



Imam Khomeini International University

Vol. 9, No. 3, Autumn 2024



نشریه مهندسی منابع معدنی

Journal of Mineral Resources Engineering  
(JMRE)

## Research Paper

# Prediction of the Maximum Surface Settlement Induced by Urban Area Tunneling Using Multiple Linear Regression Model and Random Forest Algorithm

Dehghani H.<sup>1\*</sup>, Matinpour F.<sup>2</sup>, Mohammadi Bolbanabad Sh.<sup>3</sup>, Monjezi M.<sup>4</sup>

1- Associate Professor, Mining Engineering Faculty, Hamedan University of Technology, Hamedan, Iran

2- Ph.D Student, School of Mining Engineering, University College of Engineering, University of Tehran, Tehran, Iran

3- Ph.D Student, Faculty of Engineering, Tarbiat Modares University, Tehran, Iran

4- Professor, Faculty of Engineering, Tarbiat Modares University, Tehran, Iran

Received: 27 Jul. 2023

Accepted: 04 Nov. 2023

**Abstract:** The occurrence of settlement induced by tunneling operations in urban environments is an inevitable phenomenon. The risk of damage to nearby infrastructures and surface structures can be greatly reduced by predicting and controlling this event. This paper uses multiple linear regression (MLR) model and random forest (RF) algorithm to predict the maximum surface settlement ( $S_{max}$ ) due to shallow tunnel excavation. Nine input parameters, including the distance of the tunnel center from the ground surface ( $H$ ), height of the underground water level above the tunnel ( $W.T$ ), tunnel diameter ( $D$ ), elastic modulus of soil ( $E$ ), undrained shear strength of soil ( $C_u$ ), earth pressure coefficient ( $K_0$ ), unit weight of soil ( $\gamma$ ), gap parameter ( $g$ ), and stability number ( $N$ ) were selected from 24 data sets related to 14 tunneling projects. The MLR and RF techniques were then implemented for predicting  $S_{max}$ . Three performance indicators of coefficient of determination ( $R^2$ ), root mean square error (RMSE), and mean absolute error (MAE) were employed for the training and test phases to evaluate the efficiency of the suggested models. The coefficient of determination values for MLR and RF for training data were 0.814 and 0.957, respectively, while for test data were 0.793 and 0.96, indicating that the RF approach is more efficient than MLR. Moreover, the findings reveal that the RF algorithm exhibits lower RMSE and MAE values in both the training and testing phases compared to the MLR method. This suggests that the RF algorithm exhibits reduced error and higher reliability and accuracy when compared to the MLR model. Also, the performance study demonstrates that among the input parameters, the gap parameter ( $g$ ) and the undrained shear strength of soil ( $C_u$ ) have the greatest and least influence on  $S_{max}$ , respectively.

**Keywords:** Tunneling operations, Maximum surface settlement ( $S_{max}$ ), Multiple linear regression (MLR) model, Random forest (RF) algorithm.

### How to cite this article

Dehghani, H., Matinpour, F., Mohammadi Bolbanabad, Sh., and Monjezi, M. (2024). "Prediction of the maximum surface settlement induced by urban area tunneling using multiple linear regression model and random forest algorithm". Journal of Mineral Resources Engineering, 9(3): 77-94.

DOI: 10.30479/JMRE.2023.19111.1654

\*Corresponding Author Email: [dehghani@hut.ac.ir](mailto:dehghani@hut.ac.ir)

### COPYRIGHTS



©2024 by the authors. Published by Imam Khomeini International University.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/>)

## INTRODUCTION

Ground settlement resulting from tunnelling operations is an inevitable occurrence that poses a significant concern in urban underground projects. The ability to predict and control this phenomenon is essential to reduce possible damage to surface structures and nearby infrastructures. Due to the importance of the settlement and its direct impact on the project costs, so far many studies have been conducted based on different methods as well as presenting relationships to predict it. Various methods, including empirical, analytical, numerical, and artificial intelligence approaches are available. In practice, ground movement prediction is mainly based on empirical relationships that do not consider the intrinsic parameters of the soil [1-4]. Although analytical methods are useful tools for predicting  $S_{\max}$  and present a closed-form solution, they cannot include all the parameters affecting the settlement [5-8]. In numerical methods, mathematical equations are solved using the finite element [9] or discrete element [10] method to predict the movement of the overburden layers and, finally, the settlement of the ground surface is calculated.

In contrast to the three aforesaid methods, artificial intelligence methods provide the possibility of analysing a larger amount of data by integrating complex nonlinear combinations to predict the target parameter. In this paper, using the two methods of multiple linear regression (MLR) model and random forest (RF) algorithm, the amount of  $S_{\max}$  resulting from tunneling operations is predicted. The results showed that RF algorithm can predict  $S_{\max}$  more accurately than MLR model.

## METHODS

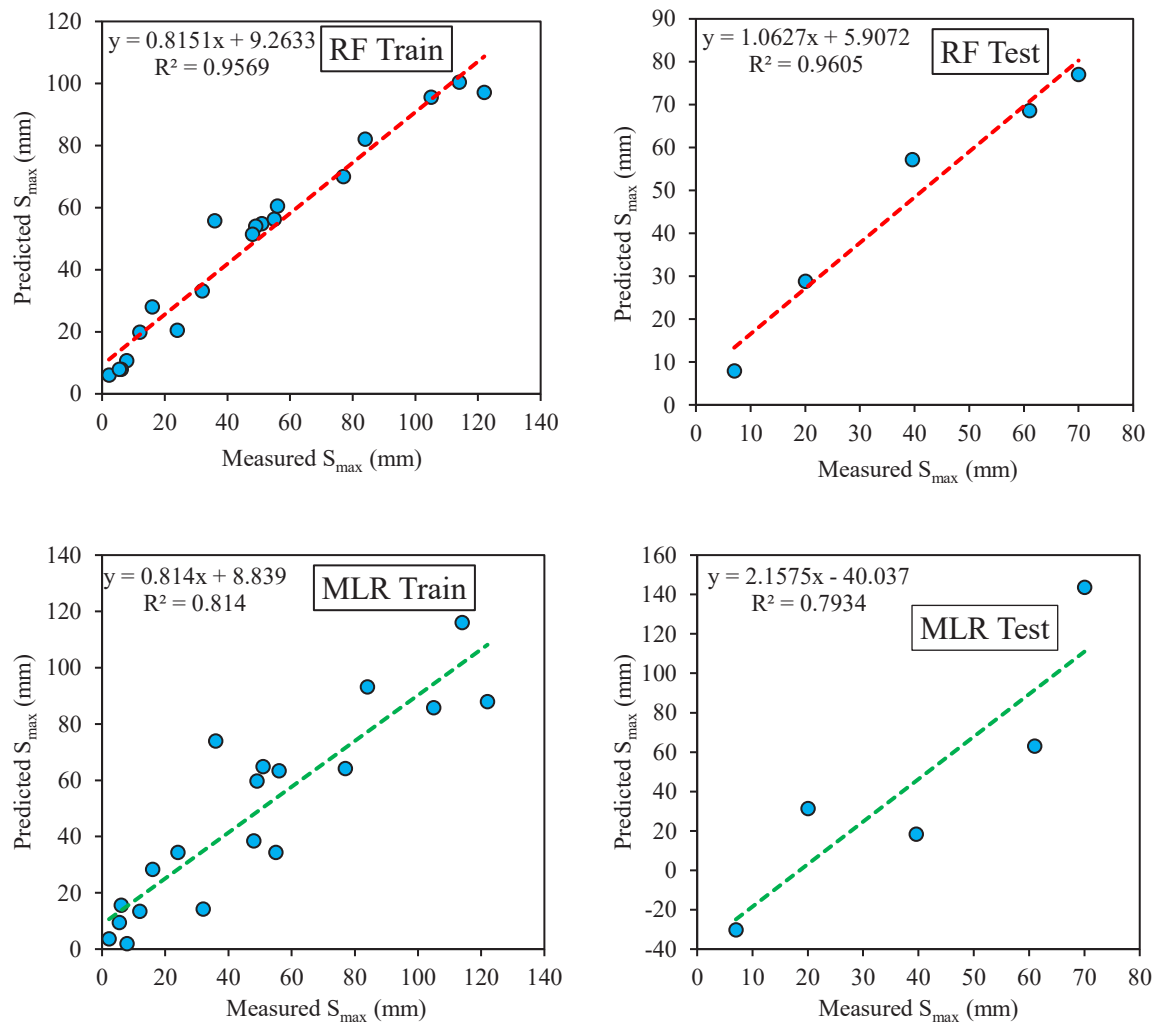
In this study, multiple linear regression (MLR) model and random forest (RF) algorithm were used to predict  $S_{\max}$  induced by tunneling operations. MLR is a statistical methodology that is widely employed to estimate the result of a variable predicted by the value of two or more variables. Random Forest (RF) is a type of ensemble learning technique, originally introduced by Breiman in 2001 [11], designed to enhance the accuracy of classification and regression tree (CART) models by integrating a vast collection of decision trees. Decision trees can be classified into two categories, namely, classification and regression trees [11]. Given that the primary goal of this research is to predict  $S_{\max}$  caused by tunneling operations, we only focus on the regression mode. The RF regression technique is classified as a nonparametric regression methodology.

## FINDINGS AND ARGUMENT

In order to attain the optimal version of MLR and RF models for predicting  $S_{\max}$ , their performances are assessed throughout the training and testing phases by utilizing three performance indicators: determination coefficient ( $R^2$ ), root mean square error (RMSE), and mean absolute error (MAE). Through the utilization of the trial and error methodology, the optimal MLR and RF models are achieved. The values of the aforesaid indicators for the training and test phase data are shown in Table 1. Figure 1 displays the acquired relationship between the measured and predicted values of  $S_{\max}$  from both the RF and MLR models within the training and testing stages. As illustrated in Figure 1, the MLR and RF models exhibit  $R^2$  values of 0.814 and 0.957, respectively, when predicting  $S_{\max}$  during the training phase. In the testing phase, these models yield  $R^2$  values of 0.793 and 0.96, respectively. Moreover, the findings reveal that the RF algorithm exhibits lower RMSE and MAE values in both the training and testing phases compared to the MLR method. This suggests that the RF algorithm exhibits reduced error and higher reliability and accuracy when compared to the MLR model. From the results obtained above, it can be inferred that the RF algorithm demonstrates greater reliability and heightened accuracy when compared to the MLR model in its ability to predict  $S_{\max}$  induced by tunneling operations. Furthermore, the RF results reveal a strong correlation with the measured

**Table 1.** Values of statistical indicators in training and testing phases of MLR and RF models

index	Train phase		Test phase	
	RF	MLR	RF	MLR
$R^2$	0.9569	0.8139	0.9605	0.7934
RMSE	9.36	15.93	9.94	38.42
MAE	6.84	12.64	8.38	29.09



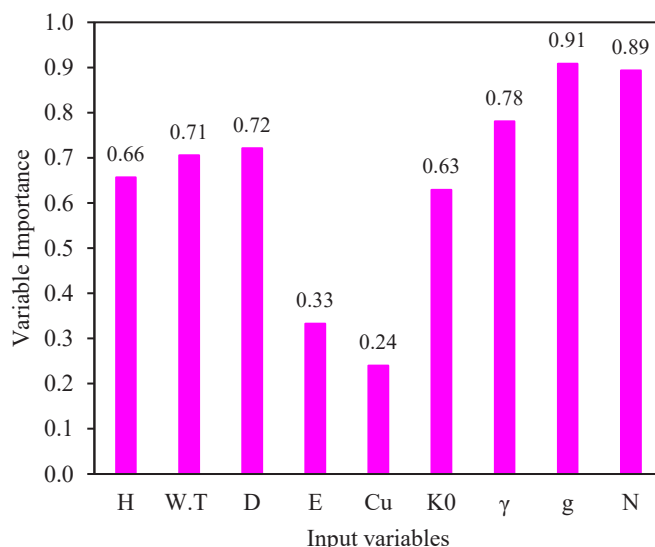
**Figure 1.** The relationship of measured and predicted values of the  $S_{max}$  in training and testing phases of RF and MLR models

data and align closely with the actual values. Consequently, it can be deduced that the RF algorithm serves as a suitable tool for the prediction of  $S_{max}$  resulting from tunneling operations conducted in urban environments.

Based on the methodology employed to assess importance, the impact values of 9 input variables on the resulting  $S_{max}$  value derived from the RF algorithm are computed and presented in Figure 2. It is evident that the parameter  $g$  holds the highest importance in determining  $S_{max}$ . Conversely, the parameter  $C_u$  is of the least importance in modeling the  $S_{max}$  parameter.

## CONCLUSIONS

In this paper, using two methods of multiple linear regression (MLR) and random forest (RF) algorithm, the prediction of  $S_{max}$  caused by tunneling operations in urban environments was studied. MLR and RF methods were implemented based on 24 data sets related to different tunneling projects and selection of 9 variables as influencing parameters on  $S_{max}$ . In the first stage, 80% of the data were used as training data and 20% as test data to implement the suggested methods. Then, in order to evaluate the efficiency of the used methods, three performance indicators of the coefficient of determination parameter ( $R^2$ ), root mean square error (RMSE), and mean absolute error (MAE) were used for the training and test data. The results of the models showed that the  $R^2$  values of MLR and RF methods for training data are 0.814 and 0.957, respectively, and for test data are 0.793 and 0.96, respectively. Also, the findings demonstrate that the RF algorithm exhibits lower RMSE and MAE values in both the training and testing phases compared to the



**Figure 2.** The result of importance study for  $S_{max}$  modeling

MLR method. The obtained results indicate the higher efficiency and accuracy of RF method compared to MLR. Finally, variable importance (VI) index was utilized to determine the most and least important parameters affecting  $S_{max}$ . The obtained results showed that among the input parameters, the gap parameter (g) and the undrained shear strength of soil ( $C_u$ ) have the greatest and least influence on  $S_{max}$ , respectively.

## REFERENCES

- [1] Herzog, M. (1985). "Surface subsidence above shallow tunnels (in German)". Bautechnik, 62: 375-377.
- [2] O'Reilly, M. P., and New, B. M. (1982). "Settlement above tunnels in the United Kingdom-their magnitude and prediction". In: Proceedings of the tunnelling conference, Brighton, 173-181.
- [3] Peck, R. B. (1969). "Deep excavations and tunnelling in soft ground". In: 7th International conference on soil mechanics and foundation engineering, State of the Art Volume, 225-290.
- [4] Schmidt, B. (1969). "A method of estimating surface settlement above tunnels constructed in soft ground". Canadian Geotechnical Journal, 20: 11-22.
- [5] Chou, W. I., and Bobet, A. (2002). "Predictions of ground deformations in shallow tunnels in clay". Tunnelling and Underground Space Technology, 17: 3-19.
- [6] Loganathan, N., and Poulos, H. G. (1998). "Analytical prediction for tunneling-induced ground movements in clays". Journal of Geotechnical and Geoenvironmental Engineering, 124(9): 846-856.
- [7] Park, K. H. (2005). "Analytical solution for tunneling-induced ground movement in clays". Tunnelling and Underground Space Technology, 20(3): 249-261.
- [8] Verruijt, A., and Booker, J. R. (1996). "Surface settlement due to deformation of a tunnel in an elastic half plane". Geotechnique, 46(4): 753-756.
- [9] Ercelebi, S. G., Copur, H., and Ocak, I. (2011). "Surface settlement predictions for Istanbul Metro tunnels excavated by EPB-TBM". Environmental Earth Sciences, 62(2): 357-365. DOI: 10.1007/s12665-010-0530-6.
- [10] Chakeri, H., and Unver, B. (2014). "A new equation for estimating the maximum surface settlement above tunnels excavated in soft ground". Environmental Earth Sciences, 71: 3195-3210.
- [11] Breiman, L. (2001). "Random forest". Machine Learning, 45(1): 5-32.



## پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی در محیط‌های شهری با مدل رگرسیون خطی چندگانه و الگوریتم جنگل تصادفی

حسام دهقانی<sup>۱</sup>، فریبرز متین‌پور<sup>۲</sup>، شادمان محمدی بلبان آباد<sup>۳</sup>، مسعود منجزی<sup>۴</sup>

- ۱- دانشیار، گروه مهندسی معدن، دانشکده مهندسی، دانشگاه صنعتی همدان، همدان
- ۲- دانشجوی دکترا، گروه مهندسی معدن، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران
- ۳- دانشجوی دکترا، گروه مهندسی معدن، دانشکده مهندسی، دانشگاه تربیت مدرس، تهران
- ۴- استاد، گروه مهندسی معدن، دانشکده مهندسی، دانشگاه تربیت مدرس، تهران

پذیرش: ۱۴۰۲/۰۸/۱۳

دریافت: ۱۴۰۲/۰۵/۰۵

### چکیده

نشست ناشی از عملیات تونلسازی در محیط‌های شهری پدیده‌ای اجتناب‌ناپذیر است. پیش‌بینی و کنترل این پدیده آسیب‌های احتمالی به سازه‌های سطحی و زیرساخت‌های مجاور را به طور قابل توجهی کاهش می‌دهد. در این مقاله، برای پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات حفاری تونل‌های کم‌عمق در محیط‌های شهری از دو روش رگرسیون خطی چندگانه (MLR) و الگوریتم جنگل تصادفی (RF) استفاده شده است. بدین منظور ۹ پارامتر ورودی موثر بر حداکثر نشست سطح زمین از جمله فاصله مرکز تونل از سطح زمین (H)، ارتفاع سطح آب زیرزمینی بالای تونل (W.T)، قطر تونل (D)، مدول الاستیسته خاک (E)، مقاومت برشی زهکشی نشده خاک (Cu)، ضریب فشار زمین ( $K_0$ )، وزن مخصوص خاک ( $\gamma$ )، پارامتر گپ (g) و عدد پایداری (N) بر اساس ۲۴ دسته داده مربوط به ۱۴ پروژه مختلف تونلسازی انتخاب و سپس روش‌های MLR و RF پیاده‌سازی شدند. برای ارزیابی کارایی مدل‌ها در پیش‌بینی حداکثر نشست از ۳ شاخص ضریب تعیین ( $R^2$ )، جذر میانگین مربعات خطا (RMSE) و میانگین خطای مطلق (MAE) برای داده‌های آموزش و تست استفاده شد. مقادیر ضریب تعیین روش‌های MLR و RF برای داده‌های آموزش به ترتیب ۰/۸۱۴ و ۰/۹۵۷ و برای داده‌های تست به ترتیب ۰/۷۹۳ و ۰/۹۶ به دست آمد که نشان‌دهنده کارایی بالای روش RF در مقایسه با MLR است. به علاوه، نتایج نشان داد که مقادیر شاخص‌های RMSE و MAE در هر دو مرحله آموزش و تست برای الگوریتم RF کمتر از روش MLR است که خطای کمتر الگوریتم RF و قابلیت اطمینان و دقت بالاتر آن نسبت به روش MLR نشان می‌دهد. همچنین، نتایج آنالیز اهمیت نشان می‌دهد که از بین پارامترهای ورودی، پارامتر گپ (g) و مقاومت برشی زهکشی نشده خاک (Cu) به ترتیب بیشترین و کمترین تاثیر را بر حداکثر نشست سطح زمین دارند.

### کلمات کلیدی

عملیات تونلسازی، حداکثر نشست سطح زمین، رگرسیون خطی چندگانه، الگوریتم جنگل تصادفی.

### استناد به این مقاله

دهقانی، ح.، متین‌پور، ف.، محمدی بلبان آباد، ش.، منجزی، م.؛ ۱۴۰۳؛ "پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی در محیط‌های شهری با مدل رگرسیون خطی چندگانه و الگوریتم جنگل تصادفی". نشریه مهندسی منابع معدنی، دوره نهم، شماره ۳، ص ۹۴-۷۷.

DOI: 10.30479/JMRE.2023.19111.1654



## ۱- مقدمه

نشست سطح زمین ناشی از عملیات تونلسازی پدیده‌ای اجتناب‌ناپذیر و یکی از دغدغه‌های اصلی در پروژه‌های زیرزمینی شهری است. پیش‌بینی و کنترل این پدیده آسیب‌های احتمالی به سازه‌های سطحی و زیرساخت‌های مجاور را به میزان قابل توجهی کاهش می‌دهد. به علت اهمیت نشست و تاثیر مستقیم آن بر هزینه‌های پروژه، تاکنون مطالعات زیادی بر اساس روش‌های مختلف صورت گرفته و روابطی برای پیش‌بینی آن ارائه شده است. از جمله این روش‌ها می‌توان به روش‌های تجربی، تحلیلی، عددی و نیز رویکردهای مبتنی بر هوش مصنوعی اشاره کرد.

در عمل پیش‌بینی جابه‌جایی زمین عمدتاً بر اساس روابط تجربی استوار است [۵-۱] که پارامترهای ذاتی خاک را در نظر نمی‌گیرند. روابط تحلیلی اگرچه حل فرم بسته و ابزار مفیدی برای پیش‌بینی حداکثر نشست سطح زمین هستند، با این حال، این روش‌ها نمی‌توانند تمام پارامترهای تاثیرگذار بر نشست را در خود جای دهند [۹-۶]. در روش‌های عددی، معادلات ریاضی به روش المان محدود و یا المان مجزا برای پیش‌بینی جابه‌جایی لایه‌های زمین حل شده و در نهایت نشست سطح زمین محاسبه می‌شود [۱۳-۱۰]. روش‌های هوش مصنوعی بر مبنای یادگیری خیره تعداد قابل توجهی داده برای یافتن رابطه بین پارامترهای مستقل و وابسته هستند [۲۳-۱۴]. برخی از روابط ارائه شده با استفاده از روش‌های مختلف برای پیش‌بینی نشست ناشی از حفر تونل در محیط‌های کم‌عمق شهری در جدول ۱ آورده شده است.

در این مقاله، با استفاده از مدل رگرسیون خطی چندگانه (MLR) و الگوریتم جنگل تصادفی (RF) به پیش‌بینی حداکثر نشست ناشی از عملیات تونلسازی در محیط‌های کم‌عمق شهری و مهم‌ترین پارامترهای تاثیرگذار بر آن پرداخته می‌شود.

## ۲- مروری بر منابع

پک<sup>۱</sup> با جمع‌آوری تعداد زیادی داده مربوط به پروژه‌های مختلف تونلسازی نشان داد که شکل منحنی گودی نشست در حالت ۲ بُعدی مشابه منحنی توزیع نرمال گوس است [۳]. او بر این اساس رابطه‌ای برای پیش‌بینی نشست سطح زمین ارائه داد که بعدها مبنای تحقیقات بعدی قرار گرفت. اشمیت<sup>۲</sup> رابطه‌ای برای پیش‌بینی حداکثر نشست سطح زمین

ارائه داد [۴]. هرزوغ<sup>۳</sup> با استفاده از نتایج یک سری تحقیقات آزمایشگاهی رابطه‌ای برای پیش‌بینی حداکثر نشست سطح زمین توسعه داد [۱]. لگانتن<sup>۴</sup> و پلوس<sup>۵</sup> یک راه حل تحلیلی برای پیش‌بینی نشست ناشی از حفر تونل در خاک‌های رسی ارائه دادند و سپس برای اعتبارسنجی رابطه ارائه شده از داده‌های اندازه‌گیری شده مربوط به ۵ مطالعه موردی استفاده کردند [۷]. آنها همچنین کارایی رابطه ارائه شده را با ۲ رابطه تجربی و تحلیلی دیگر مقایسه کردند. نتایج تحقیقات آنها نشان داد که در حالت کلی رویکرد تحلیلی پیشنهادی با قابلیت خوبی نشست سطح زمین ناشی از عملیات تونلسازی در خاک‌های کاملاً رسی را پیش‌بینی می‌کند. چوو<sup>۶</sup> و بابت<sup>۷</sup> یک راه حل تحلیلی بر اساس مفاهیم الاستیسته و با فرض شرایط کرنش صفحه‌ای برای تعیین نشست ناشی از حفر تونل‌های کم عمق در زمین‌های اشباع ارائه دادند و سپس از مقادیر نشست اندازه‌گیری شده مربوط به ۲۸ تونل برای اعتبارسنجی رابطه پیشنهادی استفاده کردند [۶]. نتایج تحقیقات آنها نشان داد که پیش‌بینی نشست با استفاده از روش پیشنهادی در حالت کلی حداکثر ۱۵٪ با مقدار واقعی اختلاف دارد. پارک<sup>۸</sup> حل الاستیکی را برای تعیین میدان‌های جابه‌جایی افقی و قائم اطراف تونل‌های کم‌عمق و عمیق در زمین‌های نرم ارائه کرد [۸].

آدنبروک<sup>۹</sup> و پوتس<sup>۱۰</sup> با استفاده از روش المان محدود و اجرای ۱۳ مدل عددی در شرایط کرنش صفحه‌ای به پیش‌بینی حداکثر نشست ناشی از حفر تونل‌های دو قلو پرداختند. آنها ۲ هندسه متفاوت قرارگیری تونل‌ها نسبت به هم یعنی: (۱) ۲ تونل با عمق یکسان و فاصله افقی معین از هم و (۲) یکی از تونل‌ها با فاصله معینی دقیقاً بالای تونل دومی، را مورد بررسی قرار دارند. نتایج تحقیقات آنها نشان داد که فاصله و نحوه قرارگیری تونل‌ها نسبت به هم تاثیر قابل توجهی بر میزان حداکثر نشست سطح زمین دارد [۱۰]. ارسلیبی<sup>۱۱</sup> و همکاران با استفاده از روش المان محدود و با به کارگیری نرم‌افزار Plaxis 2D به پیش‌بینی حداکثر نشست سطح زمین ناشی از حفر تونل‌های ۲ قلو مترو استانبول پرداختند و سپس نتایج روش عددی را با مقادیر اندازه‌گیری برجا و روش‌های تجربی و تحلیلی دیگر مقایسه کردند. یافته‌های آنها نشان داد که روش عددی با قابلیت اطمینان بالایی نسبت به دیگر روش‌ها مقدار حداکثر نشست را پیش‌بینی می‌کند [۱۳]. چاکری<sup>۱۲</sup> و همکاران با استفاده از مدلسازی عددی به پیش‌بینی حداکثر نشست ناشی از حفر تونل خط ۷ مترو تهران پرداختند [۱۱].

روش‌های معمول (تجربی، عددی و تحلیلی) کارایی نسبتاً بهتری دارند. ژوو<sup>۲۰</sup> و همکاران با استفاده از الگوریتم جنگل تصادفی و به کارگیری ۲ دسته داده (دسته اول: ۲۶ مجموعه و دسته دوم: ۴۰ مجموعه) به پیش‌بینی حداکثر نشست و طول گودی نشست سطح زمین ناشی از حفر تونل‌های کم عمق پرداختند [۲۲]. پارامترهای ورودی هر کدام از دسته داده‌ها متفاوت بود. یافته‌های آنها نشان داد که روش جنگل تصادفی نتایج قابل قبولی ارائه می‌دهد و می‌توان به عنوان یک ابزار جایگزین روش‌های تجربی و دیگر روش‌های هوش مصنوعی برای پیش‌بینی نشست ناشی از تونلسازی از آن استفاده کرد. حسنی پناه و همکاران از یک روش هوش مصنوعی ترکیبی مبتنی بر شبکه‌های عصبی مصنوعی که با استفاده از الگوریتم بهینه‌سازی ازدحام ذرات معماری آن بهینه‌سازی شده بود، برای پیش‌بینی حداکثر نشست ناشی از حفر تونل مترو خط ۲ کرج، استفاده کردند [۱۴]. آنها از ۱۴۳ دسته داده برای پیاده‌سازی روش ترکیبی استفاده و نتایج روش ترکیبی پیشنهادی را با شبکه‌های عصبی مصنوعی مقایسه کردند. نتایج تحقیقات آنها نشان داد که روش ترکیبی پیشنهادی با قابلیت اطمینان و دقت بالاتری حداکثر نشست سطح زمین را نسبت به شبکه عصبی سنتی پیش‌بینی می‌کند. بعدها مقدسی و نوریان بیدگلی به پیش‌بینی حداکثر نشست سطح زمین همان پروژه خط ۲ مترو کرج با استفاده از روش هوش مصنوعی دیگری مشابه با روش مورد استفاده حسنی پناه و همکاران با این تفاوت که در روش ترکیبی پیشنهادی از الگوریتم رقابت استعماری برای بهینه‌سازی ساختار شبکه عصبی استفاده شده بود، پرداختند [۱۶]. آنها روش ترکیبی پیشنهادی را با شبکه‌های عصبی مصنوعی سنتی و رگرسیون خطی چندگانه مقایسه کردند و دریافتند که روش ترکیبی پیشنهادی با دقت بالاتری حداکثر نشست سطح زمین را نسبت به ۲ روش دیگر پیش‌بینی می‌کند. ژانگ<sup>۲۱</sup> و همکاران از یک رویکرد هوش مصنوعی ترکیبی بر اساس شبکه‌های عصبی مصنوعی با به کارگیری یک الگوریتم تکامل دیفرانسیلی برای معماری بهینه شبکه برای پیش‌بینی حداکثر نشست ناشی از تونلسازی سپری، استفاده کردند [۲۳]. در ادامه مولفان برای بررسی قابلیت کارایی مدل، از داده‌های خط ۹ مترو گوانگژو چین استفاده کردند و دریافتند که مدل پیشنهادی دقت بالایی در پیش‌بینی نشست ناشی از تونلسازی سپری دارد. محمود زاده و همکاران از چندین روش هوش مصنوعی برای پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی استفاده

آنها سپس نتایج مدل‌سازی عددی را با مقادیر اندازه‌گیری شده و نیز مقادیر حداکثر نشست حاصل از روابط تجربی و تحلیلی دیگر مقایسه کردند و دریافتند که روش عددی مورد استفاده با دقت بیشتری نسبت به روابط تجربی و تحلیلی حداکثر نشست سطح زمین را پیش‌بینی می‌کند. به علاوه، آنها تاثیر قطر تونل، فشار جبهه‌کار و سربرار را بر مقدار حداکثر نشست سطح زمین بررسی کردند و دریافتند که با افزایش پارامترهای یاد شده مقدار حداکثر نشست به ترتیب افزایش، کاهش و افزایش می‌یابد. چاکری و آنور<sup>۱۳</sup> با استفاده از روش المان مجزا و به کارگیری نرم‌افزار FLAC 3D به ساخت مدل‌های عددی متعددی بر مبنای هندسه و زمین‌شناسی ۳ تونل مترو تهران، استانبول و مشهد اقدام کردند [۱۲]. سپس با انجام مطالعه پارامتریک، با تغییر یکی از پارامترهای مدل عددی و ثابت نگه داشتن سایر پارامترها، مقدار حداکثر نشست سطح زمین در هر مدل عددی را به دست آوردند. در نهایت بر اساس داده‌های به دست آمده در حالات مختلف، رابطه تجربی جدیدی برای تعیین حداکثر نشست سطح زمین ارائه دادند.

در سال‌های اخیر به کارگیری روش‌های مبتنی بر هوش مصنوعی برای پیش‌بینی نشست ناشی از عملیات تونلسازی در محیط‌های شهری مورد استقبال زیادی قرار گرفته است. نیوپن<sup>۱۴</sup> و آدهیکاری<sup>۱۵</sup> یک شبکه عصبی پرسپترون ۳ لایه را برای پیش‌بینی حداکثر نشست، عرض گودی نشست و جابه‌جایی افقی ایجاد شده در سطح زمین با در نظر گرفتن ۳۶ دسته داده مربوط به پروژه‌های تونلسازی مختلف، به کار گرفتند. یافته‌های آنها نشان داد که شبکه عصبی به عنوان یک ابزار قدرتمند در پیش‌بینی موفقیت آمیز حداکثر نشست ناشی از حفر تونل استفاده می‌شود [۱۸، ۱۹]. سانتوس<sup>۱۶</sup> و سلسستینو<sup>۱۷</sup> با به کارگیری شبکه‌های عصبی مصنوعی به پیش‌بینی حداکثر نشست سطح زمین ناشی از احداث تونل‌های نگهداری شده با استفاده از شاتکریت در خط ۲ متروی سائوپائولو برزیل که در رسوبات ترشیری حفر شده بودند، پرداختند [۲۱]. مولفان دریافتند که استفاده از شبکه عصبی ابزار مناسبی برای پیش‌بینی حداکثر نشست در دیگر تونل‌های با شرایط مشابه با تونل مورد مطالعه است. اک<sup>۱۸</sup> و سکر<sup>۱۹</sup> با به کارگیری ۳ روش شبکه عصبی مصنوعی، ماشین بردار پشتیبان و فرآیند گاوسی به پیش‌بینی حداکثر نشست کوتاه مدت ناشی از حفر تونل‌های ۲ قلو مربوط به متروی استانبول پرداختند [۲۰]. تحقیقات آنها نشان داد خطای روش‌های به کار گرفته شده در پیش‌بینی نشست کم بوده و این روش‌ها نسبت به دیگر

کردند [۱۵]. آنها برای اجرای روش‌های هوش مصنوعی از ۳۰۰ دسته داده مربوط به ۸ پروژه تونلسازی در ایران استفاده کردند. نتایج تحقیقات آنها نشان داد که روش شبکه‌های عصبی مصنوعی عمیق (DNNs) کارایی بالاتری نسبت به سایر روش‌های مورد استفاده دیگر در پیش‌بینی حداکثر نشست در پروژه‌های تونلسازی مورد بررسی، دارد. قیاسی و کوشکی با اجرای تعداد زیادی مدل عددی المان محدود با به کارگیری نرم‌افزار آباکوس به پیش‌بینی نشست سطح زمین ناشی از حفر تونل در خاک‌های اشباع پرداختند [۲۴]. آنها تاثیر پارامترهای عمق به قطر تونل (H/D)، چسبندگی خاک، زاویه اصطکاک داخلی و ضریب نفوذپذیری را مورد بررسی قرار دادند و سپس مقادیر نشست خروجی مدل‌های عددی را به دست آوردند. مولفان در ادامه با توجه به ۱۷۰ دسته داده به دست آمده از مدل‌سازی عددی، با به کارگیری شبکه‌های عصبی پرسپترون چند لایه با ۴ پارامتر ورودی، به پیش‌بینی نشست سطح زمین پرداختند. نتایج مدل‌سازی آنها نشان داد که شبکه‌های عصبی توانایی بالایی در پیش‌بینی نشست سطح زمین دارند. صمدی و همکاران با به کارگیری ۲ روش هوش مصنوعی ماشین بردار پشتیبان و شبکه عصبی پرسپترون چند لایه و نیز استفاده از روش رگرسیون خطی چندگانه به پیش‌بینی حداکثر نشست ناشی از حفر تونل مترو خط ۶ تهران پرداختند [۲۵]. آنها در مطالعه خود ۳ پارامتر ورودی تاثیرگذار بر حداکثر نشست شامل شاخص گشتاور، شاخص رانش و فشار زمین را در نظر گرفتند. نتایج تحقیقات آنها نشان داد که مقادیر پیش‌بینی شده حداکثر نشست با استفاده از ۳ روش به کار گرفته شده به مقادیر اندازه‌گیری شده آنها بسیار نزدیک است که کارایی بالای این روش‌ها را نشان می‌دهد. اسوائی<sup>۲۲</sup> و همکاران با استفاده از مدل‌سازی عددی تاثیر پارامترهای فشار تزریق گروت، فشار سینه‌کار، نسبت عمق به قطر تونل (H/D)، قطر تونل و مدول الاستیسته خاک را بر نشست ناشی از حفر تونل در محیط‌های رسوبی نرم مورد بررسی قرار دادند [۲۶]. نتایج تحقیقات آنها نشان داد که نشست سطح زمین با قطر تونل رابطه مستقیم و با فشار تزریق گروت، فشار سینه‌کار و مدول الاستیسته خاک رابطه عکس دارد. همچنین آنها دریافتند که با افزایش نسبت عمق به قطر تونل تا یک حد معین، نشست سطح زمین ابتدا افزایش و سپس به تدریج مقدار آن کاهش می‌یابد. ژانگ و همکاران با استفاده از الگوریتم تقویت گرادیان<sup>۲۳</sup>، شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان و رگرسیون اسپیلاین تطبیقی چندگانه به پیش‌بینی نشست ناشی از تونلسازی سپری

۳ پروژه تونلسازی در سنگاپور با در نظر گرفتن ۸ پارامتر ورودی تاثیرگذار، پرداختند [۲۷]. یافته‌های آنها نشان داد که الگوریتم تقویت گرادیان با دقت بالاتری نسبت به ۳ روش دیگر نشست سطح زمین را پیش‌بینی می‌کند. به علاوه، مولفان دریافتند که مقدار متوسط تست نفوذ استاندارد تونل<sup>۲۴</sup> و مدول الاستیسته خاک مهم‌ترین پارامترهای تاثیرگذار بر نشست هستند. کیم<sup>۲۵</sup> و همکاران از ۵ روش هوش مصنوعی برای پیش‌بینی نشست ناشی از حفر یک تونل مترو در هنگ‌کنگ پرداختند [۲۸]. یافته‌های آنها نشان داد که الگوریتم تقویت گرادیان با دقت بالاتری نسبت به الگوریتم‌های مورد استفاده دیگر نشست سطح زمین را پیش‌بینی می‌کند. لی<sup>۲۶</sup> و همکاران با به کارگیری ۳ روش هوش مصنوعی جنگل تصادفی (RF)، شبکه‌های حافظه کوتاه مدت (LSTM) و GRU به پیش‌بینی حداکثر نشست ناشی از حفر تونل چونفنگ چین پرداختند [۲۹]. روش‌های LSTM و GRU، ۲ نوع خاص از شبکه‌های عصبی بازگشتی (RNN) هستند. یافته‌های آنها نشان داد که روش‌های به کار گرفته شده قابلیت بالایی در پیش‌بینی نشست دارند. به علاوه از بین ۳ روش مورد استفاده، شبکه‌های حافظه کوتاه‌مدت نسبت به ۲ روش دیگر با دقت بالاتری حداکثر نشست سطح زمین پیش‌بینی می‌کند. حسین و میو<sup>۲۷</sup> با به کارگیری ۵ روش هوش مصنوعی به پیش‌بینی حداکثر نشست ناشی از حفر ۲ تونل مترو در هانگژو چین پرداختند [۳۰]. نتایج تحقیقات آنها نشان داد که روش رگرسیون درختان اضافی<sup>۲۸</sup> نسبت به ۴ روش دیگر با دقت بالاتری حداکثر نشست سطح زمین را پیش‌بینی می‌کند. مقتدر و همکاران با به کارگیری ۱۵۰ دسته داده مربوط به پروژه خط ۶ تونل مترو تهران به پیش‌بینی حداکثر نشست سطح زمین با استفاده از شبکه‌های عصبی مصنوعی اقدام کردند [۱۷]. نتایج تحقیقات آنها نشان داد که تزریق گروت و نفوذپذیری به ترتیب بیشترین و کمترین تاثیر را بر مقدار حداکثر نشست سطح زمین دارند. ونگ<sup>۲۹</sup> و همکاران از یک الگوریتم ترکیبی جدید (تجزیه حالت تجربی<sup>۳۰</sup>، الگوریتم فرا ابتکاری تطبیقی جستجوی گنجشک آشفته<sup>۳۱</sup> و ماشین یادگیری افراطی<sup>۳۲</sup>) برای پیش‌بینی نشست سطح زمین بالای تونل استفاده کردند [۳۱]. نتایج آنها نشان داد که مدل ترکیبی پیشنهادی با دقت و سرعت بالایی نشست سطح زمین را پیش‌بینی می‌کند. کُنْگ<sup>۳۳</sup> و همکاران یک روش ترکیبی بر اساس روش‌های یادگیری ماشین و تحلیلی برای پیش‌بینی میدان جابه‌جایی زمین ناشی از حفر تونل توسعه دادند [۳۲].



جدول ۱: برخی از روابط ارائه شده بر مبنای روش‌های مختلف برای پیش‌بینی نشست ناشی از عملیات تونلسازی در محیط‌های شهری

منبع	روش	رابطه
پک (۱۹۶۹) [۳]	تجربی	$S = S_{max} \exp\left(\frac{x^2}{2i^2}\right)$
اشمیت (۱۹۶۹) [۴]	تجربی	$S_{max} = 0.0125A \left(\frac{R^2}{i}\right) ; A = 0.87e^{0.26N}$
چوو و بابت (۲۰۰۲) [۶]	تحلیلی	$\delta_{max} = -\frac{gr}{H} + \frac{1+\nu}{E} \left\{ -\frac{1}{2}[\gamma R^2 \ln H + \gamma_b H(3-k)R^2] + \gamma_b H(1-k)R \left[ \frac{3}{4}\left(\frac{R}{H}\right)^3 - \frac{1}{4}\left(\frac{R}{H}\right)^5 \right] + \gamma_w HR \left(\frac{R}{H}\right)^2 \right\}$
لگانتن و پلوس (۱۹۹۸) [۷]	تحلیلی	$S = 4(1-\nu)R^2 \left(\frac{4gR+g^2}{4R^2}\right) \frac{Z_0}{Z_0^2+x^2} \exp\left(-\frac{1.38x^2}{(H+R)^2}\right)$
وریوجت و بوکر (۱۹۹۶) [۹]	تحلیلی	$S = 4(1-\nu)R^2 \left(\frac{4gR+g^2}{4R^2}\right) \frac{Z_0}{Z_0^2+x^2} - 2\delta R^2 \frac{Z_0(x^2-Z_0^2)}{(Z_0^2+x^2)^2}$
چاکری و آنور (۲۰۱۴) [۱۲]	تجربی (براساس مدل‌سازی عددی)	$S_{max} = 3198.744 \left(\frac{D}{H}\right) \left(\frac{\gamma H + \sigma_s - (c + 0.3\sigma_T)}{E}\right) (1 - \sin \varphi)^{0.8361}$
ارسلیمی و همکاران (۲۰۱۱) [۱۳]	تجربی (براساس داده‌های آزمایشگاهی)	$S_{max} = 0.785(\gamma H + \sigma_s) \left(\frac{D^2}{iE}\right)$
	تجربی (برای تونل‌های دوقلو)	$S_{max} = 4.71(\gamma H + \sigma_s) \left(\frac{D^2}{(3i+d)E}\right)$

$\delta$ : بیضی‌شدگی<sup>۳۴</sup> تونل (بدون بُعد)

$\gamma$ : وزن مخصوص مواد روباره بالای تونل ( $\text{ton/m}^3$ )

$\gamma_w$ : وزن مخصوص آب ( $\text{ton/m}^3$ )

$\gamma_b$ : وزن مخصوص شناور<sup>۳۵</sup> خاک ( $\text{ton/m}^3$ )

$\nu$ : نسبت پواسون

$E$ : مدول الاستیسته مواد روباره (GPa)

$c$ : مقاومت چسبندگی (kPa)

$\varphi$ : زاویه اصطکاک داخلی (درجه)

$\sigma_s$ : فشار سربار (kPa)

$\sigma_T$ : فشار سینه‌کار تونل (kPa)

$k$ : ضریب فشار زمین (بی بُعد) است.

### ۳- مواد و روش‌ها

در این بخش به معرفی و نحوه عملکرد مدل رگرسیون خطی چندگانه و الگوریتم جنگل تصادفی برای پیش‌بینی حداکثر نشست سطح زمین پرداخته می‌شود.

#### ۳-۱- رگرسیون خطی چندگانه<sup>۳۶</sup>

یکی از روش‌های مرسوم در تحلیل چندگانه، تکنیک

آنها روش یادگیری ماشین حداکثر (ELM) را برای پیش‌بینی حداکثر نشست سطح زمین به کار بردند و از الگوریتم ازدحام ذرات (PSO) برای بهینه‌سازی پارامترهای روش یادگیری ماشین حداکثر، استفاده کردند. نتایج آنها نشان داد که روش ترکیبی PSO-ELM کارایی بالایی در پیش‌بینی حداکثر نشست سطح زمین دارد.

در جدول ۱:

$S$ : نشست (m)

$S_{max}$  یا  $\delta_{max}$ : حداکثر نشست (m)

$i$ : فاصله افقی نقطه تععر پروفیل نشست از نقطه حداکثر

نشست (m)

$x$ : فاصله افقی عرضی از مرکز تونل (m)

$R$ : شعاع تونل (m)

$D$ : قطر تونل (m)

$d$ : فاصله بین تونل‌های دوقلو (m)

$g$ : پارامتر گپ (m)

$Z_0$ : فاصله نقطه دلخواه بالای تونل از تاج تونل (m)

$H$ : عمق مرکز تونل از سطح زمین (m)

$N$ : عدد پایداری

و معمولا مقدار آن برابر با جذر تعداد ورودی‌ها در نظر گرفته می‌شود [۳۵،۳۴]. همچنین، در حالت کلی هر چقدر تعداد درختان (ntree) بزرگتر باشد، پیش‌بینی دقت بالاتری دارد. با این حال، نمی‌توان مقادیر مشخصی را برای پارامترهای mtry و ntree در نظر گرفت. با انتخاب مقادیر متفاوت برای این پارامترها می‌توان به مقادیر بهینه آنها برای دستیابی به دقت بالاتر و کاهش خطا الگوریتم در پیش‌بینی هدف، دست یافت. یک جنگل تصادفی آن‌قدر بزرگ است که تفسیر آن کار بسیار دشواری است، بنابراین نیاز است که اطلاعات آن با استفاده از شاخص‌های کمی خلاصه شود. یکی از این شاخص‌ها، شاخص اهمیت متغیر<sup>۳۹</sup> است. شاخص اهمیت متغیر، شاخصی برای رتبه‌بندی متغیرها بر حسب اهمیت آنها در اثرگذاری روی پاسخ است. معروف‌ترین شاخص‌های اهمیت متغیر، شاخص اهمیت گینی<sup>۴۰</sup> و شاخص اهمیت جایگشتی است که در ادامه به آنها پرداخته خواهد شد.

### ۳-۲-۱- شاخص اهمیت گینی

در طی ساخت درخت‌های جنگل تصادفی برای تعیین اینکه گره بر اساس کدام متغیر افزاز شود، از شاخص ناخالصی گینی استفاده می‌شود [۳۴]. اهمیت متغیر  $X_i$  در یک درخت مجموع کاهش در شاخص ناخالصی گینی روی تمام گره‌هایی است که بر اساس  $X_i$  افزاز شده‌اند. میانگین اندازه اهمیت متغیر  $X_i$  روی تمام درخت‌های جنگل، اندازه شاخص اهمیت گینی است.

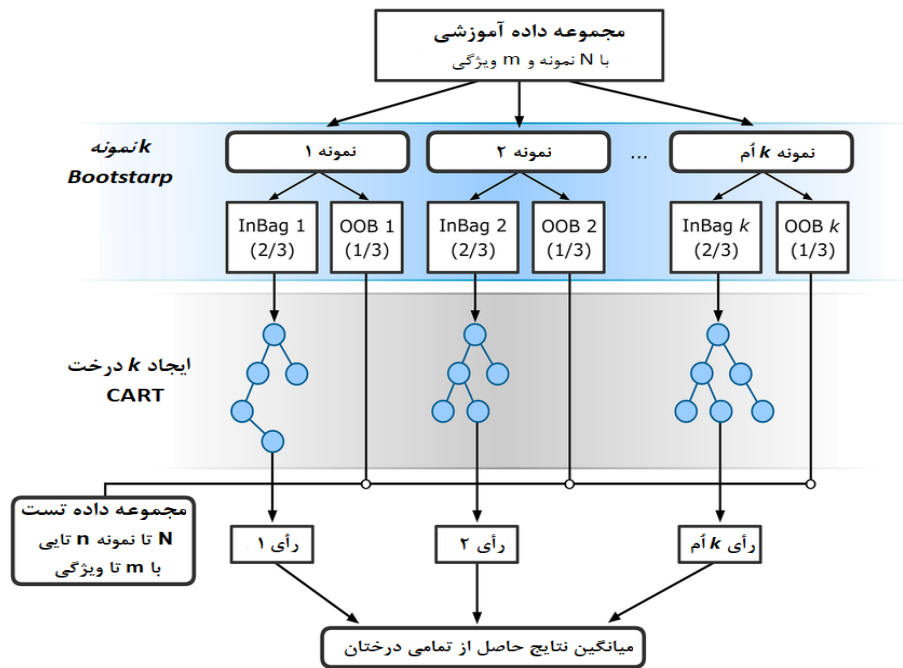
### ۳-۲-۲- شاخص اهمیت جایگشتی

برای محاسبه این شاخص، الگوریتم جنگل تصادفی از تمام مشاهدات نمونه برای ساخت درخت استفاده نمی‌کند، بلکه یک نمونه تصادفی با جایگذاری به حجم  $\Pi_1$  (معمولا دو سوم داده‌ها) برابر از مشاهدات انتخاب می‌شود که به مشاهدات انتخاب شده نمونه آموزشی<sup>۴۱</sup> و به بقیه آن‌ها نمونه خارج از کیسه<sup>۴۲</sup> گفته می‌شود. درخت‌ها با مشاهدات نمونه‌های آموزشی ساخته می‌شوند و از نمونه خارج از کیسه برای اندازه‌گیری ناخالصی درخت استفاده می‌شود. در هر درخت، ابتدا اندازه ناخالصی روی مشاهدات خارج از کیسه محاسبه می‌شود، سپس مقادیر متغیر  $X_i$  مشاهدات خارج از کیسه به طور تصادفی جابه‌جا و اندازه ناخالصی درخت روی مقادیر جابه‌جا شده محاسبه می‌شود. اندازه اهمیت متغیر  $X_i$  در هر درخت، اختلاف بین این

رگرسیون خطی چند متغیره است. بر اساس تحلیل رگرسیونی، یک رابطه خطی بین متغیر پاسخ با یک یا چند متغیر توصیفی برقرار می‌شود. البته گاه به متغیر پاسخ، متغیر وابسته و به متغیرهای توصیفی، متغیرهای مستقل نیز می‌گویند. در رگرسیون خطی چندگانه، پارامترهای یک مدل خطی به کمک یک تابع هدف و مقادیر متغیرها، برآورد می‌شوند [۳۳].

### ۳-۲- الگوریتم جنگل تصادفی<sup>۳۷</sup>

روش‌های درخت، پایه روش‌های آماری ناپارامتری<sup>۳۸</sup> (مدل آزاد) برای اجرای آنالیز کلاس‌بندی و آنالیز رگرسیونی با استفاده از الگوریتم افزازهای بازگشتی هستند [۳۴]. جنگل‌های تصادفی یک نوع مدرن از درخت پایه هستند که شامل انبوهی از درخت‌های کلاس‌بندی و رگرسیونی هستند [۳۵]. مهم‌ترین ویژگی جنگل‌های تصادفی عملکرد بالای آن‌ها در اندازه‌گیری اهمیت متغیرها برای مشخص کردن اینکه هر متغیر چه نقشی در پیش‌بینی پاسخ دارد، است. این الگوریتم در واقع یک الحاقی از درخت رگرسیون و طبقه‌بندی است که اولین بار توسط بریمان ایجاد و توسعه داده شد [۳۵]. الگوریتم جنگل تصادفی مبتنی بر دسته‌ای از درخت‌های تصمیم و در حال حاضر یکی از بهترین الگوریتم‌های یادگیری است. مدل پیش‌بینی کننده RF بر اساس میانگین‌گیری از نتایج حاصل از تمامی درخت‌های تصمیم مربوطه استوار است و برای بسیاری از مجموعه داده‌ها، طبقه‌بندی را با صحت بالایی انجام می‌دهد [۳۶،۳۷]. جنگل تصادفی درخت تصمیم‌های زیادی تولید می‌کند. برای طبقه‌بندی یک شی جدید بردار ورودی در انتهای هر یک از درختان جنگل تصادفی قرار می‌دهد که هر درخت به یک طبقه‌بندی منجر می‌شود که اصلاحا گفته می‌شود این درخت به آن کلاس رای می‌دهد. جنگل حاصل از طبقه‌بندی‌ای که بیشترین رای را داشته باشد (بین همه درخت‌های جنگل) انتخاب می‌شود. در الگوریتم جنگل تصادفی نرخ خطا به: الف) همبستگی بین هر دو درخت در جنگل و ب) قدرت هر یک از درختان در جنگل، بستگی دارد. افزایش همبستگی نرخ خطای جنگل را افزایش می‌دهد، در حالی که افزایش قدرت هر یک از درختان نرخ خطای جنگل را کاهش می‌دهد [۳۷-۳۵]. در الگوریتم جنگل تصادفی دو پارامتر مهم mtry و ntree وجود دارد که به ترتیب نشان‌دهنده تعداد متغیرها در زیرمجموعه تصادفی در هر گره و تعداد درختان هستند که باید بهینه شوند [۳۴]. پارامتر mtry تابعی از تعداد ورودی‌ها است [۲۲]



شکل ۱: روند کلی الگوریتم جنگل تصادفی

#### ۵- پیش‌بینی حداکثر نشست سطح زمین

##### ۱-۵- الگوریتم جنگل تصادفی

در این مرحله ۸۰٪ داده‌ها (۱۹ دسته داده) به عنوان داده‌های آموزش و ۲۰٪ دیگر (۵ دسته داده شامل داده‌های شماره ۲۰ تا ۲۴ جدول ۲ که به صورت تصادفی به وسیله الگوریتم انتخاب شدند) به عنوان داده‌های تست در نظر گرفته شدند. با انتخاب مقادیر متفاوت پارامترهای  $mtry$  و  $ntree$  مدل‌های مختلفی اجرا و در نهایت مقادیر بهینه آنها به ترتیب ۴ و ۱۰۰۰ به دست آمد.

##### ۲-۵- مدل رگرسیون خطی چندگانه

در این بخش بر اساس داده‌های مرحله آموزش (۱۹ دسته داده) و با استفاده از نرم‌افزار اکسل مدل، رگرسیون خطی چندگانه‌ای مطابق رابطه ۱ برای پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی در محیط‌های شهری ارایه شد:

(۱)

$$S_{max} = -5.39385 + 1.160332(H) - 8.91418(W.T) + 3.458546(D) + 0.090297(E) + 0.146881(C_u) - 69.9453(K_0) + 0.513327(\gamma) + 0.186311(g) + 18.32817(N)$$

دو اندازه ناخالصی است و میانگین این مقادیر شاخص اهمیت جایگشتی است. انگیزه این روش آن است که اگر  $X_i$  متغیر مهمی باشد، به طور تصادفی به افزایش ناخالصی درخت منجر می‌شود، در حالی که اگر متغیر تاثیرگذاری نباشد، تغییری در ناخالصی ایجاد نمی‌شود. روند کلی الگوریتم جنگل تصادفی در شکل ۱ نشان داده شده است.

#### ۴- جمع‌آوری داده‌ها

هدف اصلی این تحقیق پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی در محیط‌های شهری با استفاده از مدل‌های رگرسیون خطی چندگانه و الگوریتم جنگل تصادفی است. برای دستیابی به این مهم، ۹ پارامتر ورودی تاثیرگذار بر مقدار حداکثر نشست سطح زمین ( $S_{max}$ ) از جمله فاصله مرکز تونل از سطح زمین ( $H$ )، ارتفاع سطح آب زیرزمینی بالای تونل ( $W.T$ )، قطر تونل ( $D$ )، مدول الاستیسته خاک ( $E$ )، مقاومت برشی زهکشی نشده خاک ( $C_u$ )، ضریب فشار زمین ( $K_0$ )، وزن مخصوص خاک ( $\gamma$ )، پارامتر گپ ( $g$ ) و عدد پایداری ( $N$ ) بر اساس ۲۴ دسته داده مربوط به ۱۴ پروژه تونلسازی مختلف [۶]، برای پیاده‌سازی روش  $RF$  و  $MLR$  استفاده شد (جدول ۲). در جدول ۳ اطلاعات آماری داده‌های مورد استفاده در این مطالعه آورده شده است.

جدول ۲: پارامترهای هندسی و ژئومکانیکی تونل‌ها و زمین اطراف آنها برای پیش‌بینی حداکثر نشست سطح زمین [۶]

No.	H (m)	W.T (m)	D (m)	E (MPa)	C <sub>u</sub> (kN/m <sup>2</sup> )	K <sub>0</sub>	γ (kN/m <sup>2</sup> )	g (mm)	N	S <sub>max</sub> (mm)
۱	۱۰٫۷	۱٫۵	۲٫۴۷	۱۳	۳۵	۱	۱۸	۱۵۶	۴٫۸	۵۱
۲	۱۱٫۷	۲٫۵	۲٫۹۶	۴	۳۵	۱	۱۷	۲۸۸	۵	۱۰٫۵
۳	۲۷	۲	۶٫۲۶	۵	۱۱	۰٫۷	۱۵	۸۰	۴٫۵	۱۲۲
۴	۱۰	۰	۸	۲۵	۷۵	۱	۱۸	۳۱	۲٫۳	۲۴
۵	۷٫۹۳	۴	۲٫۲۵	۲۰	۳۵	۰٫۵	۱۹	۳۲۴	۴٫۱	۸۴
۶	۱۴٫۶	۵٫۶	۶٫۵	۷۰	۳۴۰	۱	۱۹	۶۰	۱٫۲۵	۳۲
۷	۱۹	۶	۵٫۶۶	۱۰	۵۰	۰٫۴۵	۱۶٫۵	۱۱۶	۴٫۶	۳۶
۸	۱۳٫۱	۶	۴٫۲۸	۱۴۰	۳۶۰	۰٫۹	۲۴	۱۰	۰٫۸۶	۲٫۲
۹	۱۰٫۵	۱٫۳	۲٫۴۷	۱۳	۳۵	۱	۱۸	۱۲۰	۴٫۸	۴۹
۱۰	۷٫۵	۰	۲٫۰۲	۱۵	۷۵	۱	۱۹	۱۹	۱٫۹۶	۷٫۹
۱۱	۸	۵	۳	۲۰	۱۲	۰٫۷	۱۹	۱۱۸	۵٫۹	۷۷
۱۲	۷٫۶	۵	۲٫۲۵	۲۰	۳۵	۰٫۵	۱۹	۷۸	۳٫۹	۵۵
۱۳	۱۶	۴	۳٫۲	۲۰	۳۵	۰٫۵	۱۹	۳۲۰	۴٫۷	۱۱۴
۱۴	۵٫۳	۲٫۳	۳	۶	۱۲	۰٫۷	۱۹	۶۶	۳٫۹۵	۴۸
۱۵	۱۸	۲٫۲	۲٫۶۷	۲۰	۱۵	۱	۱۷	۳۷	۳٫۲	۱۲
۱۶	۲۸٫۹	۲٫۳	۴٫۱۵	۵۰	۲۷۰	۱٫۵	۱۹	۲۳	۲٫۲	۶٫۱
۱۷	۴٫۸۵	۱٫۴۵	۲٫۷۴	۴	۱۰	۰٫۷	۱۵	۲۳	۳٫۶۴	۱۶
۱۸	۳۴٫۱	۴٫۳	۴٫۱۴	۵۶	۲۸۰	۱٫۵	۱۹	۲۳	۲٫۴	۵٫۵
۱۹	۷٫۴۴	۴	۳٫۵۷	۲۰	۳۵	۰٫۵	۱۹	۱۰۳	۴٫۵	۵۶
۲۰	۱۰٫۷۴	۸	۳٫۲۳	۲۰	۳۵	۰٫۵	۱۹	۹۸	۳٫۹	۳۹٫۶
۲۱	۱۲٫۸۱	۴	۳	۲۰	۳۵	۰٫۵	۱۹	۱۹۶	۷٫۷	۷۰
۲۲	۴٫۴	۱٫۲	۲٫۷۴	۴	۱۰	۰٫۷	۱۵	۳۰	۳٫۶۴	۲۰
۲۳	۲۰٫۱	۴٫۳	۴٫۱۴	۳۲	۲۳۰	۱٫۵	۱۹	۱۷	۱٫۷	۷
۲۴	۶٫۵	۳٫۵	۳	۶	۱۲	۰٫۷	۱۹	۸۴	۵٫۶۲	۶۱

جدول ۳: اطلاعات آماری داده‌های جمع‌آوری شده برای پیش‌بینی حداکثر نشست سطح زمین

نوع داده	پارامتر	واحد	نماد	حداقل	میانگین	حداکثر	انحراف معیار	واریانس
ورودی	عمق تونل	m	H	۴٫۴۰	۱۳٫۲۰	۳۴٫۱۰	۷٫۷۱	۵۹٫۴۳
	سطح آب زیر زمینی	m	W. T	۰٫۰۰	۳٫۳۵	۸٫۰۰	۱٫۹۸	۳٫۹۲
	قطر تونل	m	D	۲٫۰۲	۳٫۶۵	۸٫۰۰	۱٫۴۹	۲٫۲۳
	مدول الاستیسیته	MPa	E	۴۰۰	۲۵۵٫۵۴	۱۴۰۰۰	۲۸٫۹۷	۸۳۹٫۰۰
	مقاومت برشی زهکشی نشده	kN/m <sup>2</sup>	C <sub>u</sub>	۱۰۰۰	۸۶٫۵۴	۳۶۰۰۰	۱۱۰٫۹۴	۱۲۳۰۸٫۵۸
	ضریب فشار زمین	-	K <sub>0</sub>	۰٫۴۵	۰٫۸۴	۱٫۵۰	۰٫۳۲	۰٫۱۰
	وزن مخصوص	kN/m <sup>3</sup>	γ	۱۵۰۰	۱۸٫۳۱	۲۴۰۰	۱٫۸۱	۳٫۲۹
	پارامتر گپ	mm	g	۱۰۰۰	۱۰۰٫۸۳	۳۲۴۰۰	۹۲٫۲۲	۸۵۰۴٫۶۴
	عدد پایداری	-	N	۰٫۸۶	۳٫۸۰	۷٫۷۰	۱٫۵۷	۲٫۴۶
خروجی	حداکثر نشست سطح زمین	mm	S <sub>max</sub>	۲٫۲۰	۴۵٫۸۵	۱۲۲٫۰۰	۳۴٫۷۵	۱۲۰۷٫۵۷

## ۵-۳- مقایسه و اعتبارسنجی مدل‌های پیشنهادی

برای مقایسه و اعتبارسنجی مدل‌های ارائه شده بر مبنای الگوریتم جنگل تصادفی و مدل رگرسیون خطی چندگانه، از شاخص ارزیابی ضریب تعیین ( $R^2$ )، جذر میانگین مربعات خطا (RMSE) و خطای مطلق میانگین (MAE) استفاده شده است. مقدار  $R^2$  بین صفر و ۱ قرار دارد. مقادیر RMSE و MAE همواره بزرگتر یا مساوی صفر هستند. با نزدیک شدن مقادیر شاخص‌های  $R^2$ ، RMSE و MAE به ترتیب به ۱، صفر و صفر دقت مدل‌ها افزایش می‌یابد. مقادیر شاخص‌های یاد شده را می‌توان با استفاده از روابط ۲ تا ۴ به دست آورد [۲۳، ۲۲]:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_i')^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y_i')^2} \quad (3)$$

$$MAE = \frac{\sum_{i=1}^N |y_i - y_i'|}{N} \quad (4)$$

که در این روابط:

$y_i$ : مقادیر حداکثر نشست واقعی

$y_i'$ : مقادیر حداکثر نشست پیش‌بینی شده

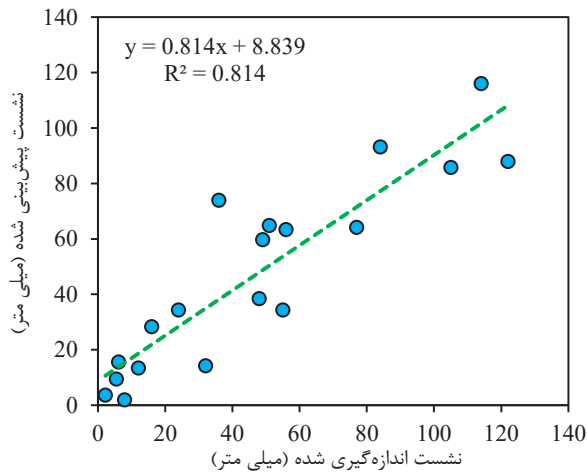
$\bar{y}$ : میانگین مقادیر اندازه‌گیری شده حداکثر نشست

$N$ : تعداد دسته داده‌ها است.

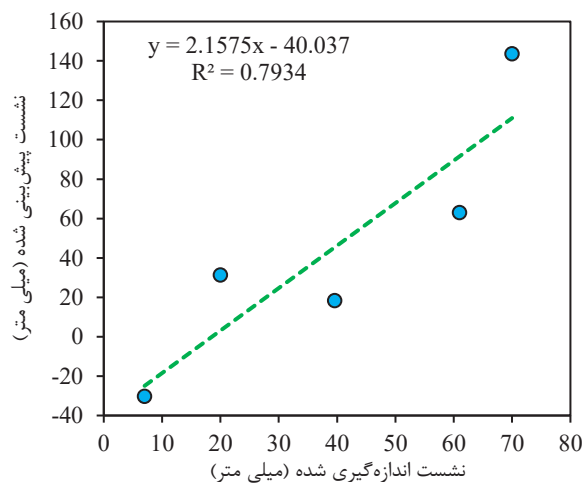
با اعمال روابط ۲ تا ۴ بر داده‌های مرحله آموزش و تست مقادیر این شاخص‌ها برای دو روش MLR و RF مطابق جدول ۴ به دست آمد.

در شکل‌های ۲ و ۳ نتایج نشست حداکثر پیش‌بینی شده در مقابل مقادیر اندازه‌گیری شده آنها با استفاده از روش MLR به ترتیب برای داده‌های آموزش و تست به تصویر کشیده شده‌اند. شکل‌های ۴ و ۵ نیز نتایج یاد شده بر اساس داده‌های مراحل تست و آموزش را برای الگوریتم RF نشان می‌دهند. با توجه به

شکل‌های ۲ تا ۵ و ردیف اول جدول ۴، بر اساس شاخص  $R^2$  می‌توان اظهار کرد که الگوریتم RF با قابلیت اطمینان بالاتر و دقت بیشتری نسبت به روش MLR توانسته است حداکثر نشست سطح زمین را پیش‌بینی کند. به علاوه، مطابق جدول



شکل ۲: عملکرد MLR در فرآیند آموزش



شکل ۳: عملکرد MLR در فرآیند تست

جدول ۴: مقادیر شاخص‌های آماری به دست آمده در مراحل آموزش و تست برای مدل‌های RF و MLR

مرحله تست		مرحله آموزش		شاخص آماری
RF	MLR	RF	MLR	
۰٫۹۶۰۵	۰٫۷۹۳۴	۰٫۹۵۶۹	۰٫۸۱۳۹	$R^2$
۹٫۹۴	۳۸٫۴۲	۹٫۳۶	۱۵٫۹۳	RMSE
۸٫۳۸	۲۹٫۰۹	۶٫۸۴	۱۲٫۶۴	MAE

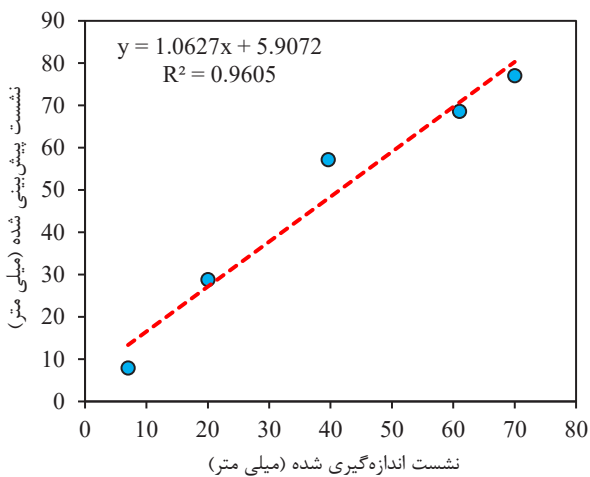
می توان گفت الگوریتم جنگل تصادفی (RF) به عنوان ابزاری مناسب برای پیش بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی در محیط های شهری استفاده می شود.

#### ۶- تحلیل اهمیت متغیر

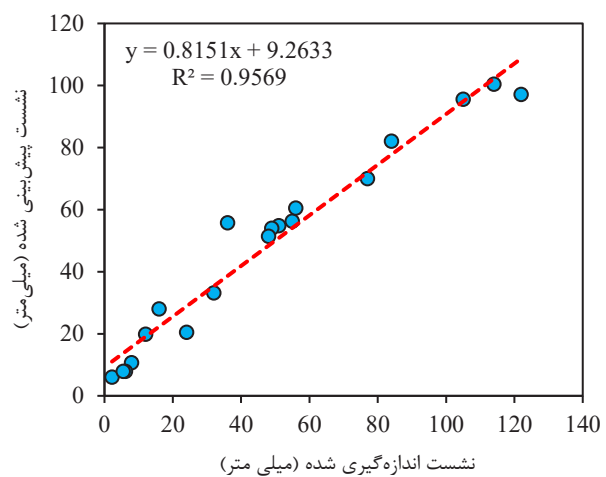
مطالعه پارامتریک معمولا برای ارزیابی اهمیت و تاثیر متغیرهای ورودی بر روی یک متغیر هدف انجام می شود. در این پژوهش، تحلیل اهمیت متغیرهای ورودی الگوریتم RF بر روی نشست سطح زمین با استفاده از نرم افزار STATISTICA 12 در طول توسعه و آزمایش مدل انجام شده است. بر اساس

۴ مقادیر شاخص های RMSE و MAE به دست آمده برای الگوریتم RF در هر دو مرحله آموزش و تست در مقایسه با مقادیر متناظر آنها برای روش MLR بسیار کمتر بوده و به صفر نزدیک تر هستند که نشان دهنده خطای کمتر الگوریتم RF و دقت و کارایی بالاتر آن نسبت به روش MLR است.

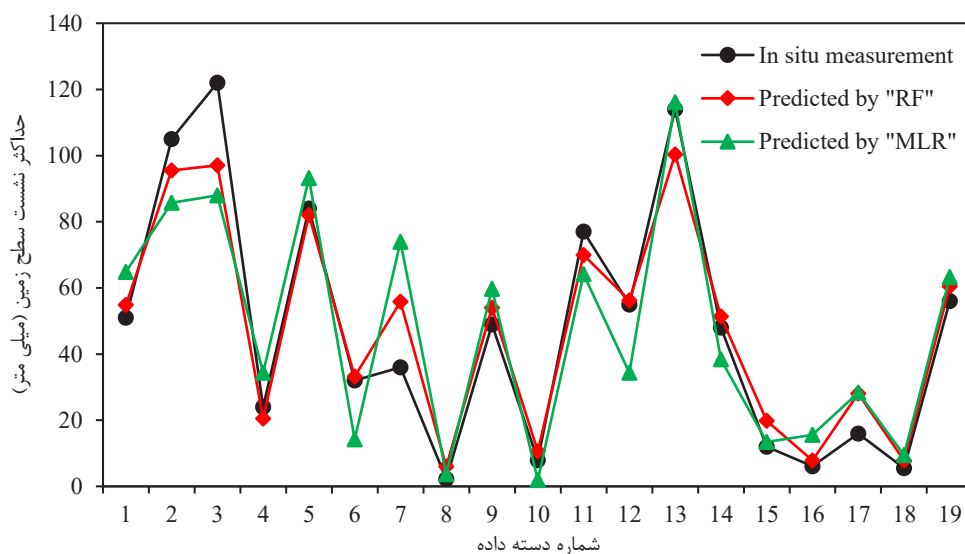
در شکل های ۶ و ۷ به ترتیب مقایسه ای بین مقادیر حداکثر نشست اندازه گیری شده مربوط به هر مطالعه موردی و مقادیر پیش بینی شده آنها با استفاده از روش های RF و MLR برای داده های آموزش و تست صورت گرفته است که نشان دهنده دقت بالای روش RF در مقایسه با روش MLR است، بنابراین



شکل ۵: عملکرد RF در فرآیند تست



شکل ۴: عملکرد RF در فرآیند آموزش

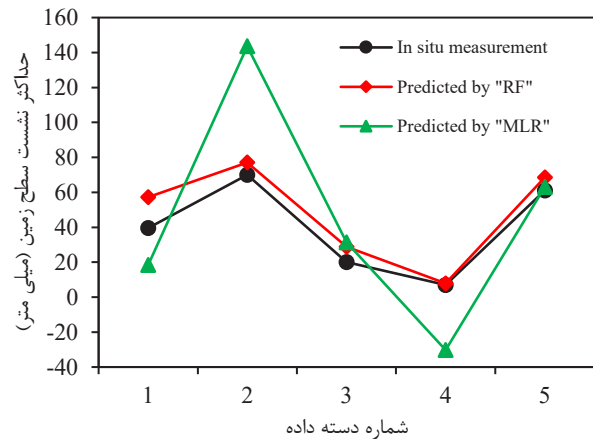


شکل ۶: مقادیر حداکثر نشست اندازه گیری و پیش بینی شده با استفاده از مدل های RF و MLR در مرحله آموزش

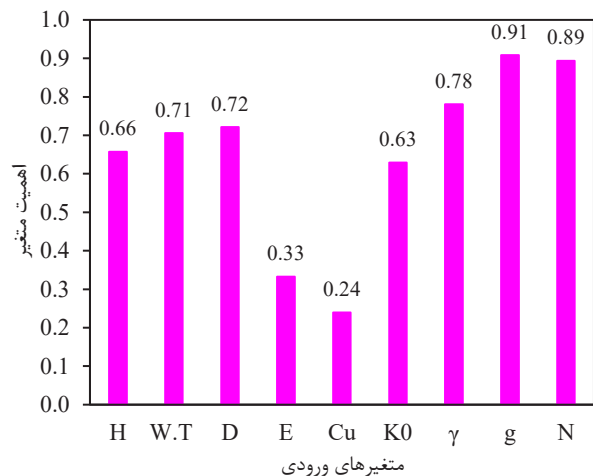
(MLR) و الگوریتم جنگل تصادفی (RF) به پیش‌بینی حداکثر نشست سطح زمین ناشی از عملیات تونلسازی در محیط‌های شهری پرداخته شد. بر اساس ۲۴ دسته داده مربوط به پروژه‌های مختلف تونلسازی و انتخاب ۹ متغیر به عنوان پارامترهای تاثیر گذار بر مقدار حداکثر نشست سطح زمین، روش‌های MLR و RF پیاده‌سازی شدند. در مرحله نخست، ۸۰٪ داده‌ها به عنوان داده‌های آموزش و ۲۰٪ دیگر به عنوان داده‌های تست برای اجرای دو روش، استفاده شدند، سپس برای ارزیابی کارایی مدل‌های مورد استفاده، از ۳ شاخص ارزیابی آماری شامل ضریب تعیین ( $R^2$ )، جذر میانگین مربعات خطا (RMSE) و خطای مطلق میانگین (MAD) برای داده‌های آموزش و تست استفاده شد. نتایج مدل‌ها نشان داد که مقادیر  $R^2$  روش‌های MLR و RF برای داده‌های آموزش به ترتیب ۰٫۸۱۴ و ۰٫۹۵۷ و برای داده‌های تست به ترتیب ۰٫۷۹۳ و ۰٫۹۶ هستند که بیانگر کارایی و دقت بالاتر روش RF در مقایسه با MLR است. به علاوه، نتایج نشان داد که مقادیر شاخص‌های RMSE و MAE در هر دو مرحله آموزش و تست برای الگوریتم RF کمتر از روش MLR هستند که بیانگر خطای کمتر الگوریتم RF و قابلیت اطمینان و دقت بالاتر آن نسبت به روش MLR است. در نهایت، از شاخص اهمیت متغیر برای تعیین مهم‌ترین و کم‌اهمیت‌ترین پارامترهای تاثیرگذار بر حداکثر نشست سطح زمین استفاده شد. نتایج نشان داد که از بین پارامترهای ورودی، پارامتر گپ (g) و مقاومت برشی زهکشی نشده خاک ( $C_u$ ) به ترتیب بیشترین و کمترین تاثیر را بر حداکثر نشست سطح زمین دارند.

## ۸- مراجع

- [1] Herzog, M. (1985). "Surface subsidence above shallow tunnels (in German)". Bautechnik, 62: 375-377.
- [2] O'Reilly, M. P., and New, B. M. (1982). "Settlement above tunnels in the United Kingdom-their magnitude and prediction". In: Proceedings of the tunnelling conference, Brighton, 173-181.
- [3] Peck, R. B. (1969). "Deep excavations and tunnelling in soft ground". In: 7th International conference on soil mechanics and foundation engineering, State of the Art Volume, 225-290.
- [4] Schmidt, B. (1969). "A method of estimating surface settlement above tunnels constructed in soft ground". Canadian Geotechnical Journal, 20: 11-22.
- [5] Vermeer, P. A., and Bonnier, P. G. (1991). "Pile settlements due to tunnelling". In: 10th European



شکل ۷: مقادیر حداکثر نشست اندازه‌گیری و پیش‌بینی شده با استفاده از مدل‌های MLR و RF در مرحله تست



شکل ۸: اهمیت متغیرهای ورودی تاثیرگذار بر حداکثر نشست سطح زمین در الگوریتم RF

این تحلیل، مقادیر اهمیت متغیرهای ورودی محاسبه شده و بین صفر و ۱ مرتب می‌شوند. با استفاده از این روش، مقادیر اهمیت ۹ پارامتر ورودی در نظر گرفته شده در مدل‌سازی حداکثر نشست سطح زمین با استفاده از الگوریتم RF محاسبه شده و در شکل ۸ نشان داده شده است. همان‌طور که مشاهده می‌شود، پارامتر گپ (g) و مقاومت برشی زهکشی نشده خاک ( $C_u$ ) به ترتیب بیشترین و کمترین تاثیر را بر مقدار حداکثر نشست سطح زمین دارند.

## ۷- نتیجه‌گیری

در این مقاله، با استفاده از دو روش رگرسیون خطی چندگانه

- 197-209.
- [17] Moghtader, T., Sharafati, A., Naderpour, H., and Gharouni Nik, M. (2023). "Estimating Maximum Surface Settlement Caused by EPB Shield Tunneling Utilizing an Intelligent Approach". Buildings, 13(4): 1051. DOI: 10.3390/buildings13041051.
- [18] Neaupane, K. M., and Adhikari, N. R. (2002). "Application of neural network for the prediction of settlement above tunnels". Engineering Journal of Research and Development, 13(1): 9-18.
- [19] Neaupane, K. M., and Adhikari, N. R. (2006). "Prediction of tunneling-induced ground movement with the multi-layer perceptron". Tunnelling and Underground Space Technology, 21: 151-159.
- [20] Ocak, I., and Seker, S. E. (2013). "Calculation of surface settlements caused by EPBM tunneling using artificial neural network, SVM, and Gaussian processes". Environmental Earth Sciences, 70: 1263-1276.
- [21] Santos Jr, O. J., and Celestino, T. B. (2008). "Artificial neural networks analysis of São Paulo subway tunnel settlement data". Tunnelling and Underground Space Technology, 23: 481-491.
- [22] Zhou, J., Shi, X., Du, K., Qiu, X., Li, X., and Mitri, H. S. (2016). "Feasibility of Random-Forest Approach for Prediction of Ground Settlements Induced by the Construction of a Shield-Driven Tunnel". International Journal of Geomechanics, 17(6).
- [23] Zhang, K., Lyu, H.M., Shen, S.L., Zhou, A., and Yin, Z.Y. (2020). "Evolutionary hybrid neural network approach to predict shield tunneling induced ground settlements". Tunnelling and Underground Space Technology, 106-103594.
- [24] Ghiasi, V., and Koushki, M. (2020). "Numerical and Artificial Neural Network Analyses of Ground Surface Settlement of Tunnel in Saturated Soil". SN Applied Sciences, 2(5): 939. DOI: <https://doi.org/10.1007/s42452-020-2742-z>.
- [25] Samadi, H., Hassanpour, J., and Farrokh, E. (2021). "Maximum Surface Settlement Prediction in EPB TBM Tunneling Using Soft Computing Techniques". Journal of Physics: Conference Series, 1973(1): 012195. DOI: <https://doi.org/10.1088/1742-6596/1973/1/012195>.
- [26] Aswathy, M. S., Vinoth, M., and Mittal, A. (2021). "Impact of governing factors on prediction of tunneling induced surface settlement in young alluvium deposit". Indian Geotechnical Journal, 52(1): 13-27. DOI: <https://doi.org/10.1007/s40098-021-00561-4>.
- [27] Zhang, W. G., Li, H. R., Wu, C. Z., Li, Y. Q., Liu, Z. Q., and Liu, H. L. (2021). "Soft computing approach Conference on Soil mechanics and foundation engineering, Florence, Balkema, 2: 869-872.
- [6] Chou, W. I., and Bobet, A. (2002). "Predictions of ground deformations in shallow tunnels in clay". Tunnelling and Underground Space Technology, 17: 3-19.
- [7] Loganathan, N., and Poulos, H. G. (1998). "Analytical prediction for tunneling-induced ground movements in clays". Journal of Geotechnical and Geoenvironmental Engineering, 124(9): 846-856.
- [8] Park, K. H. (2005). "Analytical solution for tunneling-induced ground movement in clays". Tunnelling and Underground Space Technology, 20(3): 249-261.
- [9] Verruijt, A., and Booker, J. R. (1996). "Surface settlement due to deformation of a tunnel in an elastic half plane". Geotechnique, 46(4): 753-756.
- [10] Addenbrooke, T. I., and Potts, D. M. (2001). "Finite element analysis of St. James Park greenfield reference site". In: Burland, J. B., Standing, J. R., and Jardine, F. M. (Eds), Building Response to Tunnelling, Vol. 1, Thomas Telford, London, 177-194.
- [11] Chakeri, H., Ozcelik, Y., and Unver, B. (2013). "Effects of important factors on surface settlement prediction for metro tunnel excavated by EPB". Tunnelling and Underground Space Technology, 36: 14-23.
- [12] Chakeri, H., and Unver, B. (2014). "A new equation for estimating the maximum surface settlement above tunnels excavated in soft ground". Environmental Earth Sciences, 71: 3195-3210.
- [13] Ercelebi, S. G., Copur, H., and Ocak, I. (2011). "Surface settlement predictions for Istanbul Metro tunnels excavated by EPB-TBM". Environmental Earth Sciences, 62(2): 357-365. DOI: 10.1007/s12665-010-0530-6.
- [14] Hasanipanah, M., Noorian-Bidgoli, M., Jahed Armaghani, D., and Khamesi, H. (2016). "Feasibility of PSOANN model for predicting surface settlement caused by tunneling". Engineering with Computers, 32: 705-715.
- [15] Mahmoodzadeh, A., Mohammadi, M., Daraci, A., Farid Hama Ali, H., Kameran Al-Salihi, N., and Mohammed Dler Omer, R. (2020). "Forecasting maximum surface settlement caused by urban tunneling". Automation in Construction, 120: 103375. DOI: <https://doi.org/10.1016/j.autcon.2020.103375>.
- [16] Moghaddasi, M. R., and Noorian-Bidgoli, M. (2018). "ICA-ANN, ANN and multiple regression models for prediction of surface settlement caused by tunneling". Tunnelling and Underground Space Technology, 79:



سیستم‌های هوشمند در مهندسی برق، دوره چهارم، شماره ۲، ص ۱۴-۱.

for prediction of surface settlement induced by earth pressure balance shield tunneling". *Underground Space*, 6(4): 353-363. DOI: <https://doi.org/10.1016/j.undsp.2019.12.003>.

- <sup>1</sup> Peck
- <sup>2</sup> Schmidt
- <sup>3</sup> Herzog
- <sup>4</sup> Loganathan
- <sup>5</sup> Poulos
- <sup>6</sup> Chou
- <sup>7</sup> Bobet
- <sup>8</sup> Park
- <sup>9</sup> Addenbrooke
- <sup>10</sup> Potts
- <sup>11</sup> Ercelebi
- <sup>12</sup> Chakeri
- <sup>13</sup> Unver
- <sup>14</sup> Neaupane
- <sup>15</sup> Adhikari
- <sup>16</sup> Santos
- <sup>17</sup> Celestino
- <sup>18</sup> Ocak
- <sup>19</sup> Seker
- <sup>20</sup> Zhou
- <sup>21</sup> Zhang
- <sup>22</sup> Aswathy
- <sup>23</sup> Gradient boosting
- <sup>24</sup> Mean tunnel SPT value
- <sup>25</sup> Kim
- <sup>26</sup> Li
- <sup>27</sup> Mu
- <sup>28</sup> Extra trees regression
- <sup>29</sup> Wang
- <sup>30</sup> Empirical mode decomposition (EMD)
- <sup>31</sup> Chaotic adaptive sparrow search algorithm (CASSA)
- <sup>32</sup> Extreme learning machine (ELM)
- <sup>33</sup> Kong
- <sup>34</sup> Shape ovalization
- <sup>35</sup> Bouyant unit weight

- [28] Dongku, K., Kwon, K., Pham, K., Oh, J.-Y., and Choi, H. (2022). "Surface Settlement Prediction for Urban Tunneling Using Machine Learning Algorithms with Bayesian Optimization". *Automation in Construction*, 140: 104331. DOI: <https://doi.org/10.1016/j.autcon.2022.104331>.
- [29] Li, C., Li, J., Shi, Z., Li, L., Li, M., Jin, D., and Dong, G. (2022). "Prediction of Surface Settlement Induced by Large-Diameter Shield Tunneling Based on Machine-Learning Algorithms". *Geofluids*, 2022: 1-13. DOI: <https://doi.org/10.1155/2022/4174768>.
- [30] Hussaine, S. M, and Mu, L. (2022). "Intelligent Prediction of Maximum Ground Settlement Induced by EPB Shield Tunneling Using Automated Machine Learning Techniques". *Mathematics*, 10(24): 4637. DOI: <https://doi.org/10.3390/math10244637>.
- [31] Wang, Y., Dai, F., Jia, R., Wang, R., Sharifi, H., and Wang, Z. (2023). "A novel combined intelligent algorithm prediction model for the tunnel surface settlement". *Scientific Reports*, 13(1): 9845. DOI: <https://doi.org/10.1038/s41598-023-37028-w>.
- [32] Kong, F. C., Lu, D. C., Ma, Y. D., Tian, T., Yu, H. T., and Du, X. L. (2023). "Novel hybrid method to predict the ground-displacement field caused by shallow tunnel excavation". *Science China Technological Sciences*, 66: 101-114. DOI: <https://doi.org/10.1007/s11431-022-2079-8>.
- [33] Balan, B., Mohaghegh, S., and Ameri, S. (1995). "State-of-Art- in permeability determination from well log data: Part I- A comparative study, Model development". *Society of Petroleum Engineers (SPE)*, 30978: 17-25.
- [34] Hastie, T., Friedman, J., and Tibshirani, R. (2009). "The elements of Statistical Learning: Data Mining, Inference, and prediction". 2nd Edition, Springer, New York, pp. 746. ISBN: 9780387848570. DOI: <https://doi.org/10.1007/978-0-387-84858-7>.
- [35] Breiman, L. (2001). "Random forest". *Machine Learning*, 45(1): 5-32.

[۳۶] ابراهیم خانی، س.، افضلی، م.، شکوهی، ع.؛ ۱۳۹۰؛ "پیش‌بینی و بررسی عوامل تصادفات جاده‌ای با استفاده از الگوریتم‌های داده‌کاوی". فصلنامه دانش انتظامی زنجان، دوره اول، شماره ۱، ص ۱۲۷-۱۱۱.

[۳۷] فرشاد، م.، ساده، ج.؛ ۱۳۹۲؛ "مکان‌یابی خطای اتصال کوتاه در خطوط انتقال جریان مستقیم ولتاژ بالا با استفاده از شبکه عصبی، رگرسیون تعمیم یافته و الگوریتم جنگل تصادفی".

- <sup>36</sup> Multiple Linear Regression (MLR)
- <sup>37</sup> Random Forest (RF) Algorithm
- <sup>38</sup> Nonparametric statistical methods
- <sup>39</sup> Variable importance (VI) index
- <sup>40</sup> Gani importance index
- <sup>41</sup> Learning sample (LS)
- <sup>42</sup> Out-of-bag (OOB)