Amirkabir Journal of Civil Engineering

Amirkabir J. Civil Eng., 54(4) (2022) 297-300
DOI: 10.22060/ceej.2021.19650.7219

Traffic state prediction with machine learning algorithms for short-term and mid-term prediction

A. Rasaizadi, S. H. Seyedabrishami*

School of Civil and Environmental Engineering, Tarbiat Modares University, Tehran, Iran

ABSTRACT: Predicting traffic variables and informing the passengers and the transportation network operators is one way to manage the travel demand. By informing the future state of traffic through intelligent transportation systems, there is more readiness to avoid congestion. In this study, three machine learning algorithms, including support vector machine (SVM), random forest (RF), and long short-term memory (LSTM), were used to predict the hourly traffic state, consist of light, semi-heavy and heavy states, for Karaj to Chaloos rural road in the north of Iran. Predictor variables of mid-term models are calendar information, weather, and road blockage policies. In contrast, in short-term models, in addition to the mentioned variables, the observed traffic states in the past three to eight hours have been used, and these models can only predict the future of one and two hours. The results show that short-term LSTM is the most accurate traffic state predictive model, with an accuracy equal to 90.11%. Among the mid-term models, the LSTM model has predicted traffic state more accurately than SVM and RF, and its accuracy is equal to 82.07%. Also, LSTM has the highest values of f1 measure to predict light, semi-heavy, and heavy, which are equivalent to 0.86, 0.93, and 0.81, respectively. Also, the hour, holiday, and type of holiday variables and traffic state observed in 3 to 8 hours later variables have the greatest effect on increasing the accuracy of mid-term and short-term models, respectively.

Review History:

Received: Feb. 16, 2021
Revised: Apr. 30, 2021
Accepted: Jun. 03, 2021
Available Online: Jun. 13, 2021

Keywords:

Traffic state prediction
Support vector machine
Random forest
Long short-term memory
Intelligent transportation systems

1- Introduction

Intelligent transportation systems (ITS) are one of the main components of traffic management systems. The result of the use of these systems is to achieve or maintain the balance between travel supply and demand with high efficiency and low cost. The advanced traveler information systems (ATIS) is one of the most important of subsets of ITS. In these systems, all the practical and available data of the transportation network will be given to the passenger to schedule their travels. This information can be used for the existing state of the transportation network but the effectiveness of this information can be more if the future of the transportation network is predicted and informed [1].

In this study, hourly traffic states including, light, semi-heavy and heavy, are predicted using three methods based on machine learning algorithms, including support vector machine (SVM), random forest (RF), and long short-term memory (LSTM). To predict traffic state, spatial-temporal influential predictor variables are extracted and added to the data set. Models have been divided into two groups: medium-term and short-term models. In the short-term model, in addition to the spatial-temporal variables, the traffic state observed in the past three to eight hours has also been used and these models only can predict traffic hours for one and two next hours. In the medium-term models, the previous observations of the traffic situation have not been used. Finally, the accuracy of these two sets of models is evaluated and compared. The traffic data of Karaj to Chaloos, a rural road in the north of Iran, is used to predict the traffic state.

2- Data and methodology

This used traffic data is collected for one section of the Karaj-Chaloos rural road in the north of Iran by loop detectors. Data collection has been carried out from 21 March 2017 to 11 September 2019. Data is divided into three sections; observations of the first 22 months are used to train models, the next two months for validation, and the last six months for testing the predictions. Table 1 describes the candidate features to predict the traffic state in the dataset.

Used machine learning algorithms to predict traffic state are support vector machine (SVM), random forest (RF), and long short-term memory (LSTM).

In SVM, the support vectors are a set of points that characterize the boundaries of classes in an n-dimensional space. The SVM determines class boundaries which leads to the best classification and separation of data by embedding

*Corresponding author's email: seyedabrishami@modares.ac.ir

Copyrights for this article are retained by the author(s) with publishing rights granted to Amirkabir University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit https://www.creativecommons.org/licenses/by-nc/4.0/legalcode.

Table 1. Description of candidate features

Feature Name	Description
Season	Including spring, summer, fall, and winter
Solar month	Including 12 solar months
Lunar month	Including 12 lunar months
Day of a solar month	Including 29-31 days of a solar month
Day of a lunar month	Including 29-30 days of a lunar month
Time of day	Including 24 hours a day
6 hours before the holidays	Equal to 1 if it is 1 to 6 hours before holidays
6 hours after the holidays	Equal to 1 if it is 1 to 6 hours after holidays
Day or night	Including day and night
Number of holidays	The number of sequential holidays
Holidays	Includes 1 for holidays and 0 for other days
Holiday type	Type of holidays
Holiday in three days later	Equal to 1 if three days later is a holiday
Type of holidays in three days later	Including the holiday type of three days later if it is a holiday, otherwise equals 0.
Holiday in three days ago	Equal to 1 if three days ago is a holiday
Type of holidays in three days ago	Including the holiday type of three days ago if it is a holiday, otherwise equals 0.
Holiday in two days later	Equal to 1 if two days later is a holiday
Type of holidays in two days later	Including the holiday type of two days later if it is a holiday, otherwise equals 0.
Holiday in two days ago	Equal to 1 if two days ago is a holiday
Type of holidays in two days ago	Including the holiday type of two days ago if it is a holiday, otherwise equals 0.
Holiday in a day later	Equal to 1 if a day later is a holiday
Type of holidays in a day later	Including the holiday type of a day later if it is a holiday, otherwise equals 0.
Holiday in a day ago	Equal to 1 if a day ago is a holiday
Type of holidays in a day ago	Including the holiday type of a day ago if it is a holiday, otherwise equals 0.
Weather condition	Including sunny, rainy, and snowy

support vectors [2]. The RF consists of a large number of decision trees. In this model, the training data is divided between decision tree models. Predictions are made for each decision tree. The average of predictions is determined as the RF's final prediction [3]. The LSTM networks are a special type of recurrent neural network (RNN) capable of learning long-term dependencies. Traditional RNNs are not able to train the time series with long time lags. LSTM can address this issue by incorporating memory units and learning when to forget previous memories and update memories [4].

3- Results and discussion

Table 2 shows the accuracy of models for train and test datasets.

For both short-term and medium-term models, the LSTM has the highest accuracy. Also, short-term models are more accurate compared to medium-term models. Tables 3 and 4 show the performance of models to predict each traffic state in short-term and medium-term models, respectively.

Table 2. Accuracy of models for training and test datasets

Model	Accuracy of short-term models (%)		Accuracy of medium-term models (%)	
	Train	Test	Train	Test
SVM	90.82	87.74	83.82	76.63
FR	92.45	87.51	84.13	79.04
LSTM	96.88	90.11	88.69	82.07

Table 3. Precision, recall and F1 in short-term models

Model	State	Precision	Recall	F_1
SVM	Light	0.81	0.76	0.78
	Semi-heavy	0.88	0.92	0.9
	Heavy	0.74	0.67	0.7
RF	Light	0.84	0.83	0.83
	Semi-heavy	0.91	0.92	0.91
	Heavy	0.79	0.77	0.78
LSTM	Light	0.86	0.86	0.86
	Semi-heavy	0.93	0.94	0.93
	Heavy	0.84	0.79	0.81

Table 4. Precision, recall and F1 in medium-term models

Model	State	Precision	Recall	F_1
	Light	0.68	0.69	0.69
SVM	Semi-heavy	0.83	0.85	0.84
	Heavy	0.63	0.52	0.57
	Light	0.7	0.76	0.73
RF	Semi-heavy	0.86	0.84	0.85
	Heavy	0.66	0.62	0.64
LSTM	Light	0.75	0.8	0.77

Tables 3 and 4 shows that in term of F_1, LSTM predicts all of the traffic states more accurate then SVM and RF. Other findings are as follows:The use of the hour and holiday-related variables have the greatest effect on increasing the accuracy of the medium-term models. Using observations from three to eight hours ago has the greatest effect on increasing the accuracy of the short-term model. The effective factors on traffic patterns are constant and are not a function of the used method. There is not a variable that removing it causes an increase in the accuracy of the predictions.

4- Conclusion

Traffic state is a qualitative traffic parameter that shows the performance of the road and is more understandable for travelers. After informing predicted traffic state to travelers and transportation agencies through ATIS more sustainable transportation system can be expected. This paper aims to predict the traffic state by using three machine learning techniques, SVM, RF, and LSTM. Traffic data of one section of Karaj to Chaloos is used to train and test models. Results show that LSTM outperforms SVM and RF, in both short-term and medium-term prediction time horizons. In general, short-term models are more accurate compared to medium-term models but the prediction time horizon of short-term models is limited to one and two next hours.

References

[1]Polydoropoulou, A., and Ben-Akiva, M., 2018. "The effect of advanced traveller information systems (ATIS) on travellers' behaviour", pp. 317-352, Routledge.

[2]Yang, J., Deng, J., Li, S., and Hao, Y., 2017. "Improved traffic detection with support vector machine based on restricted Boltzmann machine", Soft Computing, 21(11), pp. 3101-3112.

[3]Probst, P., Wright, M. N., and Boulesteix, A. L., 2019. "Hyperparameters and tuning strategies for random forest", Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), 1301.

[4]Sherstinsky, A., 2020. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network", Physica D: Nonlinear Phenomena, 404, 132306.

HOW TO CITE THIS ARTICLE
A. Rasaizadi, S. H. Seyedabrishami, Traffic state prediction with machine learning algorithms for short-term and mid-term prediction, Amirkabir J. Civil Eng., 54(4) (2022) 297-300.

DOI: 10.22060/ceej.2021.19650.7219

This page intentionally left blank

نشریه مهندسی عمران امیرکبیر

نشریه مهندسی عمران امیرکبیر، دوره ۵۴، شماره ۴، سال ۱۴۰۱، صفحات ۱۵۰۳ تا ۱۵۲۰
DOI: 10.22060/ceej.2021.19650.7219

پیش‌بینی وضعیت ترافیک با الگوریتم‌های یادگیری ماشین برای افق‌های زمانی کوتاه مدت و میان مدت

آرش رساءایزدی، سیداحسان سیدابریشمی*

دانشکده مهندسی عمران و محیط زیست، دانشگاه تربیت مدرس، تهران، ایران.

خلاصه: پیش‌بینی متغیرهای ترافیکی و اطلاع‌رسانی آن به مسافرین و گردانندگان شبکه حمل‌ونقل یکی از راهکارهای مدیریت تقاضای سفر است. با اطلاع‌رسانی وضعیت آینده ترافیک از طریق سیستم‌های حمل‌ونقل هوشمند، آمادگی بیشتری جهت اجتناب از وقوع تراکم ترافیک به وجود می‌آید. در این مطالعه به منظور پیش‌بینی وضعیت ترافیک ساعتی، شامل سه وضعیت روان، نیمه‌سنگین و سنگین، در جاده برون‌شهری کرج به چالوس در شمال ایران، سه مدل یادگیری ماشین، شامل ماشین بردار پشتیبان، جنگل تصادفی و حافظه طولانی کوتاه مدت به دو صورت کوتاه مدت و میان مدت آموزش داده شده‌اند. متغیرهای پیش‌بینی کننده در مدل‌های میان مدت اطلاعات تقویمی، آب و هوا و محدودیت‌های ترافیکی هستند در صورتی که در مدل‌های کوتاه مدت علاوه بر متغیرهای نام برده، وضعیت ترافیک مشاهده شده در سه تا هشت ساعت گذشته نیز استفاده شده است و این مدل‌ها تنها قادر به پیش‌بینی وضعیت ترافیک یک و دو ساعت آینده هستند. نتایج نشان می‌دهد مدل حافظه طولانی کوتاه مدت با دقتی معادل با ۹۰/۱۱ درصد دقیق‌ترین مدل پیش‌بینی کننده وضعیت ترافیک با افق کوتاه مدت است. این مدل برای افق بلند مدت نیز متغیر وضعیت ترافیک را با ۸۲/۰۷ درصد دقت، دقیق‌تر از دو مدل دیگر پیش‌بینی کرده است و بیشترین مقادیر شاخص F (F1) برای پیش‌بینی سه وضعیت ترافیک سبک، نیمه‌سنگین و سنگین را به همراه داشته که به ترتیب برابر با ۰/۸۶، ۰/۹۳ و ۰/۸۱ به دست آمده‌اند. همچنین متغیرهای ساعت و تعطیلی همان روز و نوع تعطیلی و متغیرهای مشاهدات سه تا هشت ساعت پیش وضعیت ترافیک به ترتیب بیشترین تأثیر را بر افزایش دقت مدل های میان مدت و کوتاه مدت دارند.

تاریخچه داوری:
دریافت: ۱۳۹۹/۱۱/۲۸
بازنگری: ۱۴۰۰/۰۲/۱۱
پذیرش: ۱۴۰۰/۰۳/۱۳
ارائه آنلاین: ۱۴۰۰/۰۳/۲۳

کلمات کلیدی:
پیش‌بینی وضعیت ترافیک
ماشین بردار پشتیبان
جنگل تصادفی
حافظه طولانی کوتاه‌مدت
سیستم‌های حمل‌ونقل هوشمند

۱- مقدمه

سیستم‌های حمل‌ونقل هوشمند یکی از ارکان اصلی مدیریت ترافیک شبکه حمل‌ونقل هستند. نتیجه بهره‌گیری از این سیستم‌ها، رسیدن یا حفظ تعادل میان عرضه و تقاضای حمل‌ونقل با هزینه کم و کارایی زیاد است. سیستم‌های حمل‌ونقل هوشمند زیرشاخه‌های متنوعی را شامل می‌شود که یکی از مهم‌ترین این زیرشاخه‌ها سیستم‌های هوشمند اطلاع‌رسانی به مسافر است. در این سیستم تمام اطلاعات کاربردی و در دسترس شبکه حمل‌ونقل به مسافر داده می‌شود تا بتواند با آگاهی بیشتری سفرهای خود را برنامه‌ریزی کند. این اطلاعات می‌تواند برای وضعیت فعلی شبکه اطلاع‌رسانی شود ولی زمانی اثربخشی این اطلاعات بیشتر می‌شود که برای آینده شبکه حمل‌ونقل پیش‌بینی شده و در اختیار مسافر قرار بگیرد [۱]. در چنین شرایطی فرد مسافر قبل از شروع سفر خود آمادگی بیشتری برای انتخاب مسیر و زمان عزیمت و حتی انتخاب انجام سفر یا لغو آن دارد. عموماً متغیرهای ترافیکی، شامل حجم ترافیک [۲]، سرعت متوسط [۳] و وضعیت ترافیک [۴]، پیش‌بینی شده و از طریق سیستم‌های هوشمند اطلاع‌رسانی می‌شود.

پیش‌بینی متغیرهای ترافیکی از طریق تحلیل مشاهدات پیشین متغیرهای ترافیک و کشف روابط اثرگذار بر آن انجام می‌شود. بدین منظور، استفاده از مدل‌های سری زمانی به عنوان ابزاری مبتنی بر آمار و احتمالات، قدمت بیشتری در ادبیات پیشین داشته است [۵]. در مدل‌های سری زمانی تغییرات هر متغیر ترافیکی تابعی از تغییرات مقادیر مشاهده شده‌ی پیشین همان متغیر، تغییرات متغیرهای مستقل اثرگذار بر آن متغیر ترافیکی و بخش خطا دانسته می‌شود. به عنوان مثال، کومار و وناجکشی [۶] حجم ترافیک را با استفاده از مدل میانگین متحرک خود همبسته یکپارچه فصلی[1] (ساریما) پیش‌بینی کرده‌اند که نتایج به دست آمده نشان دهنده دقت بیشتر مدل

* نویسنده عهده‌دار مکاتبات: seyedabrishami@modares.ac.ir

1 Seasonal autoregressive integrated moving average (SARIMA)

حقوق مؤلفین به نویسندگان و حقوق ناشر به انتشارات دانشگاه امیرکبیر داده شده است. این مقاله تحت لیسانس آفرینندگی مردمی (Creative Commons License) در دسترس شما قرار گرفته است. برای جزئیات این لیسانس، از آدرس https://www.creativecommons.org/licenses/by-nc/4.0/legalcode دیدن فرمائید.

ساریما به نسبت مدل‌های میانگین تاریخی و نایو[1] است. در مطالعه یان و ژو [۷] نیز پیش‌بینی جریان مسافر مترو با استفاده از مدل میانگین متحرک خودهمبسته یکپارچه[2] (آریما) انجام شده است. مدل‌های سری زمانی تنها قادر به در نظرگیری روابط خطی میان متغیرهای مستقل و وابسته هستند. از سوی دیگر با افزایش حجم داده عملکرد این مدل‌ها تحت تأثیر قرار گرفته و با ویژگی‌های کلان داده، شامل حجم، سرعت و تنوع زیاد، سازگاری ندارند [۸].

رویکرد دیگر در مسئله پیش‌بینی، رویکرد یادگیری ماشین است. مدل‌های یادگیری ماشین مدل‌هایی سازگار با ویژگی‌های کلان داده بوده که هرگونه روابطی از جمله خطی و غیرخطی را به تصویر می‌کشند. عدم تفسیرپذیری و کشف روابط علت و معلولی نقطه ضعف اصلی مدل‌های یادگیری ماشین است و از این حیث مدل‌های سری زمانی بر مدل‌های یادگیری ماشین برتری دارند [۹]. مطالعاتی که در ادامه مرور می‌شوند از رویکرد یادگیری ماشین برای پیش‌بینی متغیرهای ترافیکی استفاده کرده‌اند.

سان و همکاران [۱۰] به منظور پیش‌بینی حجم ترافیک از روش K–نزدیک‌ترین همسایه[3] استفاده کرده‌اند. در این روش برای پیش‌بینی حجم ترافیک داده‌های جدید، K مشاهده با فاصله اقلیدسی کمتر پیدا شده و میانگین حجم ترافیک این K مشاهده به عنوان پیش‌بینی نهایی انتخاب می‌شود. در این مطالعه به منظور بهبود عملکرد مدل، پارامترهای آن به صورت دینامیک تعدیل می‌شود که در نتیجه منجر به افزایش دقت پیش‌بینی به نسبت حالت پایه مدل شده است. در مطالعه رحمان [۱۱]، پیش‌بینی حجم ترافیک علاوه بر مدل K–نزدیک‌ترین همسایه با مدل‌های ماشین بردار پشتیبان[4] و شبکه عصبی مصنوعی نیز انجام شده است که K–نزدیک‌ترین همسایه دقت پیش‌بینی بیشتری به همراه دارد. لی و همکاران [۱۲] سرعت ترافیک را با مدل ماشین بردار پشتیبان بهینه شده با الگوریتم ژنتیک پیش‌بینی کرده‌اند که در مقایسه با مدل پایه دقت پیش‌بینی بیشتری دارد. جنگل تصادفی[5] نیز یکی دیگر از مدل‌های پرکاربرد مبتنی بر یادگیری ماشین است که در مطالعه لیو و وو [۱۳] برای پیش‌بینی تراکم ترافیک و مطالعه‌ی چنگ و همکاران [۱۴] برای پیش‌بینی تصادفات استفاده شده است.

شارما و همکاران [۱۵] به منظور پیش‌بینی حجم ترافیک در بزرگراهی که جریان ترافیک دو جهت آن به صورت فیزیکی از هم جدا نشده‌اند از مدل شبکه عصبی مصنوعی استفاده کرده‌اند. مقایسه نتایج مدل شبکه عصبی مصنوعی با نتایج جنگل تصادفی، ماشین بردار پشتیبان، K–نزدیک‌ترین همسایه و درخت رگرسیون نشان دهنده دقت پیش‌بینی بیشتر مدل شبکه عصبی مصنوعی است. همچنین رازا و ژونگ [۱۶] برای طراحی ساختار مدل شبکه عصبی مصنوعی از الگوریتم ژنتیک استفاده کرده‌اند که سبب افزایش دقت پیش‌بینی شده است.

پژوهش بر انواع ساختار مدل‌های شبکه عصبی منجر به پیدایش مدل‌های مبتنی بر شبکه عصبی با قابلیت‌های منحصر به فرد شده است. از جمله این مدل‌ها می‌توان به مدل شبکه عصبی بازگشتی[6] و حافظه طولانی کوتاه مدت[7] اشاره کرد. مدل شبکه عصبی مصنوعی مشاهدات متوالی را مستقل از هم در نظر می‌گیرد در صورتی که در مدل‌های شبکه عصبی بازگشتی و حافظه طولانی کوتاه مدت، هر مشاهده از مشاهدات پیشین آن تأثیر می‌پذیرد و در نظرگیری این ارتباط به واقعیت و ماهیت سری زمانی داده‌های ترافیکی نزدیک‌تر است. به عنوان مثال، راماکریشنان و سونی [۱۷] از مدل شبکه عصبی بازگشتی به منظور پیش‌بینی حجم ترافیک استفاده کرده‌اند و نشان دادند این مدل به نسبت مدل‌های آماری دقت پیش‌بینی بیشتری دارد. مدل شبکه عصبی بازگشتی تنها ارتباط مشاهدات پیشین در گذشته نزدیک را در نظر می‌گیرند. در مقابل در مدل حافظه طولانی کوتاه مدت این ارتباط علاوه بر گذشته نزدیک، برای مشاهدات مربوط به گذشته دور نیز در نظر گرفته می‌شود. ژائو و همکاران [۱۸] و رانجان و همکاران [۱۹] به ترتیب برای پیش‌بینی حجم ترافیک و تراکم ترافیک از مدل حافظه طولانی کوتاه مدت استفاده کرده‌اند.

وضعیت ترافیک یک متغیر ترافیکی دیگر است که به صورت کیفی شرایط حاکم بر راه‌ها را نشان می‌دهد. عموماً این متغیر به صورت اسمی تعریف می‌شود و تعریف آن بر اساس مقادیر سایر متغیرهای ترافیکی نظیر سرعت و حجم ترافیک انجام می‌شود. این متغیر می‌تواند به صورت رنگ‌بندی درآمده و بر روی نقشه‌های ترافیکی نشان داده شود. مزیت این متغیر به نسبت متغیرهای سرعت و حجم ترافیک قابل فهم بودن برای عموم مردم است. برای ایجاد درک مناسب از شرایط ترافیک با استفاده از متغیرهای سرعت و حجم دانستن مقادیر سرعت آزاد و ظرفیت مسیر ضروری است حال آنکه در تعیین وضعیت ترافیک این موارد دیده شده است و برای کاربران قابل درک‌تر است. موارد اندکی از مطالعات پیشین به پیش‌بینی متغیر وضعیت ترافیک پرداخته‌اند. به عنوان مثال از مطالعاتی که وضعیت ترافیک

1 Naïve
2 (Autoregressive integrated moving average (ARIMA
3 K-nearest neighbours (KNN)
4 Support vector machine (SVM)
5 Random forest (RF)
6 Recurrent neural network (RNN)
7 Long short-term memory (LSTM)

جدول ۱. خلاصه ای از سایر مطالعات پیشین در زمینه ی پیش بینی متغیرهای ترافیکی

Table 1. Summary of previous studies on traffic variables prediction

ردیف	محققین	سال	کشور	متغیر ترافیکی	رویکرد	روش
۱	روتکا [۲۲]	۲۰۰۸	لیتوانی	حجم ترافیک	سری زمانی	آریما
۲	خویی و همکاران [۲۳]	۲۰۱۳	استرالیا	زمان سفر	سری زمانی	ساریما
۳	چیکاکریشنا و همکاران [۲۴]	۲۰۱۹	هند	حجم ترافیک	سری زمانی	ساریما
۴	القمدی و همکاران [۲۵]	۲۰۱۹	ایالات‌متحده آمریکا	تراکم ترافیک	سری زمانی	آریما
۵	ارامو و همکاران [۲۶]	۲۰۲۰	ایتالیا	حجم ترافیک	سری زمانی	ساریما
۶	گوس و همکاران [۲۷]	۲۰۱۶	انگلستان	حجم ترافیک	یادگیری ماشین	شبکه عصبی مصنوعی
۷	یو و همکاران [۲۸]	۲۰۱۷	چین	حجم ترافیک	یادگیری ماشین	شبکه عصبی بازگشتی
۸	یانگ و همکاران [۲۹]	۲۰۱۸	چین	وضعیت ترافیک	یادگیری ماشین	حافظه طولانی کوتاه مدت
۹	عباس و همکاران [۳۰]	۲۰۱۸	سوئد	حجم ترافیک	یادگیری ماشین	حافظه طولانی کوتاه مدت
۱۰	سان و همکاران [۳۱]	۲۰۱۸	چین	حجم ترافیک	یادگیری ماشین	K-نزدیک‌ترین همسایه
۱۱	قو و همکاران [۳۲]	۲۰۱۹	چین	حجم ترافیک	یادگیری ماشین	یادگیری عمیق
۱۲	براتساس و همکاران [۳۳]	۲۰۲۰	یونان	سرعت ترافیک	یادگیری ماشین	ماشین بردار پشتیبان
۱۳	وانگ و همکاران [۳۴]	۲۰۲۰	چین	حجم ترافیک	یادگیری ماشین	حافظه طولانی کوتاه مدت
۱۴	مالک و همکاران [۳۵]	۲۰۲۱	مراکش	سرعت ترافیک	یادگیری ماشین	حافظه طولانی کوتاه مدت

را پیش‌بینی کرده‌اند می‌توان به مطالعه زاهید [۲۰] و مطالعه‌ی ژو [۲۱] اشاره کرد. جدول ۱ خلاصه‌ای از سایر مطالعات پیشین در زمینه‌ی پیش‌بینی متغیرهای ترافیکی را به تفکیک دو رویکرد سری زمانی و یادگیری ماشین نشان می‌دهد.

یکی از مراحل حائز اهمیت در فرآیند پیش‌بینی متغیرها ترافیکی انتخاب متغیرهای پیش‌بینی کننده اثرگذار بر متغیرهای ترافیکی است. به طور کلی در مطالعات پیشین از دو دسته متغیرهای پیش‌بینی کننده استفاده شده است. دسته اول این متغیرها شامل مشاهدات پیشین همان متغیر ترافیکی می‌شود و دسته دوم اطلاعات خارجی نظیر آب و هوا، تقویم، تعطیلات و وقایع و رویدادها می‌شود. اکثر مطالعات پیشین تنها از دسته اول متغیرها استفاده کرده‌اند. جدول ۲ متغیرهای قرار گرفته در دسته دوم که در مطالعات پیشین استفاده شده‌اند را نشان می‌دهد.

در این مطالعه متغیر وضعیت ترافیک ساعتی شامل سه حالت، سبک، نیمه‌سنگین و سنگین با استفاده از سه روش مبتنی بر یادگیری ماشین شامل، ماشین بردار پشتیبان، جنگل تصادفی و حافظه طولانی کوتاه مدت پیش‌بینی شده است. به منظور پیش‌بینی وضعیت ترافیک، متغیرهای اثرگذار بر آن به

جدول ۲. متغیرهای پیش بینی کننده استفاده شده در مطالعات پیشین

Table 2. Predictor variables used in previous studies

ردیف	محققین	سال	کشور	متغیر پیش‌بینی شونده	متغیر پیش‌بینی کننده	توضیحات
۱	یانگ و همکاران [۳۶]	۲۰۱۹	ایالات متحده آمریکا	زمان سفر	سرعت متوسط ترافیک	مربوط به بالادست و پایین‌دست محور مورد نظر
					حجم ترافیک	مربوط به بالادست و پایین‌دست محور مورد نظر
					تصادفات	وقوع تصادف در همان محور، پایین‌دست، بالادست، محور موازی و جهت مخالف
					آب و هوا	دما، بارش، مه، شرایط رطوبتی روسازی و وزش باد
					رویدادهای محلی	رویدادهای ورزشی و فستیوال‌های محلی
۲	ژنگ و همکاران [۳۷]	۲۰۱۹	چین	حجم ترافیک	زمان	ساعت روز، روز هفته و نوع روز (روز تعطیل یا عادی)
					آب و هوا	دما، وزش باد، دید و بارش
۳	ژانگ و کابوکا [۳۸]	۲۰۱۸	ایالات متحده آمریکا	حجم ترافیک	زمان	ساعت روز، روز هفته و نوع روز (روز تعطیل یا عادی)
					آب و هوا	سرعت باد، بارش، دمای هوا، یخبندان و مه
	وانگ و همکاران [۳۹]	۲۰۲۰	چین	حجم ترافیک	آب و هوا	دما، سرعت باد، رطوبت، بارش و فشار اتمسفر
					تعطیلات	تعطیلات، روزهای آخر هفته
۴	نی و همکاران [۴۰]	۲۰۱۴	ایالات متحده آمریکا	حجم ترافیک	شبکه‌های اجتماعی	توییت‌های مرتبط با رویدادها، فستیوال‌ها و مسابقات ورزشی
۵	یانگ و همکاران [۴۱]	۲۰۲۰	چین	سرعت ترافیک	اطلاعات جاده	سرعت مجاز و طول جاده
					آب و هوا	وزش باد، دما، بارش
					آلودگی هوا	درصد ذرات معلق در هوا به تفکیک نوع آن

عنوان متغیرهای مستقل پیش‌بینی کننده استخراج و به مجموعه داده اضافه شده است. مدل‌های پیش‌بینی کننده وضعیت ترافیک به دو گروه مدل‌های میان مدت و مدل‌های کوتاه مدت تقسیم شده‌اند. در مدل‌های کوتاه‌مدت علاوه بر متغیرهای مستقل پیش‌بینی کننده از وضعیت ترافیک مشاهده شده در سه تا هشت ساعت گذشته نیز استفاده شده است و این مدل‌ها تنها قادر به پیش‌بینی وضعیت ترافیک یک و دو ساعت آینده هستند. در مدل‌های میان مدت از مشاهدات پیشین وضعیت ترافیک استفاده نشده است و محدودیت افق زمانی پیش‌بینی وجود ندارد. در نهایت دقت این دو دسته مدل ارزیابی

جدول ۳. تعریف متغیر وضعیت ترافیک به صورت سه گانه

Table 3. Definition of triple traffic state

نسبت حجم به ظرفیت نسبت سرعت به سرعت آزاد	کمتر از ۰/۱	۰/۳-۰/۱	۰/۵-۰/۳	۰/۷-۰/۵	۰/۹-۰/۷	بیش از ۰/۹
بیش از ۰/۹۵	روان	روان	روان	نیمه‌سنگین	نیمه‌سنگین	سنگین
۰/۹۵-۰/۸	روان	روان	نیمه‌سنگین	نیمه‌سنگین	نیمه‌سنگین	سنگین
۰/۸-۰/۶	روان	نیمه‌سنگین	نیمه‌سنگین	نیمه‌سنگین	سنگین	سنگین
۰/۶-۰/۴۵	نیمه‌سنگین	نیمه‌سنگین	نیمه‌سنگین	سنگین	سنگین	سنگین
زیر ۰/۴۵	سنگین	سنگین	سنگین	سنگین	سنگین	سنگین

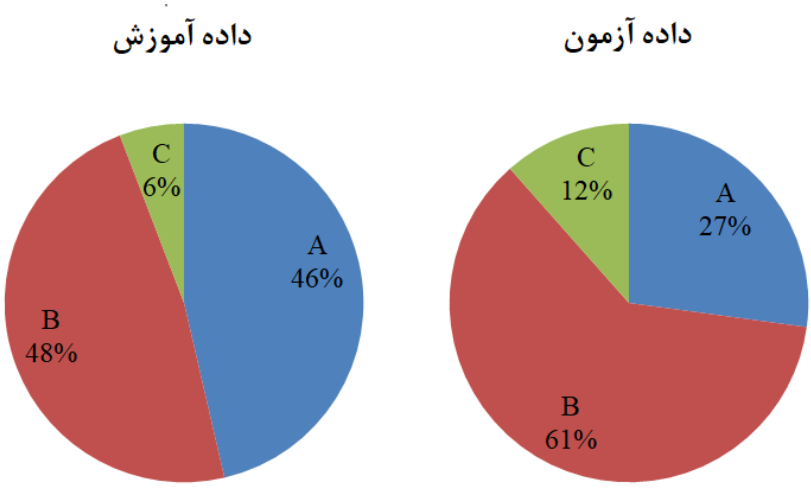
و مقایسه شده است. داده ترافیکی مورد استفاده مربوط به محور کرج به چالوس مقطع پورکان است که مشاهدات آن در بازه‌های ساعتی جمع‌آوری شده است. به طور کلی شناسایی الگوی حاکم بر متغیرهای ترافیکی در محورهای برون شهری پیچیده‌تر از محورهای درون شهری است زیرا سهم قابل‌توجهی از سفرهای درون شهری را سفرهای کاری و روزمره تشکیل شده است در صورتی که در سفرهای برون شهری سهم سفرهای روزمره و با الگوی مشخص کمتر است. به طور کلی نوآوری این پژوهش از حیث داده، معرفی متغیرهای پیش‌بینی کننده جدید در کنار استفاده از متغیرهای پیش‌بینی کننده مورد استفاده در مطالعات پیشین (معرفی شده در جدول ۲) است. متغیرهای پیش‌بینی کننده معرفی شده در این مطالعه مربوط به نوع تعطیلی، تعطیلات در روزهای گذشته و آتی و اطلاعات محدودیت‌های ترافیکی می‌شود که استخراج این متغیرها سبب تشکیل بانک اطلاعاتی منسجم و کامل شده است. همچنین پیش‌بینی وضعیت ترافیک که در مطالعات پیشین کمتر پیش‌بینی شده است برای داده برون شهری با الگوی ترافیکی نامنظم و استفاده از داده بومی کشور در حال توسعه ایران بخش دیگری از نوآوری‌های پژوهش از منظر داده مورد استفاده است. به کارگیری سه الگوریتم یادگیری ماشین شامل ماشین بردار پشتیبان، جنگل تصادفی و حافظه طولانی کوتاه مدت که به عنوان یک روش یادگیری عمیق شناخته می‌شود و ارزیابی و مقایسه نتایج بخش دیگری از نوآوری پژوهش است. همچنین پرداخت و مقایسه مدل‌های کوتاه مدت و میان مدت با افق پیش‌بینی و متغیرهای مورد استفاده متفاوت نوآوری دیگر پژوهش است.

۲- داده

داده ترافیکی استفاده شده در این مطالعه در یک مقطع از محور برون شهری کرج به چالوس توسط شمارشگرهای القایی جمع‌آوری شده است. گردآوری این داده به مدت حدود دو سال و شش ماه از ۱ فروردین ۱۳۹۶ تا ۲۰ شهریور ۱۳۹۸ انجام شده است که داده ترافیکی سال ۱۳۹۶ و ۱۰ ماه نخست ۱۳۹۷ به عنوان داده آموزش، داده دو ماه آخر سال ۱۳۹۷ به عنوان داده اعتبارسنجی و داده ترافیکی نیمه اول سال ۱۳۹۸ به عنوان داده آزمون در نظر گرفته شده است.

متغیر وضعیت ترافیکی مدنظر در این مطالعه شامل سه حالت روان، نیمه سنگین و سنگین می‌شود. ملاحظه می‌شود که این متغیر یک متغیر گسسته از نوع اسمی است و همچون سرعت و حجم ترافیک دیگر با اعداد پیوسته سروکار ندارد. نکته حائز اهمیت این است که با توجه به نوع متغیر مورد پیش‌بینی باید مدل متناسب انتخاب شود. در این پژوهش برای تشخیص وضعیت ترافیک از نسبت حجم به ظرفیت و نسبت سرعت به سرعت آزاد استفاده شده است. جدول ۳ نحوه تبدیل سرعت متوسط و حجم ترافیک به وضعیت ترافیک را نشان می‌دهد. این جدول توسط سازمان راهداری و حمل‌ونقل جاده‌ای کشور پیشنهاد شده و مورد استفاده قرار می‌گیرد. در این جدول منظور از ظرفیت، حداکثر تعداد وسیله نقلیه قابل عبور از یک مقطع از راه به ازای تمام خطوط آن در طول یک ساعت (برحسب وسیله نقلیه بر ساعت) است که بر اساس دستورالعمل‌های HCM[1] محاسبه شده است و

1 Highway capacity manual



شکل ۱. فراوانی نسبی (درصد) حالات متغیر وضعیت ترافیک به تفکیک داده آموزش و آزمون (A معادل روان، B معادل نیمهسنگین و C معادل سنگین است

Fig. 1. Relative frequency (%) of the traffic states for the train and test datasets (A is light, B is semi-heavy, and C is heavy)

منظور از سرعت آزاد سرعتی است که از ۹۵ درصد سرعت‌های مشاهده شده بیشتر باشد. ظرفیت تعیین شده بر اساس دستورالعمل‌ها به عنوان ظرفیت تئوری محور تلقی می‌شود و ظرفیت عملکردی آن می‌تواند از ظرفیت تئوری بیشتر باشد. تخمین مقادیر ظرفیت و سرعت آزاد و تعریف متغیر وضعیت ترافیک توسط سازمان راهداری و حمل‌ونقل جاده‌ای انجام شده است.

شکل ۱ فراوانی نسبی حالات متغیر وضعیت ترافیک به تفکیک داده آموزش و آزمون را نشان می‌دهد. این شکل نشان می‌دهد که این سه وضعیت ترافیک به صورت نامتوازن پراکنده شده‌اند. در مدل‌های یادگیری ماشین، هر چه فراوانی نسبی مشاهده وضعیت ترافیکی بیشتر باشد، مدل الگوی حاکم بر تغییرات آن وضعیت ترافیک را بهتر می‌آموزد و در نتیجه انتظار می‌رود دقت پیش‌بینی مدل به ازای آن وضعیت ترافیک بیشتر شود [۴۲]. همچنین شکل ۱ نشان می‌دهد در نیمه اول سال ۱۳۹۸، به نسبت سال‌های ۱۳۹۶ و ۱۳۹۷، از فراوانی نسبی وضعیت ترافیک سبک کاسته و به فراوانی نسبی وضعیت‌های نیمه‌سنگین و سنگین اضافه شده است.

داده خام اولیه شامل وضعیت ترافیک ساعتی و تاریخ است. پس از بررسی ارتباط وضعیت ترافیک ساعتی با اطلاعات تقویمی نظیر ایام خاص سال، تعطیلات و نوع تعطیلی، متغیرهای مرتبط به تقویم و تعطیلات به مجموعه داده اضافه شده است. با توجه به الگوهای ترافیکی متفاوت در ماه‌های شمسی و همچنین وجود این تفاوت در هر کدام از روزهای ماه‌های شمسی متغیرهای مربوط با تقویم شمسی شامل روز و ماه شمسی در نظر گرفته شده است. به عنوان مثال عموماً حجم سفرها به شمال کشور در ماه شهریور بیشتر از ماه مهر و در نیمه دوم ماه شهریور بیشتر از نیمه اول این ماه است. همین موارد برای تقویم قمری و متغیرهای مربوطه صادق است. به عنوان مثال حجم سفرها در ماه رمضان کمتر از سایر ماه‌ها بوده و در نیمه دوم این ماه که مصادف با شب‌های قدر است، حجم سفرها کمتر از نیمه اول این ماه است. از سوی دیگر از آنجا که ایام خاص و تعطیلی در کشور ایران بر مبنای دو تقویم شمسی و قمری هستند و این دو تقویم به نسبت یکدیگر ثابت نیستند از اطلاعات هر دو تقویم استفاده شده است. همچنین تعطیل بودن یک روز علاوه بر تغییر متغیرهای ترافیکی همان روز بر متغیرهای ترافیکی روزهای قبل و بعد آن نیز اثرگذار است زیرا بسیاری از مسافران سفر خود را قبل از روز تعطیل آغاز و تا پس از روز تعطیل ادامه می‌دهند. متغیر روز و شب بر اساس اوقات شرعی و به منظور در نظر گرفتن اثر روشنایی و نور خورشید استفاده شده است. در فصول مختلف سال اوقات شرعی در حال تغییر است و متغیر ساعت نمی‌تواند بیان‌گر روشنایی که به صورت مستقیم بر سرعت ترافیک اثرگذار است باشد. در طول روز نیز تغییرات ساعتی حجم و سرعت متوسط ترافیک ملموس است و در نظر گرفتن متغیر ساعت اهمیت دارد. اطلاعات مربوط به آب و هوا نیز بخش دیگری از عوامل اثرگذار بر وضعیت ترافیک است که این متغیر نیز استخراج و به مجموعه داده اضافه شده است. همچنین به دلیل اینکه در برخی ساعات از سال محدودیت ترافیکی یک طرفه کردن این مسیر در دو جهت رفت و برگشت و محورهای موازی آن توسط پلیس اعمال می‌شود، متغیرهای مربوطه به مجموعه داده اضافه شده است. چنانچه محورهای موازی مسدود شوند انتظار

جدول ۴. معرفی متغیرهای استخراج شده در مجموعه داده

Table 4. Introducing extracted variables in the data set

ردیف	متغیر	توضیحات
۱	وضعیت ترافیک ساعتی	وضعیت ترافیک هر مقطع در بازه زمانی یک ساعت
۲	فصل	شامل بهار، تابستان، پاییز و زمستان
۳	ماه شمسی	شامل ۱۲ ماه یک سال شمسی
۴	ماه قمری	شامل ۱۲ ماه یک سال قمری
۵	روز شمسی	شامل ۲۹-۳۱ روز یک ماه شمسی
۶	روز قمری	شامل ۲۸-۳۰ روز یک ماه قمری
۷	ساعت	شامل ۲۴ ساعت یک روز
۸	روز و شب	شامل دو حالت روز یا شب
۹	تعطیلی	بیانگر وجود تعطیلی رسمی در کشور
۱۰	نوع تعطیلی	بیانگر نوع تعطیلی رسمی در کشور
۱۱	تعداد روزهای تعطیل	تعداد روزهای تعطیل متوالی
۱۲	روزهای قبل از تعطیلی	بیانگر وجود تعطیلی ۱، ۲ و ۳ روز بعد از روز مذکور (هر کدام یک متغیر)
۱۳	نوع تعطیلی روزهای بعد	بیانگر نوع تعطیلی ۱، ۲ و ۳ روز بعد از روز مذکور (هر کدام یک متغیر)
۱۴	روزهای بعد از تعطیلی	بیانگر وجود تعطیلی ۱، ۲ و ۳ روز قبل از روز مذکور (هر کدام یک متغیر)
۱۵	نوع تعطیلی روزهای قبل	بیانگر نوع تعطیلی ۱، ۲ و ۳ روز قبل از روز مذکور (هر کدام یک متغیر)
۱۶	آب‌وهوا	شامل سه وضعیت آفتابی، بارانی و برفی
۱۷	انسداد همان مسیر	انسداد همان مسیر به علت محدودیت‌های ترافیکی یا حوادث طبیعی
۱۸	انسداد مسیر مقابل	انسداد مسیر مقابل به علت محدودیت‌های ترافیکی یا حوادث طبیعی
۱۹	انسداد محور موازی	انسداد محور موازی به علت محدودیت‌های ترافیکی یا حوادث طبیعی

می‌رود بخشی از حجم ترافیک محورهای مسدود به محور مورد نظر انتقال یابد. جدول ۴ متغیرهای استخراج شده در بانک اطلاعاتی را نشان می‌دهد.

متغیرهای معرفی شده در جدول ۴ به منظور آموزش مدل‌های میان مدت با افق پیش‌بینی امکان‌پذیر نامحدود استفاده شده‌اند. در مدل‌های کوتاه مدت علاوه بر متغیرهای جدول ۴، وضعیت ترافیک مشاهده شده در بازه‌های زمانی سه تا هشت ساعت گذشته نیز به عنوان متغیرهای پیش‌بینی کننده استفاده شده است.

۳- روش‌شناسی

در این بخش روش‌شناسی مرتبط با سه مدل ماشین بردار پشتیبان، جنگل تصادفی و حافظه طولانی کوتاه مدت بررسی شده است. هر سه مدل به دو صورت کوتاه مدت و میان مدت آموزش داده شده و متغیر وضعیت ترافیک را پیش‌بینی می‌کنند. در انتهای این بخش نیز معیارهای ارزیابی به کار گرفته شده معرفی می‌شوند.

۳- ۱- ماشین بردار پشتیبان

یکی از مدل‌های پرکاربرد در حوزه پیش‌بینی کوتاه مدت متغیرهای ترافیکی، مدل ماشین بردار پشتیبان است. بردارهای پشتیبان مجموعه‌ای از نقاط در فضای n بعدی داده‌ها هستند که مرز وضعیت‌های ترافیک را مشخص می‌کنند. به عبارت دیگر ماشین بردار پشتیبان به دنبال یافتن مرزی است که بهترین دسته‌بندی و تفکیک بین داده‌ها را مشخص می‌کند. در ماشین بردار پشتیبان فقط داده‌های قرار گرفته در مرز که با عنوان بردارهای پشتیبان شناخته می‌شوند، مبنای یادگیری ماشین قرار می‌گیرند و هدف آن هم یافتن بهترین مرز در بین داده‌ها است، به گونه‌ای که بیشترین فاصله ممکن را از تمام بردارهای پشتیبان داشته باشد. اگر داده‌ها به صورت غیرخطی توزیع شده باشند لازم است داده‌ها به کمک تابع ریاضی کرنل[1] به فضای دیگر نگاشت شوند تا در آن فضا تفکیک‌پذیر باشند. لازم به ذکر است تعیین درست این تابع نگاشت در عملکرد ماشین بردار پشتیبان مؤثر است. در نهایت پس از نگاشت داده‌ها، مدل با یافتن خطوط پشتیبان (صفحات پشتیبان در فضای چند بعدی)، سعی در یافتن معادله خطی دارد که بیشترین فاصله را بین دو دسته ایجاد کند. در این مطالعه از تابع پایه شعاعی (RBF)[2] مطابق با رابطه ۱، به عنوان تابع کرنل استفاده شده است [۴۳].

$$K(X_i, X) = exp\left(\frac{(X_i - X_j)^2}{2\varepsilon^2}\right) \quad (۱)$$

که در آن ε پارامتر آزاد مدل است و $(X_i - X_j)^2$ مجذور فاصله اقلیدسی میان دو بردار متغیر X_j و X_i است.

۳- ۲- جنگل تصادفی

مشابه با مدل ماشین بردار پشتیبان، مدل جنگل تصادفی یک مدل یادگیری ماشین است که برای مسائل رگرسیون و طبقه‌بندی کاربرد دارد. مدل جنگل تصادفی متشکل از تعداد زیادی درخت تصمیم[3] است. در این مدل داده آموزش میان مدل‌های درخت تصمیم جداگانه تقسیم شده و پس از آموزش آن‌ها، پیش‌بینی برای هر مدل درخت تصمیم انجام می‌شود. پیش‌بینی با بیشترین تکرار به عنوان پیش‌بینی کلی مدل جنگل تصادفی تعیین می‌گردد [۴۴]. مراحل زیر چگونگی کارکرد الگوریتم جنگل تصادفی را نشان می‌دهد.

گام ۱: به تعداد مدل‌های درخت تصمیم مدنظر نمونه تصادفی از مجموعه داده آموزش برداشت می‌شود.

گام ۲: با استفاده از هر نمونه یک مدل درخت تصمیم آموزش داده می‌شود.

گام ۳: پیش‌بینی هر مدل درخت تصمیم برای داده آزمون انجام می‌شود.

گام ۴: با استفاده از روش‌های رأی‌گیری، وضعیت ترافیک پرتکرار انتخاب و به عنوان پیش‌بینی نهایی گزارش می‌شود.

مدل درخت تصمیم از یک گره شروع می‌شود و گره اولیه با استفاده از شاخه‌ها به گره‌های بعدی متصل می‌شود. در این پژوهش برای مشخص شدن نحوه پیشرفت شاخه‌ها از رابطه آنتروپی (رابطه ۲) استفاده شده است [۴۴].

$$\text{Entropy} = \sum_{i=1}^{C} -p_i * \log_2(pi) \quad (۲)$$

که در آن p_i بسامد نسبی وضعیت ترافیک i و C تعداد کل وضعیت‌های ترافیک است.

۳- ۳- حافظه طولانی کوتاه مدت

مدل شبکه عصبی بازگشتی قابلیت در نظر گرفتن وابستگی مشاهدات متوالی را دارد اما نقطه ضعف اصلی این مدل عدم در نظرگیری وابستگی‌های بلند مدت است. به عنوان مثال وضعیت ترافیک ممکن است علاوه بر وضعیت ترافیک ساعات پیش به وضعیت ترافیک مدت مشابه در یک سال قبل نیز وابسته باشد. مدل حافظه طولانی کوتاه مدت که نوعی مدل شبکه عصبی بازگشتی و مدلی مبتنی بر یادگیری عمیق است، وابستگی‌های بلند مدت را نیز در نظر می‌گیرد. ورودی مدل حافظه کوتاه مدت بلند شامل حافظه کوتاه مدت، حافظه بلند مدت و برخی مشاهدات موجود در مجموعه داده آموزش می‌شود. ساختار این مدل شامل چهار دروازه می‌شود. ورودی بلند مدت مدل وارد دروازه فراموشی[4] می‌شود و این دروازه تصمیم می‌گیرد کدام یک از ورودی‌های نامرتبط حذف شود (رابطه ۳). ورودی کوتاه مدت مدل و برخی مشاهدات داده آموزش وارد دروازه آموزش[5] می‌شوند و این دروازه

1 Kernel function
2 Radial Basis Function (RBF)
3 Decision tree
4 Forget gate
5 Learn gate

تصمیم می‌گیرد کدام ورودی یاد گرفته شود (رابطه ۴). اطلاعات عبور داده شده شامل خروجی‌های دروازه‌های فراموشی و آموزش وارد دروازه یادآوری[1] می‌شوند. خروجی این دروازه حافظه‌های بلند مدت جدید است (رابطه ۵). در نهایت دروازه خروجی[2] حافظه‌های کوتاه مدت را نیز به‌روزرسانی می‌کند و خروجی نهایی مدل را تولید می‌کند (رابطه ۶) [۴۵].

$$f_t = \sigma\left(w_f\left[h_{t-1}, x_t\right] + b_f\right) \quad (۳)$$

$$l_t = \tanh\left(w_n\left[h_{t-1}, x_t\right] + b_n\right) \quad (۴)$$

$$r_t = l_t + f_t \quad (۵)$$

$$o_t = \sigma\left(w_o\left[h_{t-1}, x_t\right] + b_o\right) \quad (۶)$$

در روابط فوق f_t، l_t، r_t و o_t به ترتیب فاکتورهای دروازه‌های فراموشی، آموزش، یادآوری و خروجی هستند. σ تابع سیگموئید و w_x بردار وزن نرون‌های دروازه x است. h_{t-1} خروجی بلوک قبلی، x_t ورودی دروازه x در زمان t و b_x هم انحراف[3] دروازه x است.

۳- ۴- معیارهای ارزیابی دقت مدل‌ها

دقت[4] و شاخص F یا F_1 معیارهای مهم برای ارزیابی عملکرد نتایج مدل‌های طبقه‌بندی کننده هستند. برای محاسبه‌ی این دو معیار نتایج پیش‌بینی به چهار دسته‌ی مثبت صحیح، مثبت کاذب، منفی صحیح و منفی کاذب دسته‌بندی می‌شوند. برای مثال برای وضعیت روان، داده‌ها به صورت روان بودن و نبودن در نظر گرفته می‌شود. زمانی که وضعیت روان باشد و مدل آن را وضعیت روان پیش‌بینی کند دسته‌ی مثبت صحیح (TP) و زمانی که وضعیت روان نباشد ولی مدل وضعیت روان تشخیص دهد مثبت کاذب (FP) رخ داده است. چنانچه وضعیت غیرروان باشد و مدل آن را وضعیت غیرروان پیش‌بینی کند دسته‌ی منفی صحیح (TN) و زمانی که وضعیت غیرروان نباشد ولی مدل وضعیت غیرروان را پیش‌بینی کند منفی کاذب (FN) رخ می‌دهد [۴۶].

دقت بدین معناست که مدل تا چه اندازه خروجی را درست پیش‌بینی می‌کند. بر اساس رابطه ۷ دقت به صورت حاصل تقسیم تعداد پیش‌بینی‌های درست مدل به تعداد کل مشاهدات محاسبه می‌گردد و به صورت درصد بیان می‌شود.

$$\text{دقت} = \left(TP + TN\right) / \left(TP + FN + FP + TN\right) \quad (۷)$$

F۱ یک معیار مناسب برای ارزیابی پیش‌بینی مدل‌ها به تفکیک هر وضعیت است که با استفاده از رابطه ۸ محاسبه می‌گردد.

$$F_1 = 2 * (\text{حساسیت} \times \text{صحت}) / (\text{حساسیت} + \text{صحت}) \quad (۸)$$

که در آن صحت[5] و حساسیت[6] برای هر وضعیت ترافیک از رابطه‌های ۹ و ۱۰ محاسبه می‌شوند.

$$\text{حساسیت} = \mathrm{TP} / \left(\mathrm{TP} + \mathrm{FN}\right) \quad (۹)$$

$$\text{صحت} = \mathrm{TP} / \left(\mathrm{TP} + \mathrm{FP}\right) \quad (۱۰)$$

۴- نتایج

گام نخست مدل‌سازی انتخاب مقادیر مناسب برای پارامترهای مدل‌های یادگیری ماشین است. این پارامترها شامل پارامتر گاما (γ) و هزینه (c) در مدل ماشین بردار پشتیبان، تعداد درخت‌ها (N_T) و تعداد متغیرهای تصادفی نمونه برداری شده در هر تقسیم (N_V) در مدل جنگل تصادفی و تعداد نرون‌ها (n) و لایه‌های مخفی (m) در مدل حافظه طولانی کوتاه مدت می‌شود. در فرآیند سعی و خطا، پس از تخصیص مقادیر مختلف به این پارامترها و بررسی دقت پیش‌بینی با استفاده از داده اعتبارسنجی (این داده

1 Remember gate
2 Output gate
3 Bias
4 Accuracy
5 Precision
6 Recall

جدول ۵. مقادیر بهینه پارامترهای مدل های پیش بینی کننده

Table 5. Optimal values for parameters of predictive models

مدل	پارامتر	مدل کوتاه مدت	مدل میان مدت
ماشین بردار پشتیبان	γ	۱	۱
	c	۳۲	۲۸
جنگل تصادفی	N_T	۲۵۰	۳۲۰
	N_V	۳۰	۳۵
حافظه طولانی کوتاه مدت	n	۳۰	۴۰
	m	۳	۴

جدول ۶. دقت پیش بینی وضعیت ترافیک

Table 6. Accuracy of traffic state predictions

مدل	دقت مدل کوتاه مدت (درصد)		دقت مدل میان مدت (درصد)	
	آموزش	آزمون	آموزش	آزمون
ماشین بردار پشتیبان	۹۰/۸۲	۸۴/۷۴	۸۳/۸۲	۷۶/۶۳
جنگل تصادفی	۹۲/۴۵	۸۷/۵۱	۸۴/۱۳	۷۹/۰۴
حافظه طولانی کوتاه مدت	۹۶/۸۸	۹۰/۱۱	۸۸/۶۹	۸۲/۰۷

در فرآیند آموزش مدل‌ها استفاده نشده است)، مقادیر بهینه پارامترها به ازای مدل‌های کوتاه مدت و میان مدت مشخص شده و در جدول ۵ نشان داده شده است.

پس از آموزش مدل‌های نهایی کوتاه مدت و میان مدت، وضعیت ترافیک برای داده آزمون پیش‌بینی شده است. جدول ۶ دقت پیش‌بینی وضعیت ترافیک به تفکیک مدل‌های کوتاه مدت و میان مدت و به تفکیک داده آموزش و آزمون را نشان می‌دهد.

جدول ۶ نشان می‌دهد برای هر دو افق پیش‌بینی کوتاه مدت و میان مدت، مدل حافظه طولانی کوتاه مدت بیشترین دقت و مدل ماشین بردار پشتیبان کم‌ترین دقت را نتیجه می‌دهند. نکته قابل‌توجه دیگر دقت بیشتر مدل‌های کوتاه مدت به نسبت مدل‌های میان مدت است. همان‌طور که انتظار می‌رفت استفاده از مشاهدات پیشین به عنوان متغیر پیش‌بینی کننده باعث افزایش ۷ تا ۸ درصدی دقت مدل‌ها به نسبت زمانی که از این متغیرها استفاده نشود، شده است. در عین حال توجه به این نکته که مدل‌های کوتاه مدت تنها می‌توانند یک و دو ساعت آینده را پیش‌بینی کنند و چنین محدودیت برای مدل‌های میان مدت وجود ندارد، ضرورت دارد.

یکی دیگر از جنبه‌های مهم در ارزیابی عملکرد مدل‌های پیش‌بینی کننده ترافیک بررسی دقت پیش‌بینی به تفکیک هر وضعیت ترافیک است. بدین منظور با استفاده از ماتریس‌های آشفتگی که در جدول ۷ نشان داده شده است، شاخص F_1 به تفکیک مدل‌های کوتاه مدت و میان مدت و به تفکیک وضعیت‌های ترافیک برای داده آزمون محاسبه شده و در جدول ۸ نشان داده شده است.

جدول ۸ نشان می‌دهد بر اساس شاخص F_1، برای هر دو دسته مدل کوتاه مدت و میان مدت، مدل حافظه طولانی کوتاه مدت هر سه وضعیت ترافیک را دقیق‌تر از دو مدل دیگر پیش‌بینی می‌کند. همچنین دقت پیش‌بینی هر وضعیت ترافیک با استفاده از مدل جنگل تصادفی بیشتر از مدل ماشین بردار پشتیبان است. به عنوان یک نکته حائز اهمیت دیگر، هر سه مدل وضعیت ترافیک نیمه‌سنگین که بیشترین درصد فراوانی را در داده آموزش و آزمون دارد را دقیق‌تر از دو وضعیت ترافیک دیگر پیش‌بینی می‌کنند. در مقابل دقت پیش‌بینی وضعیت ترافیک سنگین که کمترین درصد

جدول ۷. ماتریس های آشفتگی پیش بینی وضعیت ترافیک

Table 7. Confusion matrices of traffic state predictions

مدل		وضعیت	کوتاه مدت			میان مدت		
			واقعیت					
			روان	نیمه‌سنگین	سنگین	روان	نیمه‌سنگین	سنگین
ماشین بردار پشتیبان	پیش بینی	روان	۸۸۱	۱۵۱	۵۱	۸۱۰	۳۱۲	۷۶
		نیمه‌سنگین	۲۱۷	۲۴۵۱	۱۲۲	۲۹۳	۲۲۴۵	۱۷۲
		سنگین	۶۸	۵۳	۳۴۵	۶۳	۹۸	۲۷۰
جنگل تصادفی		روان	۹۶۶	۱۶۴	۳۰	۸۹۰	۳۳۲	۴۶
		نیمه‌سنگین	۱۵۵	۲۴۳۲	۸۹	۲۱۱	۲۲۱۹	۱۵۰
		سنگین	۴۵	۵۹	۳۹۹	۶۵	۱۰۴	۳۲۲
حافظه طولانی کوتاه مدت		روان	۱۰۰۶	۱۲۵	۳۶	۹۳۵	۲۷۱	۴۵
		نیمه‌سنگین	۱۱۷	۲۴۹۶	۷۴	۱۷۱	۲۲۹۲	۱۳۹
		سنگین	۴۳	۳۴	۴۰۸	۶۰	۹۲	۳۳۴

جدول ۸. صحت، حساسیت و F^1 به تفکیک مدل های کوتاه مدت و میان مدت و به تفکیک وضعیت های ترافیک

Table 8. Precision, recall, and F1 for short-term and medium-term models for each traffic state

مدل	وضعیت	کوتاه مدت			میان مدت		
		صحت	حساسیت	F_1	صحت	حساسیت	F_1
ماشین بردار پشتیبان	روان	۰/۸۱	۰/۷۶	۰/۷۸	۰/۶۸	۰/۶۹	۰/۶۹
	نیمه‌سنگین	۰/۸۸	۰/۹۲	۰/۹	۰/۸۳	۰/۸۵	۰/۸۴
	سنگین	۰/۷۴	۰/۶۷	۰/۷	۰/۶۳	۰/۵۲	۰/۵۷
جنگل تصادفی	روان	۰/۸۴	۰/۸۳	۰/۸۳	۰/۷	۰/۷۶	۰/۷۳
	نیمه‌سنگین	۰/۹۱	۰/۹۲	۰/۹۱	۰/۸۶	۰/۸۴	۰/۸۵
	سنگین	۰/۷۹	۰/۷۷	۰/۷۸	۰/۶۶	۰/۶۲	۰/۶۴
حافظه طولانی کوتاه مدت	روان	۰/۸۶	۰/۸۶	۰/۸۶	۰/۷۵	۰/۸	۰/۷۷
	نیمه‌سنگین	۰/۹۳	۰/۹۴	۰/۹۳	۰/۸۸	۰/۸۶	۰/۸۷
	سنگین	۰/۸۴	۰/۷۹	۰/۸۱	۰/۶۹	۰/۶۴	۰/۶۷

جدول ۹. میزان کاهش دقت پیش بینی پس از حذف هر گروه متغیر

Table 9. The rate of decrease in prediction accuracy after the removal of each variable group

گروه متغیر	کوتاه مدت			میان مدت		
	ماشین بردار پشتیبان	جنگل تصادفی	حافظه طولانی کوتاه مدت	ماشین بردار پشتیبان	جنگل تصادفی	حافظه طولانی کوتاه مدت
فصل	۱/۷۵	۱/۶۳	۲/۹	۳/۲۵	۴/۷۱	۴/۸۰
ماه و روز شمسی	۲/۸۳	۳/۵	۴/۶۱	۶/۸۴	۶/۴۲	۶/۸۳
ماه و روز قمری	۰/۹۳	۰/۸۵	۱/۲۱	۲/۳۶	۲/۳۳	۲/۲۸
ساعت	۶/۸۶	۶/۴۳	۵/۱۴	۱۱/۱۱	۱۲/۷۱	۱۰/۱۴
روز و شب	۱/۵۵	۱/۹۲	۲/۷۷	۳/۸۴	۴/۳۵	۴/۱۱
تعطیلی همان روز و نوع آن	۶/۰۱	۶/۲۱	۵/۲۳	۸/۴۴	۸/۶۶	۹/۳۸
تعطیلی روزهای قبل و بعد و نوع آن	۱/۸۳	۲/۶۵	۱/۹۶	۴/۸۴	۴/۱۵	۴/۳۶
آب‌وهوا	۱/۹۱	۱/۷۵	۱/۹۸	۳/۴۶	۳/۲	۳/۹۹
انسداد مسیر، مسیر مقابل و محور موازی	۲/۴۴	۲/۰۹	۲/۳۲	۳/۲۳	۲/۷۶	۲/۸۷
مشاهدات سه تا هشت ساعت پیش وضعیت ترافیک	۸/۱۱	۸/۴۷	۸/۰۴	–	–	–

فراوانی را در داده آموزش و آزمون دارد، کمتر از وضعیت ترافیک روان است. مقایسه پیش‌بینی مدل‌های کوتاه مدت و میان مدت نیز نشان می‌دهد مقادیر F_1 مدل‌های کوتاه مدت کمتر از مقادیر F_1 مدل‌های میان مدت است و مدل‌های کوتاه مدت هر سه وضعیت ترافیک را به ازای سه مدل مدنظر، دقیق‌تر از مدل‌های میان مدت پیش‌بینی می‌کنند.

به منظور انجام تحلیل حساسیت در خصوص نقش پیش‌بینی کنندگی متغیرهای مستقل مورد استفاده، ابتدا گروه متغیرهای از یک جنس تعریف شده و مدل‌ها بدون استفاده از هر گروه متغیر آموزش داده شده است. میزان کاهش دقت کلی مدل‌ها پس از حذف هر گروه متغیر محاسبه و در جدول ۹ نشان داده شده است.

به طور کلی نتایج زیر از جدول ۹ برداشت می‌شود:

- استفاده از متغیرهای ساعت و تعطیلی همان روز و نوع تعطیلی آن بیشترین تأثیر را بر افزایش دقت مدل‌های میان مدت دارند. عموماً پراکندگی سفرهای برون شهری در طول روز یکنواخت نیست و به همین دلیل استفاده از متغیر ساعت می‌تواند تغییرات متغیر وضعیت ترافیک را به تصویر بکشد. همچنین بخش زیادی از سفرها به سمت مقاصد شمال کشور از نوع سفرهای تفریحی است لذا می‌توان انتظار داشت که در ایام تعطیل سال با توجه به حجم بیشتر سفرهای تفریحی الگوی ترافیک تغییر یابد. البته تغییر الگوی ترافیکی تابعی از نوع تعطیلی است. در برخی از ایام تعطیل سال نظیر تعطیلی‌های ماه‌های محرم و رمضان به دلیل وجود آیین‌های مذهبی حجم سفرها به سمت مقاصد شمال کشور کمتر از سایر ایام تعطیل سال است.

جدول ۱۰. دقت کلی پیش بینی وضعیت ترافیک در مطالعات پیشین

Table 10. Overall accuracy of traffic state predictions in previous studies

نویسنده	تعداد وضعیت‌های ترافیک	تعداد مشاهدات	روش	دقت کلی (درصد)
آنتونیو و همکاران [۴۷]	۳	۷۰۵	K-نزدیک‌ترین همسایه	۷۹/۷۱
			شبکه عصبی مصنوعی	۸۳/۷۰
	۴	۷۰۵	K-نزدیک‌ترین همسایه	۸۶/۳۷
			شبکه عصبی مصنوعی	۸۶/۶۳
زاهید و همکاران [۲۰]	۶	نامشخص	جنگل تصادفی	۹۷/۲۲
			ماشین بردار پشتیبان	۹۶/۷۶
			شبکه عصبی مصنوعی	۹۵/۷۷
			استنتاج قوانین[1]	۹۵/۸۱
ژو و همکاران [۲۱]	۴	۴۹۰۳۸	درخت تصمیم	۵۱
			ماشین بردار پشتیبان	۵۷
			شبکه عصبی مصنوعی	۷۲
			شبکه عصبی	۸۵
			شبکه عصبی بازگشتی	۸۹/۵
تنچاروئن و پیانتاناکولچای [۴۸]	۳	۱۰۸۰۰	شبکه عصبی	۸۹
			شبکه عصبی مصنوعی	۸۷
			K-نزدیک‌ترین همسایه	۸۷
			جنگل تصادفی	۸۷

- استفاده از مشاهدات سه تا هشت ساعت پیش وضعیت ترافیک بیشترین تأثیر را بر افزایش دقت مدل‌های کوتاه مدت دارند. حذف این متغیرها مدل را از حالت کوتاه مدت به میان مدت تبدیل می‌کند (دقت مدل میان مدت حاصل می‌شود).
- به طور کلی با توجه به اثر چشمگیر متغیرهای مربوط به مشاهدات سه تا هشت ساعت پیش وضعیت ترافیک در مدل‌های کوتاه مدت، حذف سایر متغیرها اثر کمتری بر کاهش دقت پیش‌بینی به نسبت مدل‌های میان مدت می‌گذارد. معمولاً انتظار می‌رود شرایط ترافیکی تا چندین ساعت آینده ثبات داشته باشد، لذا اهمیت این متغیر در پیش‌بینی وضعیت ترافیک با مدل‌های کوتاه مدت قابل‌انتظار است. به عنوان مثال چنانچه در محوری راه‌بندان رخ دهد، این معضل ممکن است وضعیت ترافیک آن محور تا چندین ساعت آینده را تحت تأثیر خود قرار دهد.
- هر سه مدل ماشین بردار پشتیبان، جنگل تصادفی و حافظه طولانی کوتاه مدت، تأثیر مشابهی از حذف متغیرها می‌گیرند؛ به عبارت دیگر محدوده تغییرات دقت پیش‌بینی هر سه مدل به یکدیگر نزدیک است. این یافته نشان می‌دهد عوامل اثرگذار بر الگوهای ترافیکی ثابت بوده و تابعی از روش مورد استفاده نیست.

- متغیری وجود ندارد که حذف آن سبب افزایش دقت پیش‌بینی شود. با استفاده از این یافته می‌توان نشان داد که استفاده از تمام متغیرهای تعریف شده در جدول ۲ موجه بوده و هر کدام از این متغیرها به نحوی بر پیش‌بینی دقیق‌تر وضعیت ترافیک مؤثر هستند.
-

۵- نتیجه‌گیری

یکی از کارکردهای سیستم‌های حمل‌ونقل هوشمند، اطلاع‌رسانی شرایط آینده ترافیک است. فرد مسافر با آگاهی از شرایط آینده ترافیک برنامه‌ریزی مناسب‌تری برای انتخاب انجام سفر، زمان عزیمت و انتخاب مسیر خواهد داشت و از سوی دیگر گرداننده شبکه حمل‌ونقل آمادگی بیشتری برای مواجهه با تراکم ترافیک خواهد داشت. در این مطالعه وضعیت ترافیک، شامل سه وضعیت روان، نیمه سنگین و سنگین، به عنوان یک پارامتر نشان دهنده عملکرد کمان‌های شبکه حمل‌ونقل با استفاده از سه روش ماشین بردار پشتیبان، جنگل تصادفی و حافظه طولانی کوتاه مدت برای محور برون شهری کرج به چالوس در شمال ایران پیش‌بینی شده است. در گام نخست عملیات پیش‌پردازش داده با اضافه کردن متغیرهای پیش‌بینی کننده‌ی مرتبط با تاریخ، تعطیلات، نوع تعطیلات و آب و هوا انجام شده است. همچنین در مدل‌های کوتاه مدت علاوه بر متغیرهای نام برده از مشاهدات وضعیت ترافیک سه تا هشت ساعت گذشته نیز استفاده شده است. در گام دوم جهت یافتن مقادیر بهینه پارامترهای مدل‌های کوتاه مدت و میان مدت، مدل‌ها به ازای مقادیر مختلف این پارامترها آموزش داده شده و بر اساس دقت پیش‌بینی بر روی داده اعتبارسنجی مقادیر بهینه انتخاب شده است. پس از آموزش مدل‌ها با مقادیر بهینه، پیش‌بینی انجام و نتایج سه مدل یادگیری ماشین به دو صورت کوتاه مدت و میان مدت ارزیابی و مقایسه شده است. همچنین ارزیابی به تفکیک پیش‌بینی هر وضعیت ترافیک نیز انجام شده است. در نهایت به منظور مشخص شدن تأثیر استفاده از هر گروه متغیر بر افزایش دقت پیش‌بینی مدل‌ها، مدل‌ها بدون هر گروه متغیر آموزش داده شده و دقت حاصله برای داده آزمون ارزیابی شده است. نتایج نشان می‌دهد مدل کوتاه مدت حافظه طولانی کوتاه مدت بیشترین دقت پیش‌بینی وضعیت ترافیک معادل با ۹۰/۱۱ درصد را به همراه دارد. همچنین مدل میان مدت حافظه طولانی کوتاه مدت، با افق پیش‌بینی نامحدود، با دقت معادل با ۸۲/۰۷ درصد دقیق‌تر از دو مدل میان مدت دیگر وضعیت ترافیک را پیش‌بینی می‌کند. این مدل در هر دو حالت کوتاه‌مدت و میان مدت، هر سه وضعیت ترافیک روان، نیمه سنگین و سنگین را دقیق‌تر از دو مدل دیگر پیش‌بینی می‌کند. همچنین استفاده از متغیرهای ساعت و تعطیلی همان روز و نوع آن بیشترین تأثیر را بر دقت مدل‌های میان مدت و متغیر وضعیت ترافیک مشاهدات سه تا هشت ساعت گذشته بیشترین تأثیر را بر دقت مدل‌های کوتاه مدت داشته‌اند.

به منظور مشخص شدن جایگاه دقت پیش‌بینی وضعیت ترافیک به نسبت نتایج مطالعات پیشین، جدول ۱۰ دقت پیش‌بینی وضعیت ترافیک به تفکیک تعداد وضعیت‌ها و روش مورد استفاده را نشان می‌دهد. مقایسه دقت حاصل شده در این پژوهش با نتایج ارائه شده در جدول ۱۰ نشان می‌دهد، دقت کلی معادل با ۹۱/۱۱ درصد در مدل‌های کوتاه مدت و دقت معادل ۸۲/۰۷ درصد در مدل‌های بلند مدت، به نسبت دقت به دست آمده در مطالعات پیشین قابل‌قبول به نظر می‌رسد. البته توجه به تفاوت الگوهای ترافیکی، تعداد مشاهدات، تعداد وضعیت‌های ترافیک و روش مورد استفاده در مطالعات ارائه شده در جدول ۱۰ و این پژوهش ضرورت دارد.

در پایان وضعیت ترافیک پیش‌بینی شده با استفاده از مدل‌های کوتاه مدت و بلند مدت می‌تواند از طریق سیستم‌های هوشمند اطلاعات مسافر اطلاع‌رسانی شود. به منظور استفاده توأمان از دقت پیش‌بینی بیشتر مدل‌های کوتاه مدت و افق پیش‌بینی بلندتر مدل‌های میان مدت می‌توان در سیستم هوشمند اطلاعات مسافر که به صورت به هنگام عمل می‌کند، پیش‌بینی یک و دو ساعت آینده توسط مدل‌های کوتاه مدت و پیش‌بینی‌های مربوط به بیش از دو ساعت آینده توسط مدل‌های میان مدت انجام شود تا این دو مدل به صورت مکمل یکدیگر عمل کنند.

به عنوان پیشنهاد برای ادامه این پژوهش به کارگیری سایر مدل‌های یادگیری ماشین نظیر شبکه عصبی کانولوشنالی و ارزیابی نتایج آن توصیه می‌گردد. همچنین انجام تحلیل حساسیت در زمینه‌ی بررسی اثر گذشت زمان بر دقت پیش‌بینی وضعیت ترافیک، پیشنهاد می‌گردد.

تشکر و قدردانی

بدین‌وسیله از مسئولین و کارشناسان سازمان راهداری و حمل‌ونقل جاده‌ای در مرکز مدیریت راه‌های کشور بابت حمایت این پژوهش و در اختیار قرار دادن داده ترافیکی راه‌های برون‌شهری تشکر و قدردانی می‌شود.

۶- فهرست علائم

علائم انگلیسی

c	پارامتر هزینه مدل ماشین بردار پشتیبان
C	تعداد حالات متغیر وضعیت ترافیک در مدل جنگل تصادفی
FN	منفی کاذب
FP	مثبت کاذب
f_t	فاکتور دروازه فراموشی مدل حافظه طولانی کوتاه مدت
l_t	فاکتور دروازه آموزش مدل حافظه طولانی کوتاه مدت
n	تعداد نرون‌های مدل حافظه طولانی کوتاه مدت
N_T	تعداد درخت‌ها در مدل جنگل تصادفی
N_V	تعداد متغیرهای تصادفی نمونه-برداری شده در هر تقسیم در مدل جنگل تصادفی
m	تعداد لایه‌های مدل حافظه طولانی کوتاه مدت
o_t	فاکتور دروازه خروجی مدل حافظه طولانی کوتاه مدت
p_i	بسامد نسبی وضعیت ترافیک i در مدل جنگل تصادفی
r_t	فاکتور دروازه یادآوری مدل حافظه طولانی کوتاه مدت

علائم یونانی

ε	پارامتر آزاد مدل ماشین بردار پشتیبان
γ	پارامتر گاما مدل ماشین بردار پشتیبان
σ	تابع سیگموئید

منابع

[1] A.A. Faiyetole, A Review of Intelligent Transport System and Its People's Needs Considerations for Traffic Management's Policy Framework in a Developing Country: People's Needs Considerations for ITS Policy, Global Advancements in Connected and Intelligent Mobility: Emerging Research and Opportunities, (2020) 166-195.

[2] X. Chen, J. Lu, J. Zhao, Z. Qu, Y. Yang, J. Xian, Traffic flow prediction at varied time scales via ensemble empirical mode decomposition and artificial neural network, Sustainability, 12(9) (2020) 3678.

[3] C.D. Harper, S. Qian, C. Samaras, Improving Short-Term Travel Speed Prediction with High-Resolution Spatial and Temporal Rainfall Data, Journal of Transportation Engineering, Part A: Systems, 147(3) (2021) 04021004.

[4] D. Xu, C. Wei, P. Peng, Q. Xuan, H. Guo, GE-GAN: A novel deep learning framework for road traffic state estimation, Transportation Research Part C: Emerging Technologies, 117 (2020) 102635.

[5] A. Pankratz, Forecasting with univariate Box-Jenkins models: Concepts and cases, John Wiley & Sons, 2009.

[6] S.V. Kumar, L. Vanajakshi, Short-term traffic flow prediction using seasonal ARIMA model with limited input data, European Transport Research Review, 7(3) (2015) 1-9.

[7] D. Yan, J. Zhou, Y. Zhao, B. Wu, Short-term subway passenger flow prediction based on ARIMA, in: International Conference on Geo-Spatial Knowledge and Intelligence, Springer, 2017, pp. 464-479.

[8] B.L. Smith, B.M. Williams, R.K. Oswald, Comparison of parametric and nonparametric models for traffic flow forecasting, Transportation Research Part C: Emerging Technologies, 10(4) (2002) 303-321.

[9] E.I. Vlahogianni, J.C. Golias, M.G. Karlaftis, Short-term traffic forecasting: Overview of objectives and methods, Transport reviews, 24(5) (2004) 533-557.

[10] B. Sun, W. Cheng, P. Goswami, G. Bai, Flow-aware WPT k-nearest neighbours regression for short-term traffic prediction, in: 2017 IEEE Symposium on Computers and Communications (ISCC), IEEE, 2017, pp. 48-53.

[11] F.I. Rahman, SHORT TERM TRAFFIC FLOW PREDICTION USING MACHINE LEARNING-KNN, SVM AND ANN WITH WEATHER INFORMATION, International Journal for Traffic & Transport Engineering, 10(3) (2020).

[12] Y. Li, M. Chen, X. Lu, W. Zhao, Research on optimized GA-SVM vehicle speed prediction model based on driver-vehicle-road-traffic system, Science China Technological Sciences, 61(5) (2018) 782-790.

[13] Y. Liu, H. Wu, Prediction of road traffic congestion based on random forest, in: 2017 10th International Symposium on Computational Intelligence and Design

(ISCID), IEEE, 2017, pp. 361-364.

[14] R. CHENG, M.-M. ZHANG, Y. Xue-Mei, Prediction Model for Road Traffic Accident Based on Random Forest, DEStech Transactions on Social Science, Education and Human Science, (icesd) (2019).

[15] B. Sharma, S. Kumar, P. Tiwari, P. Yadav, M.I. Nezhurina, ANN based short-term traffic flow forecasting in undivided two lane highway, Journal of Big Data, 5(1) (2018) 1-16.

[16] A. Raza, M. Zhong, Lane-based short-term urban traffic forecasting with GA designed ANN and LWR models, Transportation research procedia, 25 (2017) 1430-1443.

[17] N. Ramakrishnan, T. Soni, Network traffic prediction using recurrent neural networks, in: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, 2018, pp. 187-193.

[18] Z. Zhao, W. Chen, X. Wu, P.C. Chen, J. Liu, LSTM network: a deep learning approach for short-term traffic forecast, IET Intelligent Transport Systems, 11(2) (2017) 68-75.

[19] N. Ranjan, S. Bhandari, H.P. Zhao, H. Kim, P. Khan, City-Wide Traffic Congestion Prediction Based on CNN, LSTM and Transpose CNN, IEEE Access, 8 (2020) 81606-81620.

[20] M. Zahid, Y. Chen, A. Jamal, M.Q. Memon, Short term traffic state prediction via hyperparameter optimization based classifiers, Sensors, 20(3) (2020) 685.

[21] J. Zhu, C. Huang, M. Yang, G.P.C. Fung, Context-based prediction for road traffic state using trajectory pattern mining and recurrent convolutional neural networks, Information Sciences, 473 (2019) 190-201.

[22] G. Rutka, Network traffic prediction using ARIMA and neural networks models, Elektronika ir Elektrotechnika, 84(4) (2008) 53-58.

[23] A.M. Khoei, A. Bhaskar, E. Chung, Travel time prediction on signalised urban arterials by applying SARIMA modelling on Bluetooth data, in: 36th Australasian transport research forum (ATRF) 2013,

2013.

[24] N.K. Chikkakrishna, C. Hardik, K. Deepika, N. Sparsha, Short-term traffic prediction using sarima and FbPROPHET, in: 2019 IEEE 16th India Council International Conference (INDICON), IEEE, 2019, pp. 1-4.

[25] T. Alghamdi, K. Elgazzar, M. Bayoumi, T. Sharaf, S. Shah, Forecasting traffic congestion using ARIMA modeling, in: 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), IEEE, 2019, pp. 1227-1232.

[26] V. Eramo, T. Catena, F. Lavacca, F. Di Giorgio, Study and investigation of SARIMA-based traffic prediction models for the resource allocation in NFV networks with elastic optical interconnection, in: 2020 22nd International Conference on Transparent Optical Networks (ICTON), IEEE, 2020, pp. 1-4.

[27] C. Goves, R. North, R. Johnston, G. Fletcher, Short term traffic prediction on the UK motorway network using neural networks, Transportation Research Procedia, 13 (2016) 184-195.

[28] H. Yu, Z. Wu, S. Wang, Y. Wang, X. Ma, Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks, Sensors, 17(7) (2017) 1501.

[29] G. Yang, Y. Wang, H. Yu, Y. Ren, J. Xie, Short-term traffic state prediction based on the spatiotemporal features of critical road sections, Sensors, 18(7) (2018) 2287.

[30] Z. Abbas, A. Al-Shishtawy, S. Girdzijauskas, V. Vlassov, Short-term traffic prediction using long short-term memory neural networks, in: 2018 IEEE International Congress on Big Data (BigData Congress), IEEE, 2018, pp. 57-65.

[31] B. Sun, W. Cheng, P. Goswami, G. Bai, Short-term traffic forecasting using self-adjusting k-nearest neighbours, IET Intelligent Transport Systems, 12(1) (2017) 41-48.

[32] L. Qu, W. Li, W. Li, D. Ma, Y. Wang, Daily long-term traffic flow forecasting based on a deep neural network,

Expert Systems with applications, 121 (2019) 304-312.

[33] C. Bratsas, K. Koupidis, J.-M. Salanova, K. Giannakopoulos, A. Kaloudis, G. Aifadopoulou, A comparison of machine learning methods for the prediction of traffic speed in urban places, Sustainability, 12(1) (2020) 142.

[34] Z. Wang, X. Su, Z. Ding, Long-term traffic prediction based on lstm encoder-decoder architecture, IEEE Transactions on Intelligent Transportation Systems, (2020).

[35] Y.N. Malek, M. Najib, M. Bakhouya, M. Essaaidi, Multivariate deep learning approach for electric vehicle speed forecasting, Big Data Mining and Analytics, 4(1) (2021) 56-64.

[36] S. Yang, S. Qian, Understanding and predicting travel time with spatio-temporal features of network traffic flow, weather and incidents, IEEE Intelligent Transportation Systems Magazine, 11(3) (2019) 12-28.

[37] C. Zheng, X. Fan, C. Wen, L. Chen, C. Wang, J. Li, Deepstd: Mining spatio-temporal disturbances of multiple context factors for citywide traffic flow prediction, IEEE Transactions on Intelligent Transportation Systems, 21(9) (2019) 3744-3755.

[38] D. Zhang, M.R. Kabuka, Combining weather condition data to predict traffic flow: a GRU-based deep learning approach, IET Intelligent Transport Systems, 12(7) (2018) 578-585.

[39] J. Wang, W. Zhu, Y. Sun, C. Tian, An effective dynamic spatiotemporal framework with external features information for traffic prediction, Applied Intelligence, (2020) 1-15.

[40] M. Ni, Q. He, J. Gao, Using social media to predict traffic flow under special event conditions, in: The 93rd annual meeting of transportation research board, 2014.

[41] X. Yang, Y. Yuan, Z. Liu, Short-term traffic speed prediction of urban road with multi-source data, IEEE Access, 8 (2020) 87541-87551.

[42] C. Ranjan, M. Reddy, M. Mustonen, K. Paynabar, K. Pourak, Dataset: rare event classification in multivariate time series, arXiv preprint arXiv:1809.10717, (2018).

[43] D.A. Pisner, D.M. Schnyer, Support vector machine, in: Machine Learning, Elsevier, 2020, pp. 101-121.

[44] G. Biau, E. Scornet, A random forest guided tour, Test, 25(2) (2016) 197-227.

[45] R.C. Staudemeyer, E.R. Morris, Understanding LSTM--a tutorial into Long Short-Term Memory Recurrent Neural Networks, arXiv preprint arXiv:1909.09586, (2019).

[46] D.M. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, arXiv preprint arXiv:2010.16061, (2020).

[47] C. Antoniou, H.N. Koutsopoulos, G. Yannis, Dynamic data-driven local traffic state estimation and prediction, Transportation Research Part C: Emerging Technologies, 34 (2013) 89-107.

[48] R. Toncharoen, M. Piantanakulchai, Traffic state prediction using convolutional neural network, in: 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE, 2018, pp. 1-6.

چگونه به این مقاله ارجاع دهیم

A. Rasaizadi, S. H. Seyedabrishami , Traffic state prediction with machine learning algorithms for short-term and mid-term prediction time horizons, Amirkabir J. Civil Eng., 54(4) (2022) 1503-1520.

DOI: 10.22060/ceej.2021.19650.7219

این صفحه عمداً خالی است