



Predicting construction project scheduling issues using LSTM neural network (long-term short-term memory)

E. Farzad, H. Dehghan Menshadi*, M. A. Dashti

Department of Civil Engineering, Yazd Branch, Islamic Azad University, Yazd, Iran

ABSTRACT: As the purpose of monitoring the project is to make accurate decisions that can have significant effects on the project's success, predicting the project's characteristics becomes more important. According to experts, schedule delays are a frequent issue in many construction projects. This research aims to propose a model that can address project scheduling problems. For this purpose, this study proposes new applications of recurrent neural network architectures based on short-term long-term memory (LSTM) prediction models. Subsequently, the prediction results of the presented models are compared and verified with the historical data of a real project. The data used in this study has been obtained from the South Extension Project of Tehran Metro Line 6. The project started in October 2016 and ended in July 2018, lasting for a total of 21 months. In this study, the training dataset consisted of the initial 14 months' data, which accounted for 83 percent of the total data. We used the construction project progress as a forecasting variable. To evaluate the performance of LSTM models, we used the mean square error (MSE) metric as the evaluation criterion. The results show that the model accurately forecasts the project's future progress based on its past progress.

Review History:

Received: May, 08, 2022
Revised: Apr. 29, 2023
Accepted: Aug. 12, 2023
Available Online: Aug. 20, 2023

Keywords:

Project management
scheduling
artificial intelligence
short term long term memory
forecasting

1- Introduction

According to published statistics by the Office of Supervision and Evaluation of Plans, more than 50% of the country's construction projects face delays every year. The typical duration for project execution is nearly eight years, whereas it was originally expected to be completed within a three-years. Scheduling is crucial in construction projects as it enables effective project management, resource allocation, and timely completion. Uncertainty in project planning refers to the unpredictable events that can affect the project's scope, quality, schedule, or budget [1]. Furthermore, accurate forecasting allows for enhanced project monitoring. scheduling under uncertainty is primarily investigated through the utilization of stochastic optimization methods. Additionally, to overcome project scheduling challenges, several approaches based on genetic algorithms have been proposed [2]. However, deep learning models are a more effective way to handle the uncertainties of project scheduling. They can forecast the uncertainties better than other methods [1]. By utilizing deep ensemble networks, it is possible to achieve high-quality uncertainty estimation with a small value of the prediction interval width and a high confidence of prediction interval coverage probability [3]. Another approach in the prediction problem involves the use of neural networks and machine learning. Artificial neural

networks are inspired by the brain, and their computations could be implemented in biological neurons. These artificial neural networks, with their non-linear mapping capability and effective topology, are widely used in pattern recognition, data prediction, and more [4]. Among the existing models in neural networks, Recurrent Neural Networks (RNNs) have the ability to consider sequential observations. In other words, this category of RNN models processes sequential data as time series [5]. The main drawback of RNNs is that they have difficulty in learning long-term dependencies in sequential data [6]. The Long Short-Term Memory (LSTM), which is a special kind of RNN that can overcome this problem by using the gates to control the amount of information that flows into and out of the memory cell, which prevents the gradients from vanishing or exploding. Therefore, in this research, the progress of civil engineering projects is predicted by applying the LSTM method.

2- Framework of LSTM Model

The LSTM consists of unique units known as memory blocks, which replace the hidden layer found in neural networks. The input gate, forget gate, output gate, and input modulation gate are the features of the LSTM structure. This study used three hidden layers. The equations governing these are

*Corresponding author's email: sh.manshadi@iau.ac.ir



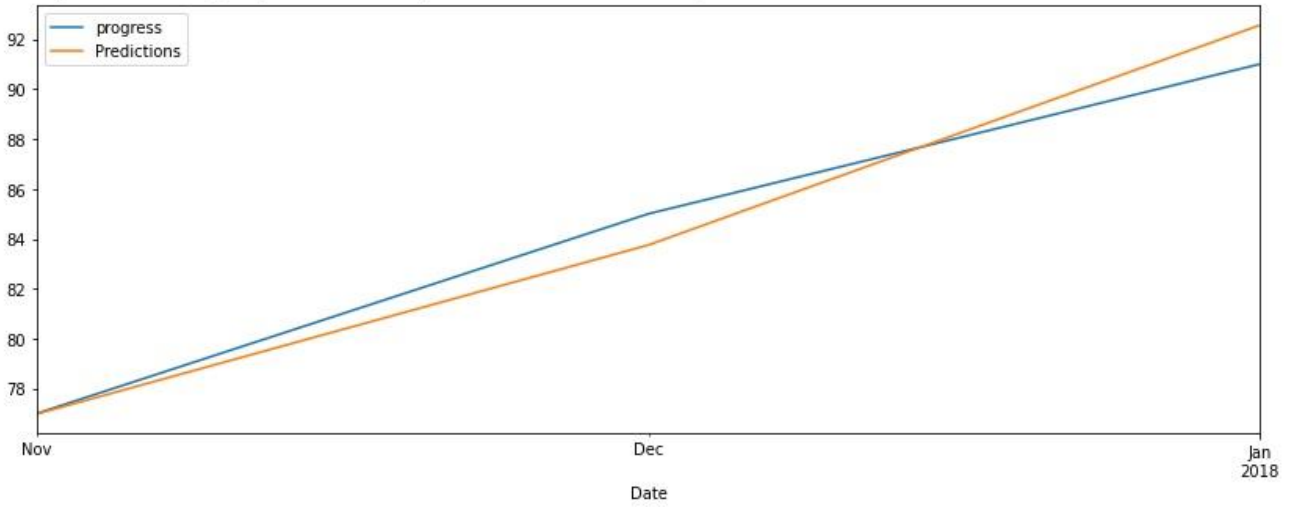


Fig. 1. Comparison of actual versus predicted project progress

$$i_t = \sigma(w_i X_t + U_i h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(w_f X_t + U_f h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(w_o X_t + U_o h_{t-1} + b_o) \quad (3)$$

$$c_t = f_t C_{t-1} + i_t \tanh(w_c X_t + U_c h_{t-1} + b_c) \quad (4)$$

$$h_t = o_t \tanh(C_c) \quad (5)$$

where X_t represents the current input; h_{t-1} denotes the previous hidden state; σ denotes the logistic sigmoid function; $i_t, f_t, o_t, g_t, c_t, c_{t-1}$, and h_t refer to the input gate, forget gate, output gate, input modulation gate, cell state, previous cell state, and current hidden state, respectively. The weight matrices are represented by w_i, w_f, w_o, w_c and the biases are denoted as b_i, b_f, b_o, b_c .

3- Experiments, results, and their discussion

This section describes the simulation experiments to evaluate the effectiveness of the proposed approach.

3- 1- Dataset

The data used in this study has been obtained from the South Extension Project of Tehran Metro Line 6, which includes parking, station construction, ventilation systems, and emergency exits. The project started in October 2016 and

ended in July 2018, lasting for a total of 21 months. In this study, the training dataset consisted of the initial 14 months' data, which accounted for 83 percent of the total data. The remaining three months' data, equivalent to 17 percent of the total, were used as the test dataset to evaluate the model's performance after the training phase. We used the construction project progress as a forecasting variable.

3- 2- Experimental setting

The simulation is carried out using a system running on the Windows 8 operating system, equipped with an Intel(R) Core(TM) i7-2600 CPU operating at 3.40 GHz and 8 GB of RAM. The implementation of the LSTM was done using Python language via Keras library and TensorFlow at the back-end. Google Colab API environment was used for the implementation.

3- 3- Performance evaluation metric

To evaluate the performance of LSTM models, we used the mean square error (MSE) metric as the evaluation criterion. The smaller percentage value of the MSE metric expresses better performance.

4- Results and discussion

The implementation of the proposed approach and the forecasting results are presented in this section. Figure 1 illustrates the actual and forecasted results of construction project progress from LSTM model. This figure compares the actual progress of the project (blue line) and the LSTM model's prediction (orange line) for months 15, 16, and 17. As the figure illustrates, the model accurately forecasts the project's future progress based on its past progress. As the project nears its completion, the pace of progress tends to decline. Additionally, the model's forecasts are based on data

from earlier stages of the project. Consequently, the accuracy of the model's predictions may be slightly affected during the final months, leading to slight deviations from the actual outcomes. However, each project is unique and its progress, especially in the last months, is influenced by various factors that this study does not account for these variables.

5- Conclusions

The ability to forecast the physical progress in the scheduling of construction projects provides project managers with valuable insights to adapt and adjust the initial schedule. By accurately predicting the progress, project managers can identify potential delays or advancements, enabling them to make informed decisions and implement necessary modifications to the schedule. This proactive approach helps enhance project management, optimize resource allocation, and improve overall project efficiency. In this research, we utilized the Long Short-Term Memory (LSTM) model, a type of recurrent neural network architecture designed to capture long-term dependencies, for predicting the project's progress. We used the construction project progress as a forecasting variable. To evaluate the performance of LSTM models, the mean square error (MSE) metric was used as the evaluation criterion. The results indicated that the Long Short-Term Memory (LSTM) method yielded reliable predictions of the project's progress by effectively considering long-term dependencies.

References

- [1] S. Liu, W. Hao, Forecasting the scheduling issues in engineering project management: Applications of deep learning models, *Future Generation Computer Systems*, 123 (2021) 85-93.
- [2] X. Tian, S. Yuan, Genetic algorithm parameters tuning for resource-constrained project scheduling problem, in: *AIP Conference Proceedings*, AIP Publishing, 2018.
- [3] R. Hu, Q. Huang, S. Chang, H. Wang, J. He, The MBPEP: a deep ensemble pruning algorithm providing high quality uncertainty prediction, *Applied Intelligence*, 49(8) (2019) 2942-2955.
- [4] M. Li, W. Chen, Application of BP neural network algorithm in sustainable development of highway construction projects, *Physics Procedia*, 25 (2012) 1212-1217.
- [5] Z. Cen, J. Wang, Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer, *Energy*, 169 (2019) 160-171.
- [6] Z. Liu, L. Li, X. Fang, W. Qi, J. Shen, H. Zhou, Y. Zhang, Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network, *Automation in Construction*, 125 (2021) 103647.

HOW TO CITE THIS ARTICLE

E. Farzad, H. Dehghan Menshadi, M. A. Dashti, *Predicting construction project scheduling issues using LSTM neural network (long-term short-term memory)*, *Amirkabir J. Civil Eng.*, 55(9) (2023) 371-374.

DOI: 10.22060/ceej.2023.21383.7701





پیش بینی مسائل مربوط به زمان بندی پروژه های عمرانی با استفاده از شبکه عصبی LSTM (حافظه طولانی کوتاه مدت)

عرفان فرزاد، هادی دهقان منشادی*، محمدعلی دشتی رحمت آبادی

گروه مهندسی عمران، دانشگاه آزاد اسلامی واحد یزد، یزد، ایران.

تاریخچه داوری:

دریافت: ۱۴۰۱/۰۲/۱۸
بازنگری: ۱۴۰۲/۰۲/۰۹
پذیرش: ۱۴۰۲/۰۵/۲۱
ارائه آنلاین: ۱۴۰۲/۰۵/۲۹

کلمات کلیدی:

مدیریت پروژه
زمان بندی
هوش مصنوعی
حافظه کوتاه مدت بلند مدت
پیش بینی

خلاصه: از آنجایی که هدف نظارت بر پروژه، تصمیم گیری دقیق بوده که می تواند تأثیرات به سزایی بر موفقیت پروژه داشته باشد، پیش بینی ویژگی های پروژه از اهمیت بیشتری برخوردار می گردد. متخصصان بر این باورند که قریب به اتفاق پروژه های عمرانی از تاخیرات رنج می برند. لذا یکی از مهمترین ویژگی پروژه های عمرانی، مسائل مربوط به زمان است. این پژوهش مدلی را پیشنهاد می کند تا بتوان مشکلات زمان بندی پروژه ها را تا حدودی مرتفع نمود. برای این منظور، این پژوهش کاربردهای جدید مدل های پیش بینی حافظه ای کوتاه مدت بلند مدت (LSTM) را که یک معماری از شبکه عصبی بازگشتی است، ارائه می کند. از سوی دیگر به منظور مقایسه و اعتبار سنجی روش LSTM مدل یادگیری دروازه ای (GRU) مورد بررسی قرار می گیرد. متعاقب آن، نتایج پیش بینی مدل های ارائه شده، با داده های یک پروژه واقعی مقایسه و راست آزمایی می شود. در این مطالعه برای پیش بینی از داده های پروژه ای توسعه جنوبی خط شش مترو تهران که در سال ۹۷ به پایان رسیده است، استفاده شده و جهت سنجش دقت از معیار ریشه میانگین مربعات استفاده شده است. مدت زمان در نظر گرفته شده برای مدل سازی ۱۷ ماه بوده به طوری که اطلاعات چهارده ماه ابتدایی به عنوان داده ای آموزش (معادل ۸۳ درصد کل داده ها) و سه ماه آتی پروژه (معادل ۱۷ درصد کل داده ها) به عنوان داده ای آزمون استفاده شده است. متغیر پیش بینی کننده در این پژوهش درصد پیشرفت فیزیکی پروژه می باشد. نتایج نشان می دهد که حافظه ای کوتاه مدت بلند مدت و کاربردهای پیشنهادی مدل، می تواند با دقت خوبی پیشرفت پروژه را پیش بینی کند.

۱- مقدمه

روش های قوی بهینه سازی و بهینه سازی تصادفی بررسی می شود. همچنین جهت حل مشکل زمان بندی پروژه ها چندین الگوریتم مبتنی بر ژنتیک نیز پیشنهاد شده است [۳]. اما مدل های یادگیری عمیق راه حل بهتری را برای پیش بینی مشکلات عدم قطعیت ارائه می دهند [۱]. در تحقیقات انجام شده مدل هایی برای پیش بینی عدم قطعیت با خطا و فاصله عرضی بسیار کم ایجاد شده است [۴].

رویکرد دیگر در مساله پیش بینی، استفاده از شبکه های عصبی و یادگیری ماشین است. شبکه عصبی مصنوعی یک سیستم پردازش اطلاعات است که بر مبنای مقدماتی از ساختار درک مغز انسان بوجود آمده است. این شبکه های عصبی مصنوعی با توانایی نگاشت غیرخطی و توپولوژی خوب، اکنون به طور گسترده در تشخیص الگو، پیش بینی داده ها و ... استفاده می شود [۵]. در میان مدل های موجود در شبکه های عصبی، شبکه های عصبی بازگشتی^۱ قابلیت در نظر گرفتن وابستگی مشاهدات متوالی را

پروژه ای که در زمان مقرر و بدون تاخیر، حتی در عصر حاضر با وجود تکنولوژی و آشنایی با تکنیک های مدیریت پروژه، به پایان برسد، بسیار نادر است. مطابق آمار منتشر شده از سوی دفتر نظارت و ارزیابی طرح ها هر ساله بیش از ۵۰ درصد پروژه های عمرانی کشور با تاخیر مواجه می شوند. به طور میانگین عمر پروژه ها نزدیک به هشت سال است که انتظار می رفت سه ساله به پایان برسند. زمان بندی یک کار مکرر برای کنترل سیستم های مختلف، از جمله پروژه های ساختمانی، قبل از شروع پروژه ها است. عدم قطعیت در برنامه ریزی پروژه باعث تاخیر در زمان بندی و افزایش هزینه انجام کار می گردد [۱]. از سوی دیگر با پیش بینی می توان نظارت بر پروژه را افزایش داده و در زمانهای مورد نیاز اصلاحاتی را در روند پروژه ایجاد کرد. لذا، پیش بینی پروژه، یکی از جنبه های حیاتی کنترل پروژه در زمینه های زمان و هزینه می باشد [۲]. زمان بندی در شرایط عدم قطعیت عمدتاً با استفاده از

1 Recurrent Neural Network (RNN)

* نویسنده عهده دار مکاتبات: sh.manshadi@iau.ac.ir

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.



با استفاده از شبکه عصبی حافظه‌ی کوتاه مدت بلند مدت به پیش‌بینی لایه‌های سرعت باد با فرکانس پایین پرداختند [۱۳]. سانگ و همکاران در سال ۲۰۲۱ مصرف انرژی گرمایشی را برای ساختمان‌های غیرمسکونی با استفاده از مدل ترکیبی CNN-LSTM بررسی کردند [۱۴]. در پژوهشی دیگر جینا و همکاران به ارزیابی آسیب پذیری زلزله برای شبه قاره هند با استفاده از مدل حافظه‌ی کوتاه مدت بلند مدت پرداختند و نشان دادند این مدل می‌تواند با نسبت بسیار خوبی آسیب پذیری مناطق مختلف در زلزله را آشکار کند [۸]. پیش‌بینی وضعیت ترافیک با الگوریتم‌های یادگیری ماشین برای افق‌های زمانی کوتاه و میان‌مدت توسط رسا ایزدی و ابریشمی مورد بررسی قرار گرفت. نتایج تحقیق بیانگر دقت قابل توجه روش LSTM در پیش‌بینی وضعیت ترافیک بود [۱۵].

روش پیشنهادی این پژوهش به منظور پیش‌بینی روند پیشرفت پروژه‌های عمرانی که دارای روابط پیچیده‌ی غیر خطی هستند، به طور کلی مبتنی بر شبکه عصبی بازگشتی می‌باشد. روش LSTM به عنوان زیر مجموعه‌ای از شبکه‌های عصبی بازگشتی دسترسی به تمامی مقادیر گذشته در سری‌های زمانی را مقدور می‌سازد و از این جهت یک ابزار قدرتمند جهت پیش‌بینی رویدادهایی است که دارای روابط پیچیده غیر خطی در طول زمان هستند. لذا این پژوهش برای نخستین بار کاربرد روش LSTM در پیش‌بینی روند پیشرفت پروژه‌های عمرانی بر اساس متغیر درصد پیشرفت فیزیکی را مورد بررسی قرار می‌دهد. مقایسه و ارزیابی نتایج روش LSTM با GRU بخش دیگری از نوآوری پژوهش حاضر در ارتباط با پروژه‌های عمرانی است.

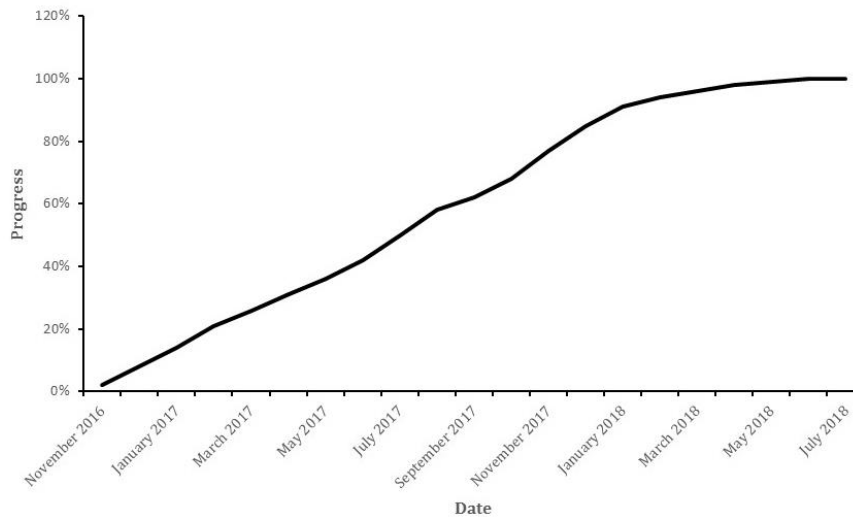
۲- داده

داده‌های استفاده شده در این مطالعه مربوط به پروژه‌ی توسعه جنوبی خط ۶ مترو تهران بوده که شامل پارکینگ شبانه، احداث ایستگاه‌ها، هواکش، خروجی‌های اضطراری و ... می‌باشد. این پروژه، دهم آبان ماه سال ۹۵ آغاز و در تاریخ دهم تیرماه ۹۷ به پایان رسیده است. کل مدت زمان پروژه ۲۱ ماه می‌باشد. براساس نظریه‌های یادگیری ماشین و پژوهش‌های پیشین، به طور کلی بین ۵ تا ۲۰ درصد داده‌های سری زمانی به عنوان داده‌های آزمایش در نظر گرفته می‌شود. این تقسیم بندی معمولاً باعث جلوگیری از برآزش بیش از حد مدل و مقادیر بهینه‌ی پارامترهای مدل خواهد شد [۱۶ و ۱۷]. لذا در این پژوهش از داده‌های ۱۴ ماه ابتدایی (معادل ۸۳ درصد داده‌ها) به عنوان داده‌ی آموزش و سه ماه بعد پروژه (معادل ۱۷ درصد داده‌ها) به عنوان

دارد. به بیانی دیگر، این دسته از مدل‌های شبکه عصبی داده‌های متوالی را به عنوان سری زمانی پردازش می‌کند [۶]. تحقیقات مختلفی با استفاده از شبکه‌های عصبی بازگشتی جهت ترجمه‌ی زبان‌های مختلف به یکدیگر انجام شده است [۶]. نقطه ضعف اصلی این دسته از مدل‌های شبکه عصبی، عدم در نظرگیری وابستگی‌های بلند مدت است.

مدل حافظه‌ی کوتاه مدت بلند مدت^۱ که نوعی مدل شبکه عصبی بازگشتی و مدلی مبتنی بر یادگیری عمیق است، نقیصه یاد شده را برطرف کرده و وابستگی‌های بلند مدت را نیز در نظر می‌گیرد. این مدل در ابتدا برای پردازش زبان طبیعی استفاده شد. پس از آن، حافظه‌ی کوتاه مدت بلند مدت به طور گسترده‌ای برای حل پیش‌بینی موضوع در حوزه‌های مختلف پژوهشی استفاده شد [۷]. لی و همکاران در مقاله‌ای با استفاده از روش واحد بازگشتی دروازه‌ای و الگوریتم ازدحام ذرات به پیش‌بینی عملکرد پارامترهای روسازی پرداختند و نشان دادند این مدل می‌تواند تا ۲۱ درصد در مقایسه با روش‌های سنتی دقت پیش‌بینی را افزایش دهد [۸]. از آنجا که معمولاً عملیات حفاری با TBM^۲ از بازدهی پایین به دلیل عدم تشخیص شرایط زمین شناسی رنج می‌برد. لیو و همکاران در مقاله‌ای تلاش نمودند با استفاده از داده‌های زیاد، مدلی هوشمند مرتبط با زمان را مبتنی بر حافظه‌ی کوتاه مدت بلند مدت جهت پیش‌بینی سنگ شناسی تونل توسعه دهند [۹]. شی و همکاران در مقاله‌ای یک مدل پیش‌بینی چند مرحله‌ای ترکیبی با استفاده از حافظه‌ی کوتاه مدت بلند مدت برای گشتاور سر برش ماشین تونل‌زنی پیشنهاد دادند و در آن نشان دادند که میزان دقت به ۹۸ درصد می‌رسد [۱۰]. با هدف کاهش خطرات ایمنی ناشی از موقعیت قرارگیری ماشین آلات ساختمانی لو و همکاران در مقاله‌ای با استفاده از شبکه عصبی بازگشتی و شبکه عصبی دروازه‌ای^۳ مدلی را پیشنهاد دادند که در آن بیل مکانیکی به عنوان شی در نظر گرفته می‌شود و موفق شدند موقعیت صحیح قرارگیری را تا حدود ۹۰ درصد به صورت مناسب پیش‌بینی کنند [۱۱]. فانگ و همکاران با هدف بهبود پیش‌بینی مصرف برق در ساختمان‌ها از مدل ترکیبی حافظه‌ی کوتاه مدت بلند مدت و شبکه عصبی عمیق^۴ استفاده نمودند و در مقاله‌ای نشان دادند که اکثر مدل‌های یادگیری انتقال چند منبعی می‌توانند عملکرد پیش‌بینی مصرف برق ساختمان هدف را در مقایسه با مدل‌های یادگیری انتقال تک منبع مربوطه افزایش دهند [۱۲]. لیو و همکاران در سال ۲۰۱۸

- 1 Long Short-Term Memory (LSTM)
- 2 Tunnel Boring Machine (TBM)
- 3 Gated Recurrent Unit (GRU)
- 4 Deep Neural Network (DNN)



شکل ۱. منحنی S پروژه توسعه جنوبی مترو تهران

Fig. 1. S curve of south Tehran's metro development

استفاده از داده‌های مربوط به فعالیت ماه‌های قبل، میزان پیشرفت پروژه در ماه‌های بعد پیش‌بینی خواهد شد. داده‌های مورد نظر در این تحقیق، پیش‌پردازش و نرمال سازی شده‌اند. سپس با استفاده از الگوی مناسب مبتنی بر شبکه‌های عصبی بازگشتی سعی بر پیش‌بینی پیشرفت فیزیکی پروژه در ماه‌های بعد شد. طبق شکل ۲، پس از جمع‌آوری و نرمال‌سازی داده‌ها، جهت ارزیابی نهایی روش بکار گرفته شده در این تحقیق، داده‌ها به دو دسته‌ی آزمایش و آموزش تقسیم شدند. آموزش شبکه عصبی با دسته اول انجام شد. داده‌های آزمایش، نمونه‌های جدید و دیده نشده برای ارزیابی روش است.

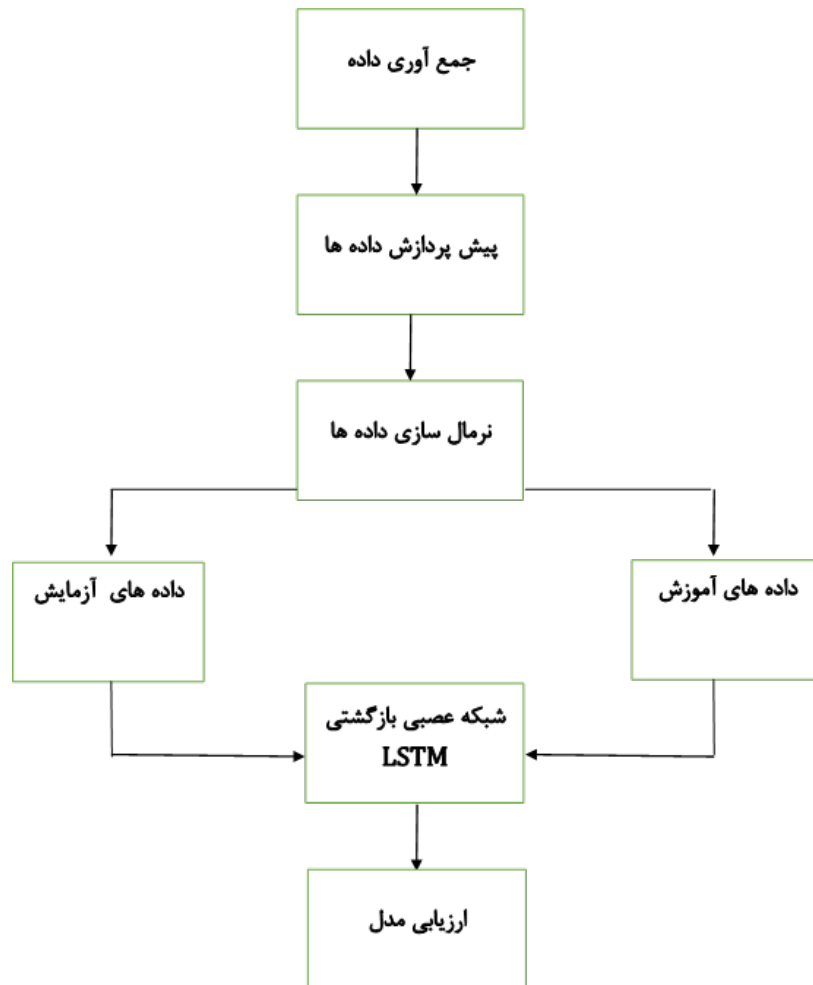
۳-۱- مدل حافظه بلند مدت کوتاه مدت

در این بخش، روش شناسی مرتبط با مدل حافظه‌ی طولانی کوتاه مدت بررسی شده است. مدل شبکه عصبی بازگشتی قابلیت در نظر گرفتن وابستگی مشاهدات متوالی را دارد اما نقطه ضعف اصلی این مدل عدم در نظر گرفتن وابستگی‌های بلند مدت است. مدل حافظه‌ی بلند مدت 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 حافظه‌ی کوتاه مدت بلند مدت، شامل حافظه‌ی کوتاه مدت، حافظه‌ی بلند مدت و

داده‌ی آزمون استفاده شده است. متغیر پیش‌بینی کننده در این پژوهش درصد پیشرفت فیزیکی پروژه می‌باشد. اعداد مشخص شده در نمودار شکل ۱ پیشرفت فیزیکی تجمعی مربوط به انتهای هر ماه از پروژه است. به این صورت که پیشرفت پروژه در انتهای هر ماه مشخص شده با پیشرفت ماه‌های قبل از خود جمع شده است. با توجه به اطلاعات موجود، منحنی S متناظر با پروژه‌ی مورد بررسی مطابق شکل ۱ می‌باشد. منحنی S در واقع یک نمایش گرافیکی از هزینه‌های تجمعی، ساعت نیروی کار، درصد پیشرفت کار یا مقادیر دیگر، بر اساس زمان در یک پروژه می‌باشد [۱۸]. شکل این منحنی، معمولاً با یک شیب کم در ابتدا و نزدیک به انتهای پروژه و شیب بزرگتر در وسط، نشان می‌دهد که پیشرفت در شروع و بسته شدن منابع آهسته، اما زمانی که کار اصلی انجام می‌شود سریعتر است. این نمودار می‌تواند پیشرفت کلی پروژه را در اعداد واحد نشان دهد. کارهای تحقیقاتی اولیه، استفاده از نمودارهای تجمعی هزینه در مقابل زمان و پیشرفت تجمعی را در مقابل زمان پیشنهاد کردند، که نقش پررنگ آنها در تسهیل مدیریت پروژه برای مدیران بر کسی پوشیده نیست [۱۹].

۳- روش شناسی

مراحل انجام تحقیق در شکل ۲ نشان داده شده است. در این مطالعه با



شکل ۲. روند آموزش در مدل های LSTM

Fig. 2. Training process in LSTM models

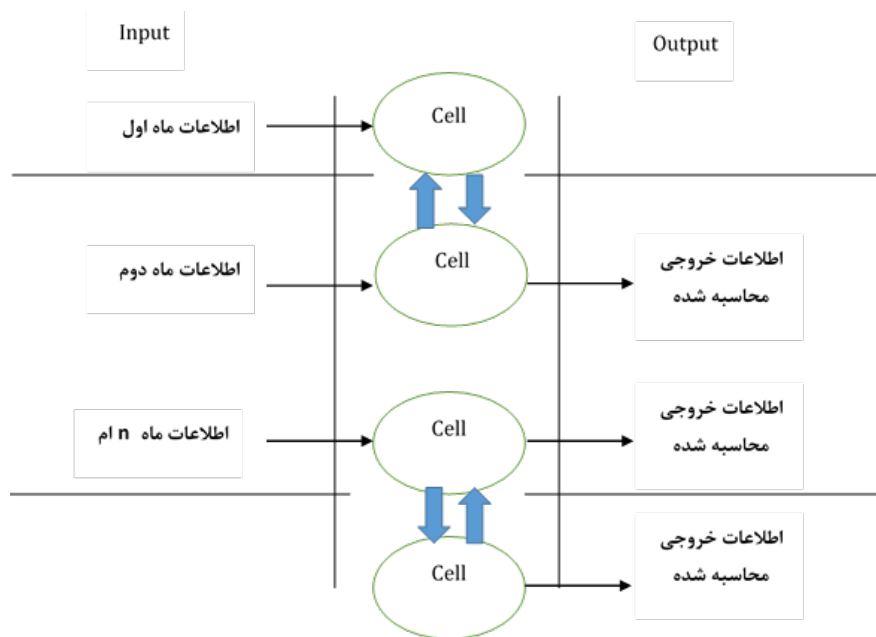
$$i_t = \sigma(w_i X_t + U_i h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(w_f X_t + U_f h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(w_o X_t + U_o h_{t-1} + b_o) \quad (3)$$

$$c_t = f_t C_{t-1} + i_t \tanh(w_c X_t + U_c h_{t-1} + b_c) \quad (4)$$

- (۱) برخی مشاهدات موجود در مجموعه داده آموزش می شود. ساختار این مدل شامل چهار دروازه می شود. ورودی بلند مدت وارد دروازه‌ی فراموشی می شود و این دروازه تصمیم می گیرد کدام یک از ورودی‌های نامرتب حذف شود (رابطه‌ی ۱). ورودی کوتاه مدت مدل و برخی مشاهدات داده آموزش وارد دروازه آموزش می شوند و این دروازه تصمیم می گیرد کدام ورودی یاد گرفته شود (رابطه‌ی ۲). اطلاعات عبور داده شده شامل خروجی‌های دروازه‌های فراموشی و آموزش، وارد دروازه یادآوری می شوند. خروجی این دروازه حافظه‌های بلند مدت جدید است. (رابطه‌ی ۳). در نهایت دروازه خروجی، حافظه‌های کوتاه مدت را نیز به روز رسانی می کند (رابطه‌ی ۴).



شکل ۳. ساختار آموزش در مدل های LSTM

Fig. 3. Training structure in LSTM models

دروازه از دو دروازه به روز رسانی و باز نشانی استفاده می کند که در نتیجه از نظر سرعت، بهبود نسبتاً خوبی خواهد داشت [۲۱]. دروازه بروز رسانی تصمیم می گیرد تا چه مقداری، واحد عصبی به روز رسانی شود و دروازه باز نشانی تصمیم می گیرد چه مقدار از واحد عصبی قبلی را فراموش کند. آخرین حالت h_{t-1} و ورودی x_t و ماتریس های وزن برای اتصالات دروازه های هستند، بردار بایاس است که در حین آموزش مقدار دهی می شود. ورودی مدل به صورت دنباله ای از X بیان می شود. بنابراین x_t بردار ورودی و h_t بردار خروجی است. کل فرآیندی که در یک مرحله ای زمانی در یک سلول انجام می شود، محاسبه جمع کردن حالت موجود و همچنین محاسبه حالت بروز رسانی است [۲۲].

$$h_t = ((1 - z_t)h_{t-1} + z_t h_t) \quad (6)$$

$$h_t = \tanh(w_h x_t + U_h (r_t h_{t-1}) + b_h) \quad (7)$$

در همان زمان، بردار دروازه به روز رسانی z_t تصمیم می گیرد تا چه میزان از محتوای واحد فعال را به روز رسانی کند، بردار z_t به صورت زیر محاسبه

$$h_t = o_t \tanh(C_c) \quad (5)$$

در روابط بالا پارامترهای c, o, f و h به ترتیب معرف ورودی، فراموشی، خروجی، حافظه و حالت سلول هستند. در معادله ای بالا σ تابع سیگموید، w_f و b_f نیز به ترتیب بیانگر b_f وزن و بایاس در دروازه فراموشی هستند. در طول ساخت شبکه ی LSTM، در مراحل ابتدایی مدل، نیازی به حذف اطلاعات نیست و این کار توسط تابع سیگموید انجام می شود که خروجی آخرین واحد را در زمان $(t-1)$ و همچنین ورودی جدید را در زمان حال (t) دریافت می کند. در معادلات بالا C_t و C_{t-1} حالت های سلول در زمان $t-1$ و t می باشند.

۳-۲- شبکه عصبی بازگشتی GRU

شبکه عصبی بازگشتی GRU توسط چو و همکاران در سال ۲۰۱۴ پیشنهاد شده است [۲۰]. این شبکه، نوع خاصی از شبکه های عصبی بازگشتی هستند که با تنظیم انواع مختلف واحد چرخه، مشکل وابستگی به انواع مقیاس های زمانی را حل کرده است. شبکه عصبی بازگشتی GRU ماند ساز و کار LSTM عمل می کند و فرق آن این است که به جای سه

جدول ۱. مقایسه LSTM و GRU

Table 1. Comparison of GRU and LSTM

معیار ارزیابی	LSTM	GRU
عرضه حافظه کنترل شده	مقدار حافظه عرضه شده توسط واحد های دیگر شبکه بوسيله دروازه خروجی کنترل می گردد	کل حافظه در شبکه عرضه می گردد
محاسبه حافظه جدید	هیچ کنترل جداگانه ای برای میزان جریان اطلاعات از مرحله زمان قبلی وجود ندارد	جریان اطلاعات را از فعال سازی قبلی کنترل می کند
پیچیدگی نسبت به عملکرد	با یک دروازه اضافی پیچیدگی بالاتری دارد	پارامترهای کمتری دارد و آموزش سریعتر
تعداد دروازه	دارای سه دروازه ورودی خروجی و فراموشی است	دارای دو دروازه به روز رسانی و باز نشانی است

می شود:

۳-۳- مقایسه شبکه عصبی بازگشتی GRU و شبکه عصبی LSTM

تفاوت بین این دو شبکه در جدول (۱) ارائه شده است [۲۵]. اگرچه این دو نوع شبکه با توجه به آزمایش های اولیه خود در ترجمه ماشین مشابه یکدیگر رفتار می کنند ولی معلوم نیست که آیا این امر به کاربردهایی غیر از ترجمه ماشین نیز قابل تعمیم هست یا خیر.

$$z_t = \sigma(w_z X_t + U_z h_{t-1} + b_z) \quad (8)$$

بردار z_t, r_t دروازه های باز نشانی است. هنگامی که دروازه بسته است، اجازه می دهد که سلول، گذشته را فراموش کند. پس در نتیجه هر کدام از بردارهای Z و r یک بردار نرمال شده هستند که مقادیر آن ها بین صفر و یک است. مقدار صفر یعنی هیچ مقداری عبور نمی دهد و مقدار یک یعنی کل مقدار را عبور می دهد. بردار طبق رابطه زیر محاسبه می شود.

$$r_t = \sigma(w_r X_t + U_r h_{t-1} + b_r) \quad (9)$$

دروازه های به روز رسانی کنترل می کند که چه میزان از حالت گذشته مهم هستند. واحدهایی با وابستگی های بلند مدت با دروازه به روز رسانی فعال خواهد شد [۲۳]. شبکه عصبی بازگشتی GRU می تواند از واحد حافظه در یک شبکه، برای مقابله با هر توان داده و ورودی استفاده کند. بنابراین توانایی یادگیری سری های زمانی را دارد. GRU نه تنها می تواند سری های زمانی طولانی را مرور کند، بلکه به طور خودکار تعادل زمان مطلوب برای پیش بینی را نیز تعیین می کند. در سال های اخیر، GRU با موفقیت در شناسایی دست خط، شناسایی حرکت انسان و کنترل ربات و ... مورد استفاده قرار گرفته است [۲۴].

۴- نتایج

در این بخش آزمایش های شبیه سازی برای ارزیابی اثر بخشی و رویکرد پیشنهادی مورد بررسی قرار می گیرد.

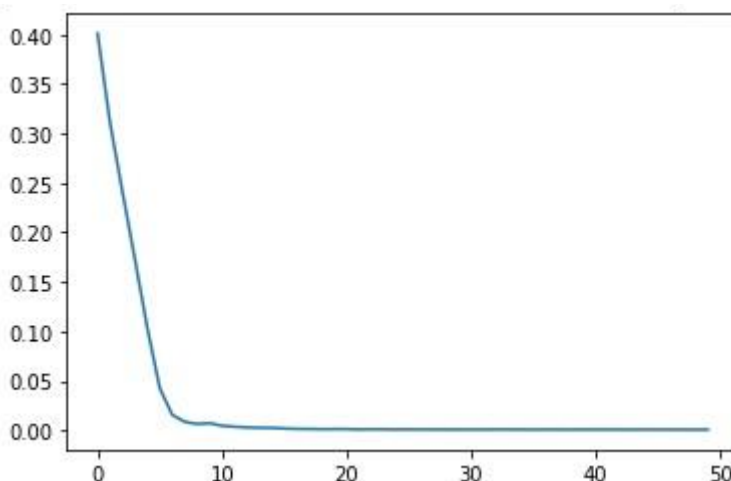
۴-۱- تنظیمات مدل

رویکرد پیشنهادی، مبتنی بر شبکه های عصبی بازگشتی و متشکل از سه لایه که هاپر پارامترهای آن در جدول ۲ آمده است. شبیه سازی در پردازنده Intel(R) Core(TM) i7-2600 CPU @3.40 GHz و با رم ۸ گیگا بایت انجام شده است. سیستم عامل مورد استفاده ویندوز ۸ و برنامه نویسی در بستر زبان برنامه نویسی پایتون انجام شده و از کتابخانه های Keras و Tensorflow استفاده گردیده است. در گام نخست عملیات پیش پردازش داده با اضافه کردن متغیر پیش بینی کننده انجام شده است. در گام دوم یافتن مقادیر بهینه پارامترهای مدل با توجه به حداقل رساندن تفاوت بین خروجی پیش بینی شده و خروجی واقعی تنظیم می شوند. این کار از طریق فرآیندی به نام انتشار پس انداز انجام شد، جایی که خطا از طریق شبکه منتشر می شود تا پارامترها بر اساس سهم آنها در خطا تنظیم

جدول ۲. هایپر پارامتر های مدل

Table 2. Model hyperparameters

نام دیتا ست	تعداد پارامترهای آزمایش	تعداد نورون های لایه LSTM	تعداد نورون های لایه DENCE	تعداد EPOCH
پیشرفت فیزیکی پروژه توسعه جنوبی مترو تهران	۴۰۹۰۱	۱۰۰	۱	۵۰



شکل ۴. نحوه کاهش میانگین مربعات خطا در پی هر بار تکرار

Fig. 4. Variations of mean squared error (MSE) versus epoch

۴-۲- معیار ارزیابی عملکرد

پیشنهاد می‌شود از میانگین مربعات خطا (MSE) برای ارزیابی مدل LSTM استفاده شود. هر چه میزان این خطا کمتر باشد مدل عملکرد بهتری خواهد داشت. میانگین مربعات خطا در رابطه‌ی (۱۰) بیان شده است.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_{pred})^2 \quad (10)$$

MSE یک معیار رایج برای پیش بینی خطاها بوده و نشان دهنده‌ی مقدار متوسط خطا بین مقدار پیش بینی شده و مقدار واقعی است.

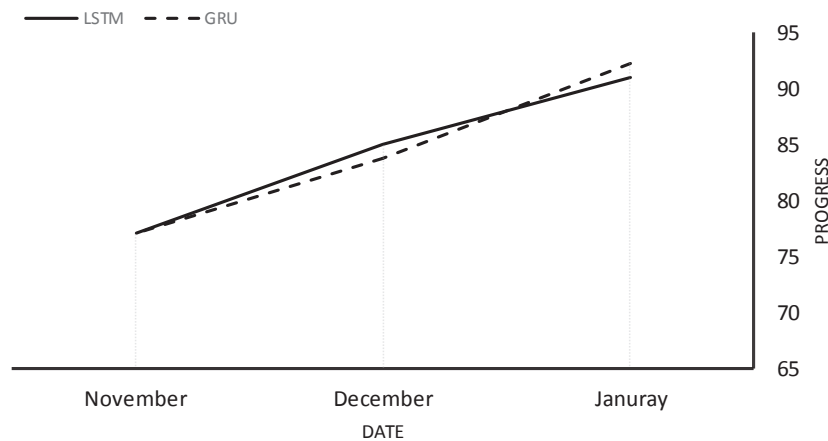
شوند. در این مرحله، تعیین تعداد لایه های LSTM و لایه‌های متراکم پس از آزمایش و اصلاح به دست آمد. به طوری که برای لایه LSTM از تعداد ۱۰۰ نورون استفاده شده است. لایه LSTM در این مدل هوش مصنوعی به یک لایه شبکه عصبی بازگشتی اشاره دارد که شامل ۱۰۰ واحد حافظه کوتاه مدت است. لایه بعدی در این مدل، لایه متراکم (Dense) است که نشان دهنده یک لایه کاملاً متصل است که در آن هر گره در لایه به هر گره در لایه قبلی متصل است. برای این لایه از یک نورون استفاده شد.

در شکل ۴ نحوه کاهش خطا در پی هر بار تکرار مدل، آمده است و همان گونه که مشاهده می‌شود تقریباً پس از تکرار دهم میزان خطا تقریباً ثابت می‌شود.

جدول ۲. پارامترهای LSTM و GRU

Table 2. LSTM and GRU parameters

نام مدل	کل پارامترها	پارامترهای آموزش	تعداد نورون‌ها	تعداد نورون های لایه Dense	ریشه میانگین مربعات
LSTM	۴۰۹۰۱	۴۰۹۰۱	۱۰۰	۱	۱/۲۶۹۲
GRU	۳۱۰۰۱	۳۱۰۰۱	۱۰۰	۱	۱/۲۳۷۴



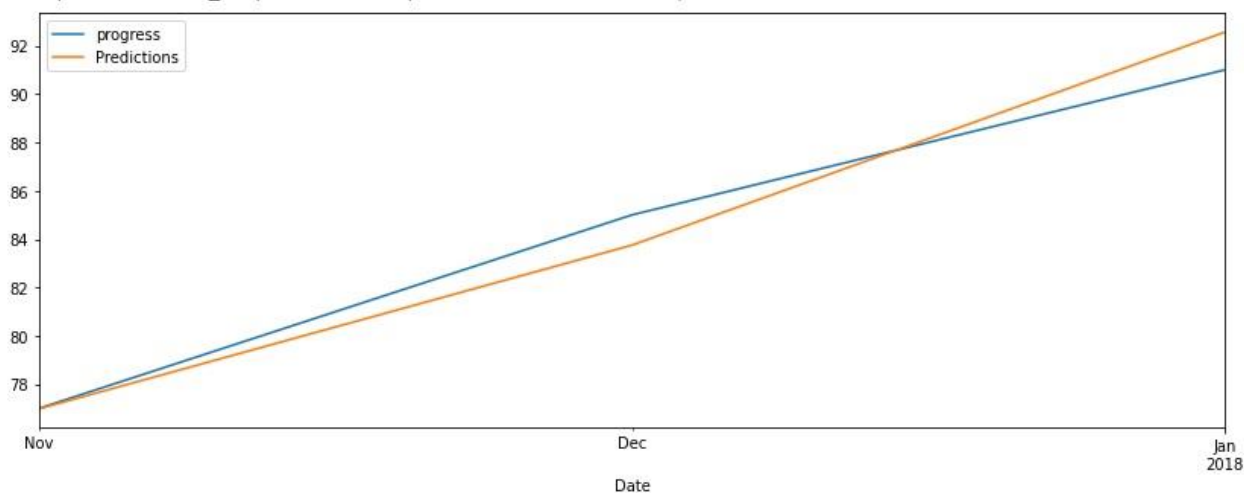
شکل ۵. مقایسه پیشرفت پیش بینی شده به روش LSTM و GRU

Fig. 5. Comparison of forecasted progress by LSTM and GRU method

به طور گسترده‌ای برای پیش بینی دقیق مسائل سری زمانی استفاده شده است. در این تحقیق GRU به کار گرفته شده تا نتایج مدل LSTM را تأیید کند و مشکل ناپدید شدن مسائل گرادیان در مدل های استاندارد RNN را حل کند. در جدول ۳ پارامترهای و نتایج بدست آمده از دو روش LSTM و GRU با هم مقایسه شده‌اند. از سوی دیگر نتایج مربوط به پیش بینی سه ماه انتهایی پروژه‌ی مورد نظر توسط هر دو روش در شکل ۵ نشان داده شده است. همان گونه که مشاهده می‌شود، روند پیش بینی هر دو روش مشابه و نزدیک به هم می‌باشد.

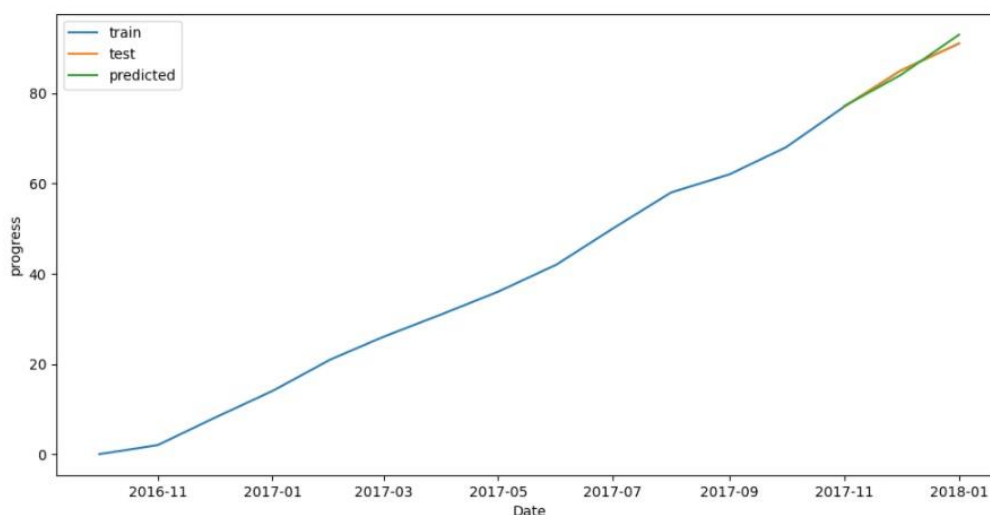
۴-۳- شبکه مکرر دروازه ای (GRU)

در این تحقیق از مدل مکرر دروازه ای (GRU) برای اعتبار سنجی روش LSTM استفاده شده است. معماری GRU در سال ۲۰۱۴ توسط Cho و همکاران معرفی شد. این معماری به منظور برطرف سازی کاستی‌های شبکه عصبی بازگشتی سنتی نظیر مشکل محو شدگی گرادیان و همچنین کاهش سربار موجود در معماری LSTM ارائه شده است. GRU عموماً به عنوان نسخه ای تغییر یافته از LSTM در نظر گرفته می‌شود چرا که هر دو این معماری‌ها از طراحی مشابهی بهره می‌برند و در بعضی از موارد بصورت یکسان نتایج عالی بدست می‌دهند. مدل GRU



شکل ۶. مقایسه پیشرفت پیش بینی شده و پیشرفت واقعی

Fig. 6. Comparison of actual versus predicted project progress



شکل ۷. روند واقعی پروژه از ابتدای پروژه تا ماه شانزدهم و مقایسه آن با روند پیش بینی شده

Fig. 7. Comparing the projected and actual progress of the project from the beginning of the project to the 16th month

رنگ متعلق به پیشرفت پیش بینی شده با مدل LSTM در ماه‌های مذکور می‌باشد. همان‌گونه که در شکل مشخص می‌باشد مدل توانسته است با دقت بسیار خوبی پیشرفت آتی پروژه را با توجه به نحوه پیشرفت پروژه در ماه‌های قبل بسنجد. در شکل ۷ نیز می‌توان روند پیشرفت پروژه در ماه‌های قبل، از شروع پروژه تا انتهای ماه ۱۷ و پیش بینی مدل از سه ماه آخر را در یک قاب مشاهده نمود. در شکل ۷ خط آبی متعلق به روند واقعی پروژه از ابتدای پروژه تا ماه چهاردهم بوده که به عنوان داده‌ی آموزش از آن استفاده

۴-۴- بحث و گفتگو

همان‌گونه که بیان شد، مدت زمان بررسی شده ۱۷ ماه می‌باشد، که پیشرفت فیزیکی مربوط به ۱۴ ماه ابتدایی پروژه به عنوان داده‌ی آموزش (معادل ۸۳ درصد کل داده‌ها) و ماه‌های ۱۵ و ۱۶ و ۱۷ (معادل ۱۷ درصد به عنوان داده‌ی آزمایش در نظر گرفته می‌شود. نتایج مربوط به مدل‌سازی در شکل ۶ نشان داده شده است. در این شکل خط آبی رنگ متعلق به پیشرفت واقعی پروژه مطابق با اعلام دفتر فنی در ماه‌های ۱۵، ۱۶ و ۱۷ و خط نارنجی

شده است. خط نارنجی روند واقعی پروژه و خط سبز رنگ پیش بینی مدل از روند پروژه در سه ماه آخر است.

۵- محدودیت ها

از آنجا که سرعت کار در پروژه‌ها در ماه‌های منتهی به پایان پروژه کاهش می‌یابد و با عنایت به اینکه پیش بینی انجام شده در مدل با توجه به روند پروژه در ماه‌های قبل انجام می‌شود، نتایج پیش بینی مدل در ماه‌های انتهایی ممکن است با واقعیت کمی تفاوت داشته باشد. از سوی دیگر با توجه به منحصر به فرد بودن هر پروژه، پیشرفت یک پروژه مخصوصاً در ماه‌های انتهایی آن وابسته به عوامل مختلفی است که پیش بینی انجام شده در این مقاله این متغیرها را در نظر نمی‌گیرد. از جمله متغیرهای تاثیر گذار بر روند پیشرفت پروژه در ماه‌های انتهایی محدودیت منابع مالی و نیروی انسانی، درصد تغییر مبلغ قرارداد، تغییرات در قیمت مصالح، تغییرات شرایط آب و هوایی و تغییرات طراحی در هر مرحله از اجرای پروژه می‌باشد.

همچنین دقت مدل استفاده شده وابستگی زیادی به مقدار داده‌های آموزش دارد لذا هر چه مقدار داده‌های آموزش بیشتر باشد، بدون شک نتایج پیش بینی شده دقیق‌تر خواهند بود. به عنوان نکته آخر بایستی به این مسئله اشاره کرد که فرآیند زمان بندی مدیریت پروژه در کشور بیشتر جنبه‌ای نمایشی دارد و صرفاً در ابتدای پروژه مورد توجه قرار می‌گیرد و در طول پروژه به دلیل تاخیرات متعددی که اتفاق می‌افتد پیمانکاران رغبتی برای بروز رسانی مستمر برنامه زمان بندی ندارند که این می‌تواند مشکلات عدیده‌ای را در اتمام پروژه متوجه سازمان‌ها کند.

۶- نتیجه گیری

پیش بینی پیشرفت فیزیکی در زمان بندی پروژه‌های عمرانی می‌تواند کمک شایانی به مدیران پروژه در جهت اصلاح برنامه زمان بندی اولیه کند. در این شرایط یک مدیر پروژه می‌تواند از تاخیرات احتمالی پروژه مطلع شده و نسبت به اصلاح برنامه پروژه اقدام کند.

بدین منظور، در این پژوهش مدل حافظه‌ی کوتاه مدت بلند مدت (LSTM) که یک معماری از شبکه عصبی بازگشتی بوده و قادر به در نظر گرفتن وابستگی‌های بلند مدت است، جهت پیش بینی روند پیشرفت پروژه به کار گرفته شد. متغیر پیش بینی کننده در این پژوهش درصد پیشرفت فیزیکی پروژه به صورت ماهیانه بوده که برگرفته از داده‌های پروژه‌ی توسعه جنوبی خط شش مترو تهران می‌باشد. در گام اول جهت یادگیری مدل،

داده‌های آموزش متناسب با پیشرفت فیزیکی تجمیعی مربوط به چهارده ماه ابتدای پروژه در نظر گرفته شد و از تابع فعال سازی Relu در مدل استفاده گردید. در مرحله بعد، از مدل آموزش دیده جهت پیش بینی پیشرفت فیزیکی پروژه در ماه‌های آتی استفاده شد. سپس مقادیر پیش بینی شده با مقادیر واقعی پروژه مقایسه گردید و میزان خطای آن با استفاده از شاخص میانگین مربعات خطا سنجیده شد. از سوی دیگر به منظور مقایسه و اعتبار سنجی روش LSTM مدل یادگیری دروازه‌ای (GRU) مورد بررسی قرار گرفت. نتایج مدل سازی نشان داد که روش حافظه‌ی کوتاه مدت بلند مدت با توجه به در نظر گرفتن وابستگی‌های بلند مدت قادر است پیش بینی قابل قبولی از روند پیشرفت پروژه ارائه دهد. نتایج پیش بینی مدل در ماه‌های انتهایی ممکن است با واقعیت کمی تفاوت داشته باشد. دلیل این تفاوت را می‌توان در متغیرهایی که در ماه‌های پایانی بر سرعت اجرای پروژه تاثیر گذار است، جستجو کرد. در میان این عوامل می‌توان به محدودیت‌های منابع، تغییرات طراحی در مراحل پایانی پروژه و شرایط آب و هوایی اشاره کرد.

منابع

- [1] S. Liu, W. Hao, Forecasting the scheduling issues in engineering project management: Applications of deep learning models, *Future Generation Computer Systems*, 123 (2021) 85-93.
- [2] M. Khoshgoftar, A.H.A. Bakar, O. Osman, Causes of delays in Iranian construction projects, *International Journal of Construction Management*, 10(2) (2010) 53-69.
- [3] X. Tian, S. Yuan, Genetic algorithm parameters tuning for resource-constrained project scheduling problem, in: *AIP Conference Proceedings*, AIP Publishing, 2018.
- [4] R. Hu, Q. Huang, S. Chang, H. Wang, J. He, The MBPEP: a deep ensemble pruning algorithm providing high quality uncertainty prediction, *Applied Intelligence*, 49(8) (2019) 2942-2955.
- [5] M. Li, W. Chen, Application of BP neural network algorithm in sustainable development of highway construction projects, *Physics Procedia*, 25 (2012) 1212-1217.
- [6] Z. Cen, J. Wang, Crude oil price prediction model with long short term memory deep learning based on prior

- Journal of Civil Engineering, 54(4) (2022) 1503-1520.
- [16] J. Wang, B. Ashuri, Predicting ENR'S construction cost index using the modified K nearest neighbors (KNN) algorithm, in: Construction research congress 2016, 2016, pp. 2502-2509.
- [17] J. Wang, B. Ashuri, Predicting ENR construction cost index using machine-learning algorithms, International Journal of Construction Education and Research, 13(1) (2017) 47-63.
- [18] H.P.M. Institute, A guide to the project management body of knowledge (PMBOK® guide), in, Project Management Institute, 2017.
- [19] W. Lu, Y. Peng, X. Chen, M. Skitmore, X. Zhang, The S-curve for forecasting waste generation in construction projects, Waste Management, 56 (2016) 23-34.
- [20] K. Cho, B. Van Merriënboer, D. Bahdanau, Y. Bengio, On the properties of neural machine translation: Encoder-decoder approaches, arXiv preprint arXiv:1409.1259, (2014).
- [21] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, arXiv preprint arXiv:1412.3555, (2014).
- [22] B. Liu, C. Fu, A. Bielefield, Y.Q. Liu, Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network, Energies, 10(10) (2017) 1453.
- [23] Z. Wu, S. King, Investigating gated recurrent networks for speech synthesis, in: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2016, pp. 5140-5144.
- [24] X. Zhang, F. Shen, J. Zhao, G. Yang, Time series forecasting using GRU neural network with multi-lag after decomposition, in: Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part V 24, Springer, 2017, pp. 523-532.
- [25] D. Shah, H. Isah, F. Zulkernine, Stock market analysis: A review and taxonomy of prediction techniques, International Journal of Financial Studies, 7(2) (2019) 26.
- knowledge data transfer, Energy, 169 (2019) 160-171.
- [7] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, arXiv preprint arXiv:1409.0473, (2014).
- [8] J. Li, Z. Zhang, X. Wang, W. Yan, Intelligent decision-making model in preventive maintenance of asphalt pavement based on PSO-GRU neural network, Advanced Engineering Informatics, 51 (2022) 101525.
- [9] Z. Liu, L. Li, X. Fang, W. Qi, J. Shen, H. Zhou, Y. Zhang, Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network, Automation in Construction, 125 (2021) 103647.
- [10] G. Shi, C. Qin, J. Tao, C. Liu, A VMD-EWT-LSTM-based multi-step prediction approach for shield tunneling machine cutterhead torque, Knowledge-Based Systems, 228 (2021) 107213.
- [11] H. Luo, M. Wang, P.K.-Y. Wong, J. Tang, J.C. Cheng, Construction machine pose prediction considering historical motions and activity attributes using gated recurrent unit (GRU), Automation in Construction, 121 (2021) 103444.
- [12] X. Fang, G. Gong, G. Li, L. Chun, P. Peng, W. Li, A general multi-source ensemble transfer learning framework integrate of LSTM-DANN and similarity metric for building energy prediction, Energy and Buildings, 252 (2021) 111435.
- [13] H. Liu, X.-w. Mi, Y.-f. Li, Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network, Energy conversion and management, 156 (2018) 498-514.
- [14] J. Song, L. Zhang, G. Xue, Y. Ma, S. Gao, Q. Jiang, Predicting hourly heating load in a district heating system based on a hybrid CNN-LSTM model, Energy and Buildings, 243 (2021) 110998.
- [15] A. Rasaizadi, S.E. Seyedabrishami, Traffic state prediction with machine learning algorithms for short-term and mid-term prediction time horizons, Amirkabir

چگونه به این مقاله ارجاع دهیم

E. Farzad, H. Dehghan Menshadi, M. A. Dashti, Predicting construction project scheduling issues using LSTM neural network (long-term short-term memory), Amirkabir J. Civil Eng., 55(9) (2023) 1753-1764.

DOI: 10.22060/ceej.2023.21383.7701

