMA(0,1,1)(0,1,1) برای مرگ حوادث ترافیکی در ایران طی سال‌های 1390-1398

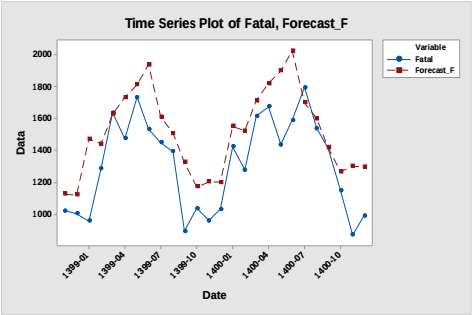
نوع	مقدار	SE	T-Value	سطح معنی داری
MA 1	1/012	0/002	478/960	0/000
SMA 12	0/805	0/086	9/390	0/000
ثابت	1/815	0/088	20/540	0/000

جدول شماره 3: برآورد پارامترهای مدل SARIMA(1,0,0)(0,1,4) برای مصدومان حوادث ترافیکی در ایران طی سال‌های 1390-1398

نوع	مقدار	SE	T-Value	سطح معنی
AR 1	0/609	0/085	7/190	0/000
SMA 12	0/270	0/135	2	0/049
SMA 24	0/179	0/160	1/140	0/258
SMA 36	-0/235	0/170	-1/380	0/170
SMA 48	0/557	0/178	3/12	0/002
ثابت	217/380	54/330	4	0/000

برای مرحله آزمون با استفاده از روش باکس-جنکینز همان‌گونه که در نمودار شماره 4 آمده است، مقادیر برازش شده (خط قرمز) در هر مدل، به‌جز در چند مورد، کم‌تر از مقادیر اصلی مشاهده شده است که می‌تواند ناشی از اعمال محدودیت‌های ترافیکی و کاهش ترددهای ناشی از محدودیت‌های کرونایی بوده باشد(17).





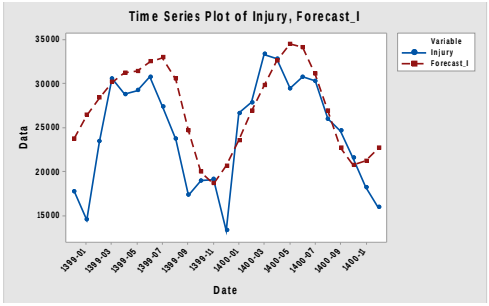
نمودار شماره 4: مقادیر پیش‌بینی‌شده سال‌های 1398 تا 1400 برای تعداد فوتی‌ها و مصدومان ناشی از تصادفات به روش باکس-جنکینز

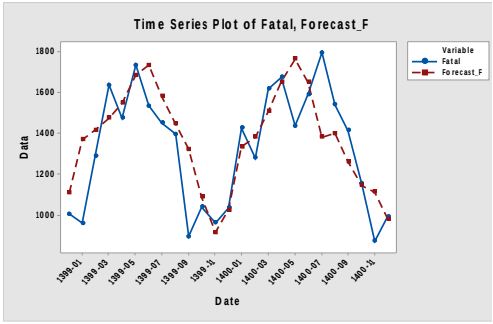
نمودار شماره 5 مربوط به برآورد سال‌های 1398 تا 1400 توسط روش LSTM است. R^2 توابع رگرسیونی میان مقادیر پیش‌بینی‌شده و مشاهده‌شده برای فوتی‌ها در مرحله یادگیری 83/54، در مرحله اعتبارسنجی 81/62 و در مرحله آزمون 55/6 و برای مصدومان در مرحله یادگیری 89/29، در مرحله اعتبارسنجی 86/86 و در مرحله آزمون 41/57 بوده است که مناسب ارزیابی می‌شود.

درپایان MSE هر چهار آزمون در جدول شماره 4 آورده شده است؛ بنابر جدول مشخص است که MSE مدل LSTM از مدل SARIMA کم‌تر است.

با استفاده از مدل LSTM برای سه سال بعدی در نمودار شماره 6 پیش‌بینی صورت گرفت. خط قرمز که مقادیر پیش‌بینی‌شده هستند؛ یک روند رو به افزایش در تعداد فوتی‌ها و مصدومان را نشان می‌دهد.

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]





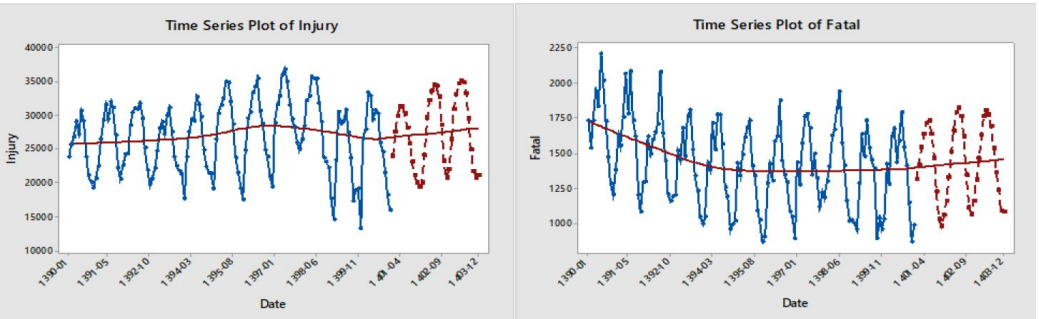
نمودار شماره 5: مقادیر پیش‌بینی‌شده سال‌های 1398 تا 1400 برای تعداد فوتی‌ها و مصدومان ناشی از تصادفات به روش یادگیری عمیق

جدول شماره 4: مقایسه میانگین مربعات خطا روش باکس-جنکینز و شبکه عصبی در برآورد داده‌ها

آزمون	MSE مرحله آزمون مدل فوتی‌های حوادث ترافیکی	MSE مرحله آزمون مدل مصدومان حوادث ترافیکی
SARIMA	130118/300	36865957
LSTM	35292/930	20804913/590

بحث

طبق آن‌چه که WHO بر مبنای آخرین مطالعه جهانی حوادث ترافیکی 2018 در تارنامه خود بارگذاری کرده است، در کل جهان هر 23 ثانیه یک نفر بر اثر حوادث ترافیکی جان خود را از دست می‌دهد. هم‌چنین این سازمان مرگ‌ومیر ناشی از حوادث ترافیکی را در کل دنیا را افزایشی برآورد نموده است. بر این اساس اروپا و غرب اقیانوس آرام دارای روند نزولی، آمریکا و خاورمیانه با شیب بسیار کمی روند افزایشی و آفریقا و جنوب غربی آسیا روند افزایشی شدیدی را دارند(18) طبق نمودار شماره 6 و نتایج حاصل از پیش‌بینی در این مطالعه، وضعیت تعداد فوتی‌ها و مصدومان در ایران نیز به صورت افزایشی پیش‌بینی می‌شود. هم‌چنین بر مبنای گزارش WHO در بازه زمانی مورد مطالعه، کم‌ترین نرخ فوتی به ترتیب مربوط به اروپا، آمریکا، غرب اقیانوس آرام، خاورمیانه، جنوب شرقی آسیا و آفریقا بوده است. در مطالعه حاضر نیز علاوه بر مشخص بودن روند نزولی فوتی‌ها، این نرخ برای کل کشور در حدود 21 نفر به ازای هر 100/00 نفر جمعیت برآورد شده است و نشان می‌دهد که در ایران علاوه بر اینکه نرخ مرگ‌ومیر به نسبت کل دنیا بالاتر است با در نظر گرفتن جمعیت کشور نیز نسبت به کشورهای خاورمیانه نرخ بالاتری در مرگ‌ومیر ناشی از حوادث ترافیکی مشاهده می‌شود. همانند این مقاله، در مطالعه بهادری منفرد و همکاران، مدل SARIMA(0,1,1)(0,1,1) پیش‌بینی مرگ‌ومیر ناشی از حوادث ترافیکی در ایران در سال‌های 2004 تا 2011 برازش شده است(4) که این امر نشان می‌دهد روند کلی مرگ‌ومیر حوادث ترافیکی به تشخیص این مدل ثابت مانده است.



نمودار شماره 6: پیش‌بینی روند تعداد فوتی‌ها و مصدومان حوادث ترافیکی در ایران طی سال‌های 1401 تا 1403

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

البته در این خصوص مطالعه دیگری که توسط نصیری و همکاران و برای مدل‌بندی فوتی‌های حوادث ترافیکی از سال‌های 2016 تا 2020 صورت گرفته، مدل پیشنهادی SARIMA(1,0,0)(1,0,0) بوده است(5).

در مقام مقایسه، در مطالعات گوناگونی برتری مدل LSTM در پیش‌بینی سری‌های زمانی مختلف بر مدل SARIMA (باکس-جنکینز) به اثبات رسیده است و این مدل می‌تواند پیش‌بینی دقیق‌تری ارائه دهد(6-8)؛ امری که در این مطالعه نیز تأیید می‌شود.

در این مطالعه از آن‌جایی‌که مدل SARIMA برای فوتی‌ها تنها شامل مؤلفه‌های تفاضل و میانگین متحرک است، یک روند ثابت خطی (در اینجا روند کلی رو به کاهش) را مدل می‌کند؛ حال آنکه از نظر مدل LSTM در سال‌های آتی تعداد مرگ در حوادث ترافیکی افزایش نسبی خواهد داشت. برآوردهای مدل SARIMA تقریباً همواره بیش از مقادیر اصلی بوده است درحالی‌که مدل LSTM برآوردهایی تقریباً در اطراف مقادیر اصلی داشت بنابراین مدل LSTM از نظر تشخیص نوسانات بهتر عمل کرده است (مدل LSTM توانسته است خود را با کاهش تعداد مرگ‌ومیر ناشی از کاهش تعداد تردد وفق دهد).

محدودیت‌های مطالعه

با وجود نقص و داده‌های گمشده در داده‌های ثبتی اولیه، امکان تحلیل‌های جامع‌تر یا تحلیل روند در بازه زمانی کوچکتر میسر نبوده و اطلاعات موجود نیز برای بررسی ابعاد دیگر حوادث ترافیکی ناکافی بودند.

References

1. De Silva V, Tharindra H, Vissoci JRN, Andrade L, Mallawaarachchi BC, Østbye T, et al. Road traffic crashes and built environment analysis of crash hotspots based on local police data in Galle, Sri Lanka. Int J Inj Contr Saf Promot 2018; 25(3): 311-8.
2. WHO. Global status report on road safety: WHO; 2018.
3. WHO. Global status report on road safety: WHO; 2015.
4. Bahadorimonfared A, Soori H, Mehrabi Y, Delpisheh A, Esmaili A, Salehi M, et al. Trends of fatal road traffic injuries in Iran (2004–2011). PloS one 2013; 8(5): e65198.
5. Nassiri H, Mohammadpour SI, Dahaghin M. Forecasting time trend of road traffic crashes in Iran using the macro-scale traffic flow characteristics. Heliyon 2023; 9(3) :14481.
6. ArunKumar K, Kalaga DV, Kumar CMS, Kawaji M, Brenza TM. Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. Alexandria engineering journal 2022; 61(10): 7585-603.
7. Dubey AK, Kumar A, García-Díaz V, Sharma AK, Kanhaiya K. Study and analysis of SARIMA and LSTM in forecasting time series data. Sustainable Energy Technologies and Assessments 2021; 47: 101474.
8. He K, Ji L, Wu CWD, Tso KFG. Using SARIMA–CNN–LSTM approach to forecast daily tourism demand. Journal of Hospitality and Tourism Management 2021; 49: 25-33.
9. Organization ILM. Traffic Accidents: Iranian Legal Medicine Organization; 2022 [cited 2022 9 April]. Annual reports of the dead and injured in traffic accidents]. [Available from: http://lmo.ir/web_directory/53999-.
10. Pankratz A. Forecasting with univariate Box-

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]

Jenkins models: Concepts and cases: John Wiley & Sons 2009.

11. Zhang GP, Qi M. Neural network forecasting for seasonal and trend time series. European journal of operational research 2005; 160(2): 501-14.
12. Fausett L. Fundamentals of neural networks: architectures, algorithms, and applications: Prentice-Hall 1994.
13. Nielsen MA. Neural networks and deep learning. San Francisco, CA, USA: Determination press 2015
14. Azzouni A, Pujolle G. A long short-term memory recurrent neural network framework for network traffic matrix prediction. arXiv preprint arXiv 2017.
15. Graves A, Mohamed A-r, Hinton G, editors. Speech recognition with deep recurrent neural networks. 2013 IEEE international conference on acoustics, speech and signal processing 2013: 6645-6649.
16. Zhao Z, Chen W, Wu X, Chen PC, Liu J. LSTM network: a deep learning approach for short-term traffic forecast. IET Intelligent Transport Systems 2017; 11(2): 68-75.
17. Ministry of Health TaMEoI. New restrictions on intercity traffic: Ministry of Health and Medical Education; 2020 24 March [cited 2022 18 April]. The decision of the national headquarters of Corona about traffic restrictions. Available from: https://behdasht.gov.ir/.
18. WHO. Death on the rouds (Data visualization).

[Downloaded from jmums.mazums.ac.ir on 2026-11-09]