این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

Journal of Natural Environmental Hazards, Vol.---, Issue ----, ------ ------

Performance Comparison of Logistic Regression and Machine Learning Algorithms for Rockfall Susceptibility Mapping (Case Study: Firuzkuh Road, Pol-e Sefid–Gaduk)

Paria Esmailnejad Nasirkandi[1], Eisa Jokar Sarhangi[2*]

1. *MSc. in Environmental Hazards, Faculty of Humanities, University of Mazandaran, Babolsar, Iran*
2. Corresponding Author, *Associate Professor, Department of Geography, University of Mazandaran, Babolsar, Iran*
Email: *e.jokar@umz.ac.ir* ***ORCID:*** *https://orcid.org/0000-0002-7737-5148*

Article Info

Article type:
Research Article

Article history:
Received: ----------
Revised: ----------
Accepted: ----------
Published: ----------

Keywords:
Rockfall,
Logistic Regression,
Artificial Neural Network,
Random Forest,
Firuzkuh Road.

ABSTRACT

The Firuzkuh Road, which traverses the Central Alborz Mountains, is highly susceptible to rockfall occurrences. This study aimed to produce a rockfall susceptibility map and compare the performance of Logistic Regression (LR), Artificial Neural Network (ANN), and Random Forest (RF) models along the Pol-e Sefid–Gaduk section of the Firuzkuh Road. Nine conditioning factors, including elevation, slope, slope aspect, slope curvature, lithology, land use, distance to faults, distance to roads, and distance to rivers, were considered. Rockfall presence and absence data were divided into training (70%) and testing (30%) datasets using a spatially stratified sampling approach for model development and validation. All three models were implemented in Python, and their performance was evaluated using Accuracy, Recall, and the Area Under the Receiver Operating Characteristic Curve (AUC). Variable importance analysis based on the Mean Decrease in Impurity (MDI) and SHAP values indicated that elevation and lithology were the most influential factors controlling rockfall occurrence. Furthermore, advanced statistical evaluations using the DeLong test and bootstrap confidence intervals confirmed that the Random Forest model, with an AUC of 0.993 and an Accuracy of 0.947, achieved a statistically significant and definitive superiority over both the Artificial Neural Network and Logistic Regression models in rockfall susceptibility zonation. The resulting susceptibility map can serve as a valuable tool for identifying rockfall-susceptible areas, prioritizing slope stabilization measures, and improving safety along the Pol-e Sefid–Gaduk section of the Firuzkuh Road.

آماده انتشار (بدون شماره مجله)

How to Cite: Last Name, Initial., Last Name, Initial., & Last Name, Initial. (2021). Title of paper. *Journal of Natural Environmental Hazards*, -- (--), ----.

© The Author/Authors

DOI: 0000000000000000000

Publisher: University of Sistan and Baluchestan

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

EXTENDED ABSTRACT

INTRODUCTION

Rockfall is one of the most important geomorphological processes in mountainous regions and steep slopes, which, due to its sudden and destructive nature, poses a serious threat to human lives, infrastructure, and transportation networks. The occurrence of this phenomenon is controlled by various geological, topographic, and land-use factors. Lithology, slope gradient, slope aspect, elevation, vegetation cover, land use type, and distance from roads and drainage networks can play a decisive role in slope instability and rockfall occurrence. The Firuzkuh Road is one of the main transportation corridors in northern Iran and, owing to its mountainous topography, active tectonic setting, and complex geomorphological conditions, is continuously exposed to various natural instabilities, including rockfalls. Rockfall occurrences in the study area cause casualties, economic losses, road blockages, and traffic disruptions. Previous studies indicate that susceptibility mapping and spatial assessment of rockfalls remain active research areas in which the application of advanced models and performance comparison within GIS environments can improve modeling accuracy. The present study aimed to compare the performance of the Logistic Regression (LR), Artificial Neural Network (ANN), and Random Forest (RF) in rockfall susceptibility mapping along the Firuzkuh Road between Pol-e Sefid and Gaduk, and to analyze the relative importance of the factors influencing rockfall occurrence.

DATA AND METHODOLOGY

In this study, a statistical model, Logistic Regression (LR), and two machine learning algorithms, Artificial Neural Network (ANN) and Random Forest (RF), were employed to produce rockfall susceptibility maps along the Firuzkuh Road and to compare their predictive performance. Data processing and spatial analyses were conducted in ArcGIS 10.3, while modeling procedures were performed using the Python programming language and the Pandas and Scikit-learn libraries. The dependent variable in this study was defined as the occurrence or non-occurrence of rockfall as a binary variable with values of one (occurrence) and zero (non-occurrence). To prepare the rockfall distribution map, several field surveys were conducted throughout the study area, and the locations of rockfall occurrences were recorded using a GPS device. The collected dataset was spatially stratified and divided into two groups, where 70% of the data (396 rockfall occurrence points and 397 non-occurrence points) were used for model training and development, while the remaining 30% (171 occurrence points and 170 non-occurrence points) were used for model evaluation and validation.

آمادگی انتشار (صدور شماره مجله)

Independent variables consisted of conditioning factors affecting rockfall occurrence, which were selected based on previous studies and the geomorphological characteristics of the study area. Continuous variables (such as slope, elevation, and distance from rivers) were used as continuous values without reclassification, whereas categorical variables (such as lithology and land use) were considered as nominal variables and, to prevent creating artificial ordinal relationships between classes, were converted into binary variables using the One-Hot Encoding method to enable their use in modeling procedures. After training and implementing the Logistic Regression, Artificial Neural Network and Random Forest models in Python, rockfall occurrence probability was calculated for the entire study area. Model accuracy and performance were evaluated using statistical indices derived from the confusion matrix and the Receiver Operating Characteristic (ROC) curve.

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

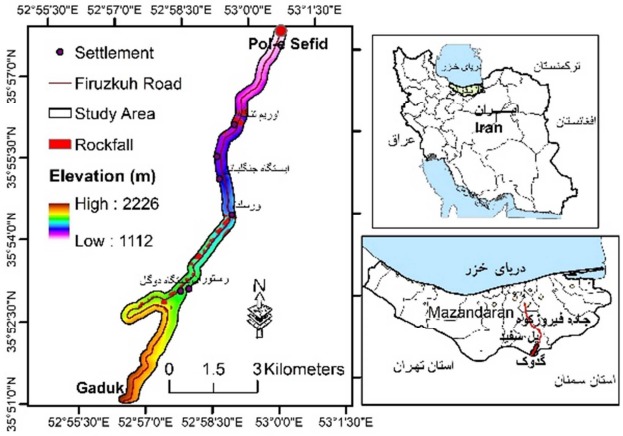


Figure 1: Geographical Location Map of the Study Area in Mazandaran Province Iran

RESULTS AND DISCUSSION

The performance evaluation results of the Logistic Regression, Artificial Neural Network and Random Forest models indicate that all three models have very high capability for rockfall susceptibility mapping. However, considerable differences were observed in their predictive accuracy and discrimination ability. The Random Forest model showed better performance, with an Accuracy value of 0.947 and an AUC value of 0.993. Furthermore, the Recall value of 0.959 indicates that the model successfully identified most of the actual rockfall occurrences. The Artificial Neural Network also demonstrated highly competitive performance (AUC = 0.976, Accuracy = 0.941). However, statistical significance testing (DeLong test) and bootstrap confidence intervals confirmed the definitive and statistically significant superiority of the Random Forest model over both ANN ($p < 0.001$) and Logistic Regression ($p < 0.001$). Analysis of the Receiver Operating Characteristic (ROC) curves further demonstrated that all three models exhibited acceptable discrimination capability for identifying rockfall-prone areas. However, the ROC curve of the Random Forest model remained above those of the Artificial Neural Network and Logistic Regression models across most false positive rate (FPR) values, indicating its superior ability to discriminate between rockfall and non-rockfall areas.

The prioritization of variables controlling rockfall occurrence was performed based on the Mean Decrease Impurity (MDI) index and SHAP values within the Random Forest model. The results indicate that rockfall occurrence in the study area is mainly controlled by morphometric factors, particularly elevation and lithology. To produce the final susceptibility maps, the trained models were applied to the entire study area. Accordingly, approximately 24.64% of the study area was classified as having high and very high susceptibility by the Logistic Regression model, compared with 24.18% by the Artificial Neural Network model and 19.24% by the Random Forest model. These findings indicate that the Random Forest model identified high-susceptibility areas selectively and in a concentrated manner (with higher spatial accuracy).

Table 1: Performance Evaluation of the Logistic Regression, Artificial Neural Network, and Random Forest Models for Rockfall Susceptibility Mapping Using Different Statistical Metrics

Model	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-score	AUC
Logistic Regression	157	153	17	14	0.909	0.902	0.918	0.910	0.957
Artificial Neural Network	162	159	11	9	0.941	0.936	0.947	0.942	0.976
Random Forest	164	159	11	7	0.947	0.937	0.959	0.948	0.993

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

CONCLUSION

The comparison of Logistic Regression, Artificial Neural Network, and Random Forest models using different statistical indices and ROC curve analysis demonstrated that although all three models demonstrated predictive performance in rockfall susceptibility mapping, the machine learning algorithms, particularly Random Forest, provided higher predictive accuracy and better spatial discrimination owing to their ability to model complex nonlinear relationships among the influencing factors. These findings highlight the importance of applying machine learning algorithms in rockfall susceptibility mapping studies conducted within complex geomorphological environments. According to the produced susceptibility map, approximately 160.7 hectares of the study area along the Pol-e Sefid-Gaduk section were classified within the high and very high rockfall susceptibility classes. Therefore, greater attention should be given to the management of this phenomenon in planning and maintenance practices along this mountainous corridor. The produced susceptibility map can serve as a decision-support tool for identifying susceptible areas, prioritization of stabilization measures, and infrastructure planning within the region.

ETHICAL CONSIDERATIONS

Conflict of Interest Statement: The authors declare no conflict of interest.

Ethical Statement: This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

References [in Persian]

Abedini, M., Mozafari, H., & Faal Naziri, M. (2022). Investigating and comparing the effectiveness of information value models and frequency ratio coefficient and Shannon's entropy in zoning rock fall risk (case study: Zanjan-Teham-Taram Road). *Geographical Studies of Mountainous Regions*, 3(1), 55-75. https://doi.org/10.52547/gsma.3.1.55. [In Persian]

Asghari Saraskanrood, S., Hasan Zadeh, R. and Raoofi, S. (2020). Investigation of Factors Influencing Rockfall and Its Zoning with Logistic Regression in Ali Bayad Basin of Horand. Abstract. *Hydrogeomorphology*, 7(23), 38-21. https://doi.org/10.22034/hyd.2020.11004. [In Persian]

Asghari Saraskanrood, S., & Mozafari, H. (2020). Estimate and Comparison of Frequency Ratio and Network Analysis Models in Rockfall Zoning (Case study: Zanjan road to Taham and Tarom). *Journal of Spatial Analysis and Environmental Hazards*, 6(4), 123-142. https://doi.org/10.29252/jsaeh.6.4.123. [In Persian]

Esfandyari Darabad, F., Rahimi, M., & Vahabzadeh, M. (2024). Rockfall risk zoning on the Khalkhal to Shahroud road using multilayer perceptron algorithm. *Journal of Geography and Environmental Hazards*, 13(4), 1-26. https://doi.org/10.22067/geoeh.2024.87847.1482. [In Persian]

Jokar sarhangi, E., & Alizadeh, A. (2014). An Investigation into the Efficiency of Frequency Ratio and Information Value Models in Rockfall Hazard Zonation (Case Study: Haraz Valley). *Geographical planning of space quarterly journal*, 4(11), 151- 163. https://gps.gu.ac.ir/article_7402.html?lang=en. [In Persian]

Momeni, A. and Heidari, M. (2021). Rockfall hazard analysis in the Ganjnameh tourist area, west of Iran. *Journal of Natural Environmental Hazards*, 10(28), 1-14. https://doi.org/10.22111/jneh.2021.31304.1553. [In Persian]

Rahimi, M., Esfandiari, F., Zakizadeh, O., & Vahabzadeh Zargari, M. (2025). Identification of Rockfall Hazard Zones Along the Meshginshahr-Ahar Road Using Artificial Neural Networks. *Journal of Natural Environmental Hazards*, 14(46), 23–42. https://doi.org/10.22111/jneh.2025.50028.2072. [In Persian]

References [in English]

Bhagat, N. K., Mishra, A. K., Singh, R. K., Sawmliana, C., & Singh, P. K. (2022). Application of logistic regression, CART, and random forest techniques in prediction of blast-induced slope failure during reconstruction of railway rock-cut slopes. *Engineering Failure Analysis*, 137, 106230. https://doi.org/10.1016/j.engfailanal.2022.106230.

Breiman, L., 2001. Random forests. *Machine Learning*, 45(1), 5-32. https://doi.org/10.1023/a:1010933404324.

Chang, Z., Huang, F., Huang, J., Jiang, Sh.H., Yuting Liu, Y., Meena, S.R., & Catani, F. (2023). An update of landslide susceptibility prediction from the perspective of space and time. Geoscience Frontiers, 14(5), 101619. https://doi.org/10.1016/j.gsf.2023.101619.

Fanos, A. M., & Pradhan, B. (2019). A Spatial Ensemble Model for Rockfall Source Identification from High-Resolution LiDAR Data and GIS. *IEEE Access*, 7, 74570–74585. https://doi.org/10.1109/ACCESS.2019.2919977.

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

مجله علمی پژوهشی مخاطرات محیط طبیعی، دوره ----، شماره ----- ، -------- -----

مقایسه کارایی رگرسیون لجستیک و الگوریتم‌های یادگیری ماشین در پهنه‌بندی حساسیت وقوع ریزش سنگ (مطالعه موردی: جاده فیروزکوه، پل سفید–گدوک)

پریا اسمعیل نژاد نصیرکندی[1]، عیسی جوکار سرهنگی[2]*

۱. دانش آموخته کارشناسی ارشد مخاطرات محیطی، دانشگاه مازندران، بابلسر، ایران

۲. دانشیار گروه جغرافیا، دانشگاه مازندران، بابلسر، ایران (**نویسنده مسئول**)

ایمیل: e.jokar@umz.ac.ir *ORCID: https://orcid.org/0000-0002-7737-5148*

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

تاریخ دریافت: --/--/--
تاریخ ویرایش: --/--/--
تاریخ پذیرش : --/--/--
تاریخ انتشار : --/--/--

واژه‌های کلیدی:
ریزش سنگ،
رگرسیون لجستیک،
شبکه عصبی مصنوعی،
جنگل تصادفی،
جاده فیروزکوه.

چکیده

جاده فیروزکوه به دلیل عبور از رشته‌کوه‌های البرز مرکزی همواره در معرض وقوع ریزش سنگ قرار دارد. این پژوهش با هدف تهیه نقشه پهنه‌بندی حساسیت وقوع ریزش سنگ و مقایسه کارایی مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی در محدوده پل سفید تا گدوک انجام شده است. برای این منظور، عوامل موثر شامل ارتفاع، شیب، جهت دامنه، شکل دامنه، زمین‌شناسی، کاربری اراضی، فاصله از گسل، فاصله از جاده و فاصله از رودخانه مورد بررسی قرار گرفت. برای آموزش و ارزیابی مدل‌ها از داده‌های وقوع و عدم وقوع ریزش سنگ استفاده شد که با استفاده از روش نمونه‌برداری طبقه‌بندی‌شده مکانی، ۷۰ درصد برای آموزش و مدل‌سازی و ۳۰ درصد برای آزمون و اعتبارسنجی مدل‌ها اختصاص یافت. هر سه مدل در محیط پایتون اجرا شده و عملکرد مدل‌ها با استفاده از شاخص‌های آماری Accuracy، Recall و مساحت زیر منحنی AUC ارزیابی شد. نتایج تحلیل اهمیت متغیرهای موثر بر اساس میانگین کاهش ناخالصی (MDI) و مقادیر SHAP نشان داد که سهم نسبی ارتفاع و جنس‌سنگ در وقوع ریزش بیشتر است. همچنین ارزیابی‌های پیشرفته آماری به کمک آزمون دلانگ و فواصل اطمینان بوت‌استرپ نشان داد که مدل جنگل تصادفی با مقدار AUC برابر با ۰/۹۹۳ و دقت ۰/۹۴۷ در مقایسه با شبکه عصبی مصنوعی و رگرسیون لجستیک، برتری معنادار و قطعی در پهنه‌بندی حساسیت وقوع ریزش سنگ دارد. نقشه حساسیت حاصل می‌تواند به‌عنوان ابزاری برای شناسایی نواحی مستعد وقوع ریزش سنگ، اولویت‌بندی اقدامات پایدارسازی و افزایش ایمنی محور پل سفید-گدوک مورد استفاده قرار گیرد.

آماده انتشار (ویرایش شماره مجله)

استناد: نام خانوادگی، نام؛ نام خانوادگی، نام؛ و نام خانوادگی، نام (۱۴۰۰). عنوان مقاله. *مخاطرات محیط طبیعی*، -- (--)، ----

© نویسنده/ نویسندگان
DOI: 00000000000000000000000000000
ناشر: دانشگاه سیستان و بلوچستان

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

مقدمه

ریزش سنگ یکی از مهم‌ترین فرآیندهای ژئومورفولوژیک در مناطق کوهستانی و دامنه‌های شیب‌دار است که به‌دلیل ماهیت ناگهانی و مخرب خود، تهدیدی جدی برای جان انسان‌ها، زیرساخت‌ها و شبکه‌های حمل و نقل محسوب می‌شود (مومنی و حیدری، ۱۴۰۰؛ عابدینی و همکاران، ۱۴۰۱؛ نگوین[1] و همکاران، ۲۰۲۴). وقوع این پدیده تحت تأثیر عوامل متعدد زمین‌شناسی، توپوگرافی و کاربری اراضی قرار دارد. جنس سنگ بستر، شیب و جهت دامنه، ارتفاع، پوشش گیاهی، نوع کاربری اراضی و فاصله از جاده و آبراهه‌ها می‌توانند نقش تعیین‌کننده‌ای در ناپایداری دامنه‌ها و وقوع ریزش سنگ ایفا نمایند. در ایران به دلیل توپوگرافی عمدتاً کوهستانی و فعالیت‌های تکتونیکی بالا، جاده‌های کوهستانی همواره در معرض وقوع ریزش سنگ قرار دارند (اصغری سراسکانرود و همکاران، ۱۳۹۹). این حوادث منجر به خسارات مالی فراوان و تلفات جانی می‌شوند که لزوم شناسایی مناطق مستعد و تهیه نقشه‌های پهنه‌بندی حساسیت ریزش سنگ را دوچندان می‌سازد. با توجه به افزایش روزافزون جمعیت و گسترش فعالیت‌های انسانی در مناطق کوهستانی، نیاز به شناخت دقیق مناطق مستعد ریزش و اتخاذ راهبردهای پیشگیرانه و مدیریتی برای کاهش خسارات ناشی از این پدیده ضروری است (اسفندیاری درآباد و همکاران، ۱۴۰۳). با این حال، شناسایی دقیق مناطق منشأ ریزش سنگ یک چالش مهم در مطالعات ژئومورفولوژی کاربردی محسوب می‌شود؛ به‌ویژه در مناطقی که سایر اشکال حرکات توده‌ای مانند زمین‌لغزش‌های کم‌عمق و جریان‌های واریزه‌ای با ویژگی‌های مورفومتریک مشابه نیز وجود دارند (فانوس[2] و پرادهان[3]، ۲۰۱۹). همچنین، به دلیل تغییرپذیری زمانی و مکانی این پدیده‌ها، پیش‌بینی دقیق وقوع آن‌ها دشوار بوده و مدل‌سازی را با مشکل مواجه می‌سازد (فانوس و همکاران، ۲۰۲۰).. در این چارچوب، تهیه نقشه‌های پهنه‌بندی حساسیت به‌عنوان یکی از ابزارهای مهم برای شناسایی مکانی نواحی مستعد وقوع ریزش سنگ محسوب می‌شوند. این نقشه‌ها برای اولویت‌بندی و انتخاب مناسب اقدامات کاهشی موثر هستند و به مهندسان و برنامه‌ریزان کمک می‌کنند تا مکان‌های مناسبی را برای پروژه‌های عمرانی انتخاب کرده و در مدیریت ناپایداری و طراحی کاربری اراضی به کار گرفته شوند (مختاری و همکاران، ۲۰۲۰؛ فرراری[4] و همکاران، ۲۰۱۶).

آماده انتشار (بدون شماره مجله)

به‌منظور پهنه‌بندی حساسیت ریزش سنگ، رویکردهای متعددی در دهه‌های اخیر توسعه پیدا کرده که از جمله‌ی آنها می‌توان به مدل‌های آماری و الگوریتم‌های یادگیری ماشین اشاره نمود. در میان مدل‌های آماری، روش‌هایی نظیر نسبت فراوانی، ارزش اطلاعات و رگرسیون لجستیک در مطالعات مختلف به کار گرفته شده‌اند (جوکار سرهنگی و علیزاده، ۱۳۹۳؛ اصغری سراسکانرود و مظفری، ۱۳۹۸). رگرسیون لجستیک به‌دلیل ساختار پارامتریک و قابلیت تفسیر ضرایب، یکی از پرکاربردترین مدل‌ها در پهنه‌بندی حساسیت حرکات توده‌ای و ریزش سنگ است. در این راستا، روستایی و همکاران (۱۳۹۴) از مدل رگرسیون لجستیک در پهنه‌بندی وقوع ناپایداری‌های دامنه‌ای جاده کوهستانی تنگه دره دیز استفاده کردند. نتایج آنها نشان داد که عوامل شیب، فاصله از گسل و طبقات ارتفاعی به ترتیب مهم‌ترین متغیرهای مؤثر در وقوع ناپایداری‌های دامنه‌ای در این منطقه هستند. همچنین مشخص شد که بیشترین میزان ناپایداری‌ها در شیب‌های ۶۷ الی ۹۰ درصد و در ارتفاعات ۱۸۰۰ تا ۲۲۰۰ متری رخ می‌دهند. در

[1] - Nguyen
[2] - Fanos
[3] - Pradhan
[4] - Ferrari

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

پژوهشی دیگر، شیرزادی و همکاران (۱۳۹۴) به مقایسه مدل‌های رگرسیون لجستیک، نسبت فراوانی و تحلیل سلسله مراتبی در تهیه نقشه حساسیت ریزش سنگ در گردنه صلوات آباد کردستان پرداختند. بررسی مدل ها نشان داد که رگرسیون لجستیک نسبت به سایر مدل‌ها برای شناسایی مناطق مستعد ریزش سنگ از صحت بالاتری برخوردار است. مختاری و همکاران (۲۰۲۰) برای تهیه نقشه حساسیت ریزش سنگ به مقایسه سه مدل شبکه عصبی مصنوعی، رگرسیون لجستیک و نسبت فراوانی در منطقه تفت ایران مرکزی پرداختند. نتایج آنها نشان داد که هر سه مدل عملکرد خوبی در پیش‌بینی نواحی حساس به ریزش سنگ نشان دادند. اصغری سراسکانرود و همکاران (۱۳۹۹) نیز در بررسی عوامل موثر بر ریزش سنگ و پهنه‌بندی آن در حوضه آبریز علی‌آباد هوراند نشان دادند که مدل رگرسیون لجستیک عملکرد مناسبی دارد. ژانگ[1] و همکاران (۲۰۲۳) با ترکیب مدل‌های کمّی اطلاعات و ضریب قطعیت با رگرسیون لجستیک، حساسیت ناحیه کوهستانی هنگ‌داون چین به ریزش‌های سنگ و حرکات توده‌ای را با استفاده از هشت عامل مؤثر ارزیابی کردند. مقدار AUC مدل‌های ترکیبی در این تحقیق به‌ترتیب ۰/۷۹۹ و ۰/۷۷۲ به دست آمد که نشان‌دهنده دقت قابل قبول هر دو مدل در برآورد حساسیت وقوع ریزش سنگ است. به این ترتیب، از مدل رگرسیون لجستیک در بسیاری از پژوهش‌ها استفاده شده و دقت آن مورد تأیید قرار گرفته است. با این حال، محدودیت اصلی این مدل در فرض خطی بودن رابطه بین متغیرهای مستقل و احتمال وقوع پدیده‌های مورد مطالعه است که ممکن است موجب کاهش دقت پیش‌بینی شود.

پیشرفت‌های اخیر در یادگیری ماشین (ML) رویکردهای جدید و قدرتمندی را برای تهیه نقشه‌های حساسیت ناپایداری دامنه‌ای ارائه کرده‌اند. در این میان، شبکه عصبی مصنوعی (ANN) به دلیل توانایی منحصربه‌فرد در تحلیل روابط غیرخطی و پیچیده بین عوامل محیطی و رخدادهای ناپایداری دامنه، جایگاه ویژه‌ای در پژوهش‌های محیطی پیدا کرده است (نه‌نه‌کران[2] و همکاران، ۲۰۲۲). توسعه مدل‌های مبتنی بر شبکه عصبی، به‌ویژه الگوریتم پرسپترون چندلایه (MLP) در پهنه‌بندی حساسیت وقوع پدیده‌های مختلف از جمله ریزش سنگ و جریان‌های واریزه‌ای، دقت بالایی را نشان داده است. در این راستا، فانوس و همکاران (۲۰۱۹) با تلفیق شبکه عصبی در یک ساختار ترکیبی و بهره‌گیری از داده‌های با تفکیک بالا موفق شدند منابع تولید ریزش سنگ را با دقت ۰/۹۴ شناسایی کنند که این امر بیانگر کارایی بالای این الگوریتم در مواجهه با داده‌های پیچیده محیطی است. نه‌نه‌کران و همکاران (۲۰۲۲) در مطالعه‌ای که در مناطق کوهستانی استان البرز انجام دادند، مدل پرسپترون چندلایه را به عنوان دقیق‌ترین الگوریتم در مقایسه با ماشین بردار پشتیبان و جنگل تصادفی معرفی کردند. اسفندیاری درآباد و همکاران (۱۴۰۳) از این مدل در جاده خلخال به شاهرود استفاده کردند که نتایج آنها بیانگر اعتبار بسیار بالا و پایداری مدل در پیش‌بینی رخدادهای آتی ریزش سنگ است. رحیمی و همکاران (۱۴۰۴) نیز با تمرکز بر محور مواصلاتی مشگین‌شهر به اهر، از قابلیت‌های مدل MLP برای تحلیل تأثیر عوامل محیطی استفاده کردند. نتایج این پژوهش ضمن تأیید اعتبار بالای مدل در شناسایی مناطق با حساسیت بالاتر، بیانگر نقش مهم پارامترهای زمین‌شناسی و شیب در وقوع ریزش‌های این مسیر جاده‌ای است.

با این وجود در بسیاری از پژوهش‌ها، الگوریتم جنگل تصادفی عملکرد بالاتری از خود نشان داده است. فانوس و همکاران (۲۰۱۸) از مدل‌های مختلف یادگیری ماشین از جمله شبکه عصبی مصنوعی در تهیه نقشه احتمال وقوع ریزش استفاده کردند و به این نتیجه رسیدند که مدل جنگل تصادفی در مقایسه با سایر مدل‌ها بهترین عملکرد را در

آماده انتشار (بدون شماره مجلد)

[1] - Zhang
[2] - Nanehkaran

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

شناسایی مناطق مستعد ریزش سنگ دارد. فانوس و همکاران (۲۰۲۰) با استفاده از داده‌های با وضوح بالا و مدل‌های یادگیری ماشین، نقشه‌های احتمال وقوع ریزش سنگ را با دقت بالا تهیه کردند و نشان دادند این روش‌ها قابلیت تعمیم به مناطق مختلف را دارند. در این پژوهش مدل ترکیبی Random Forest – KNN عملکرد بهتری نسبت به سایر مدل‌ها نشان داده و دقت آن در اعتبارسنجی به ۸۶ درصد رسید. بهاگات[1] و همکاران (۲۰۲۲) به بررسی کاربرد مدل‌های رگرسیون لجستیک، درخت طبقه‌بندی و جنگل تصادفی برای پیش‌بینی ناپایداری شیب ناشی از انفجار در حین بازسازی ترانشه‌های سنگی راه‌آهن پرداختند. نتایج آنها نشان داد که هر سه مدل توانایی پیش‌بینی خوبی دارند؛ اما مدل جنگل تصادفی بالاترین عملکرد کلی را داشت. از طرفی، مدل رگرسیون لجستیک در برخی شاخص‌ها از دو مدل دیگر مناسب‌تر بود. پاک‌نژاد و همکاران (۱۴۰۲) در مدل‌سازی جریان‌های واریزه‌ای حوضه تنگراه، دقت الگوریتم جنگل تصادفی را کمی بالاتر گزارش کردند و شبکه عصبی مصنوعی را با AUC برابر با ۰/۸۸ به عنوان یکی از مدل‌های معتبر در تحلیل‌های ژئومورفولوژیک ذکر نمودند. محمودزاده و همکاران (۲۰۲۳) با استفاده از داده‌های ۳۴۴ شیب سنگی از معادن روباز مختلف ایران، عملکرد دوازده مدل یادگیری ماشین برای تخمین ضریب ایمنی شیب‌های سنگی را مورد ارزیابی قرار دادند و به این نتیجه رسیدند که مدل جنگل تصادفی با کمترین خطا در تخمین ضریب ایمنی دقیق‌تر است. نگوین و همکاران (۲۰۲۴) در یک مطالعه دیگر، داده‌های مربوط به حوادث ریزش سنگ در کل سیستم بزرگراهی تایوان را از طریق اخبار اینترنتی جمع‌آوری و با استفاده از الگوریتم جنگل تصادفی و تحلیل SHAP مدل‌سازی کردند. نتایج آنها نشان داد که عواملی مانند زاویه شیب و فاصله از مناطق حساس زمین‌شناسی، نقشی اصلی در تعیین مکان‌های مستعد ریزش سنگ ایفا می‌کنند. مدل نهایی با دقت کلی ۰/۸۵ و حساسیت ۰/۸۳، عملکردی قوی در شناسایی مکان‌های وقوع ریزش از خود به نمایش گذاشت. هی[2] و همکاران (۲۰۲۵) در پیش‌بینی وقوع ریزش در منطقه کوهستانی چانگ‌بای عملکرد چهار مدل یادگیری ماشین شامل SVM، RF، XGBoost و LightGBM را مورد مقایسه قرار دادند و به این نتیجه رسیدند که مدل بهینه‌سازی شده جنگل تصادفی بهترین عملکرد را در مقایسه با سایر الگوریتم‌ها دارد. با وجود پیشرفت‌های یادشده، همچنان در بسیاری از مطالعات، مقایسه عملکرد مدل‌های آماری و یادگیری ماشین در محورهای کوهستانی ایران محدود بوده و ارزیابی کارایی آن‌ها در مقیاس‌های جاده‌ای کمتر مورد توجه قرار گرفته است.

جاده فیروزکوه یکی از محورهای ارتباطی اصلی شمال ایران است که به دلیل توپوگرافی کوهستانی، فعالیت‌های زمین‌ساختی و شرایط ژئومورفولوژیکی پیچیده، همواره با استعداد وقوع ناپایداری‌های طبیعی متعددی از جمله ریزش سنگ مواجه است. در این منطقه، سایر انواع حرکات توده‌ای مانند زمین‌لغزش‌های کم‌عمق و جریان‌های واریزه‌ای با ویژگی‌های ژئومورفومتریک مشابه نیز رخ می‌دهند که در کنار ریزش سنگ، موجب تلفات جانی، خسارات مالی، انسداد مسیر جاده و اختلال در عبور و مرور می‌شوند. پیشینه پژوهش نیز نشان می‌دهد که تهیه نقشه‌های حساسیت و ارزیابی مکانی ریزش سنگ، یکی از زمینه‌های پژوهشی فعال است و استفاده از مدل‌های پیشرفته همراه با مقایسه عملکرد آن‌ها در محیط GIS می‌تواند دقت مدل‌سازی را افزایش دهد. با وجود کاربرد گسترده این مدل‌ها، بخش قابل توجهی از پژوهش‌های پیشین در مقیاس‌های منطقه‌ای یا حوضه‌ای انجام شده‌اند و توجه کمتری به ارزیابی همزمان این رویکردها در مقیاس یک محور جاده‌ای کوهستانی داشته‌اند. از این رو، وجه تمایز و نوآوری پژوهش حاضر علاوه بر ایجاد یک پایگاه داده متراکم حاصل از بررسی‌های میدانی در به‌کارگیری و مقایسه همزمان یک مدل آماری

آماده انتشار (بدون شماره مجله)

[1] - Bhagat
[2] - He

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

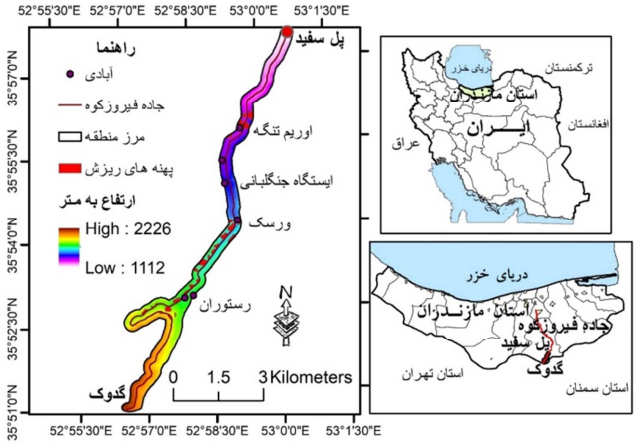
(رگرسیون لجستیک) و دو الگوریتم یادگیری ماشین (شبکه عصبی مصنوعی و جنگل تصادفی) بر اساس یک مجموعه داده یکسان از رخدادها و عوامل محیطی در مقیاس محور جاده‌ای پل‌سفید–گدوک است. نتایج حاصل از مدل‌سازی در این مقیاس، برخلاف مقیاس‌های ناحیه‌ای مستقیماً قابلیت استفاده در تصمیمات مهندسی و پایدارسازی ترانشه‌ها را دارند. علاوه بر مقایسه عملکرد مدل‌ها، اهمیت نسبی و نحوه اثرگذاری عوامل مؤثر بر وقوع ریزش سنگ نیز با بهره‌گیری از قابلیت‌های الگوریتم جنگل تصادفی و روش پیشرفته SHAP بررسی شده است تا مشخص شود کدام عوامل در شرایط ژئومورفولوژیکی پیچیده این محور نقش بیشتری در تعیین حساسیت وقوع ریزش سنگ دارند و اولویت‌بندی اقدامات مدیریتی بر مبنای آن‌ها صورت پذیرد. بر این اساس، هدف پژوهش حاضر، مقایسه کارایی مدل رگرسیون لجستیک و دو الگوریتم یادگیری ماشین شامل شبکه عصبی مصنوعی و جنگل تصادفی در تهیه نقشه پهنه‌بندی حساسیت وقوع ریزش سنگ در امتداد جاده فیروزکوه از پل‌سفید تا گدوک و تحلیل اهمیت نسبی عوامل مؤثر بر وقوع آن می‌باشد.

داده‌ها و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه در استان مازندران و شهرستان سواد کوه در دامنه شمالی رشته کوه های البرز مرکزی قرار دارد. این منطقه با مساحت ۸۳۵/۲ هکتار در موقعیت جغرافیایی "۵۵ ´۵ °۳۵ تا "۵ ´۵۰ °۳۵ عرض شمالی و "۴ ´۳۵ °۵۳ تا "۵۲ ´۵۵ °۵۲ طول شرقی واقع شده است. در شکل زیر موقعیت جغرافیایی منطقه مورد مطالعه نمایش داده شده است (شکل ۱). جاده فیروزکوه یکی از محورهای ارتباطی اصلی شمال کشور است که پس از عبور از مناطق کوهستانی البرز به سمت قائم‌شهر و در ادامه به ساری امتداد می‌یابد و نقش مهمی در حمل‌ونقل، مبادلات اقتصادی و جابه‌جایی مسافران ایفا می‌کند. طول جاده مورد مطالعه از شهر پل سفید تا انتهای حوزه استحفاظی مازندران در منطقه گدوک ۱۵/۵ کیلومتر است. ارتفاع کمینه و بیشینه منطقه از سطح دریا به ترتیب ۱۰۱۷ و ۲۲۲۶ متر می‌باشد.

آماده انتشار (بدون شماره مجله)



شکل ۱: موقعیت جغرافیایی منطقه مورد مطالعه در استان مازندران و ایران. *منبع: نگارندگان*

روش پژوهش

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

این پژوهش از نظر هدف، کاربردی و از نظر ماهیت، توصیفی-تحلیلی است که با رویکردی کمّی انجام شده و هدف آن، ارزیابی و پهنه‌بندی حساسیت وقوع ریزش سنگ در بخشی از محور کوهستانی جاده فیروزکوه می‌باشد. برای این‌منظور، از مدل آماری رگرسیون لجستیک[1] (LR) و دو الگوریتم پیشرفته یادگیری ماشین، شامل شبکه عصبی مصنوعی[2] (ANN) و جنگل تصادفی[3] (RF) استفاده شد تا ضمن مقایسه عملکرد آن‌ها، مناسب‌ترین مدل برای تهیه نقشه پهنه‌بندی حساسیت ریزش سنگ شناسایی شود. مراحل پردازش داده‌ها و تحلیل‌های فضایی در محیط ArcGIS 10.3 و مدل‌سازی با استفاده از زبان برنامه‌نویسی پایتون و کتابخانه‌های Pandas و Scikit-learn انجام گرفت.

در مرحله گردآوری داده‌ها و آماده‌سازی لایه‌ها، داده‌ها به دو دسته اصلی متغیر وابسته و متغیرهای مستقل تقسیم گردید که از منابع مختلف گردآوری شده‌اند. متغیر وابسته در این پژوهش، وقوع یا عدم وقوع پدیده ریزش سنگ است که به‌صورت یک متغیر دودویی با مقادیر یک (وقوع) و صفر (عدم وقوع) تعریف شد. برای تهیه نقشه پراکنش ریزش سنگ، بررسی‌ها و بازدیدهای مکرر میدانی از منطقه مورد مطالعه انجام گرفت و موقعیت مکانی وقوع ریزش سنگ به کمک دستگاه موقعیت‌یاب جهانی (GPS) ثبت گردید. برای تشخیص پهنه‌های ریزشی علاوه بر شواهد ژئومورفولوژیک موجود در طول مسیر جاده، از اطلاعات مهندسین اداره راه و ترابری شهرستان پل سفید و آمار و اطلاعات این اداره نیز استفاده شده است. بر اساس نقشه پراکنش ریزش، تعداد ۵۶۷ نقطه وقوع ریزش در محیط نرم‌افزار ArcGIS به صورت لایه نقطه‌ای استخراج شد. به‌منظور تعیین نقاط عدم‌وقوع علاوه بر بررسی‌های میدانی، تصاویر Google Earth نیز مورد بررسی قرار گرفت و مناطقی که فاقد شواهد قابل مشاهده از وقوع ریزش سنگ از جمله واریزه و تجمع قطعات سنگی در پای دامنه بودند و در بخش‌هایی نیز سطح دامنه با خاک و پوشش گیاهی بیشتری پوشیده شده بود؛ به‌عنوان مناطق فاقد شواهد وقوع در نظر گرفته شدند. همچنین، به‌منظور کاهش احتمال انتخاب نقاط عدم‌وقوع در مجاورت مستقیم رخدادهای ثبت‌شده، محدوده حریم ۹۰ متری پیرامون نقاط وقوع در نظر گرفته شد و نقاط واقع در این محدوده از فرآیند انتخاب حذف شدند. سپس، به تعداد نقاط وقوع یعنی ۵۶۷ نقطه، نقاط عدم‌وقوع به‌صورت تصادفی از مناطق باقی‌مانده فاقد شواهد مشاهده‌شده از وقوع ریزش انتخاب گردید. این داده‌ها شامل ۵۶۷ نقطه وقوع و ۵۶۷ نقطه عدم وقوع بودند. به منظور جلوگیری از خطای خودهمبستگی مکانی و برآورد خوش‌بینانه ناشی از تقسیم تصادفی ساده، از روش نمونه‌برداری طبقه‌بندی‌شده بر مبنای خوشه‌بندی فضایی[4] استفاده شد. بدین منظور، مختصات جغرافیایی نقاط ابتدا با الگوریتم K-Means به ۱۰ منطقه تفکیک شده و سپس از هر منطقه جغرافیایی، ۷۰ درصد داده‌ها (مجموعاً ۳۹۶ نقطه وقوع و ۳۹۷ عدم وقوع) برای آموزش و ساخت مدل‌ها و ۳۰ درصد باقی‌مانده (۱۷۱ نقطه وقوع و ۱۷۰ عدم وقوع) برای ارزیابی مدل‌ها اختصاص یافت. این رویکرد تضمین کرد که تنوع ژئومورفولوژیکی و لیتولوژیکی کل محور مورد مطالعه در هر دو مجموعه داده حفظ شود.

متغیرهای مستقل در این پژوهش، پارامترها و عوامل مؤثر بر وقوع ریزش هستند که بر اساس مرور مطالعات پیشین و ویژگی‌های ژئومورفولوژیکی منطقه مورد مطالعه انتخاب شدند. این متغیرها از نقشه‌های پایه و تصاویر ماهواره‌ای استخراج و به صورت لایه‌های رستری با اندازه سلول ۳۰ متر تهیه شدند. عوامل ژئومورفولوژیک شامل لایه‌های ارتفاع

آماده انتشار (بدون شماره مجلد)

1 - Logistic Regression
2 - Artificial Neural Network
3 - Random Forest
4 - Spatially Stratified Sampling

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

از سطح دریا، درصد شیب، جهت دامنه و شکل و انحنای دامنه از مدل رقومی ارتفاع (DEM) با تفکیک مکانی اولیه ۳۰ متر سنجنده ASTER (نسخه ۳) موجود در پایگاه اطلاعاتی USGS استخراج شدند. این لایه‌ها در سیستم مختصات WGS 1984 UTM Zone 39N تنظیم گردیدند و برای تطابق شبکه‌بندی پیکسل‌ها با سایر لایه‌ها از روش بازنمونه‌برداری درون‌یابی دوخطی[1] استفاده شد. از عوامل زمین‌شناسی، لایه‌های جنس سنگ و خطوط گسل‌های منطقه از نقشه زمین‌شناسی با مقیاس ۱:۱۰۰۰۰۰ قائم‌شهر که توسط سازمان زمین‌شناسی و اکتشافات معدنی ایران در سال ۱۳۸۳ منتشر شده، تهیه و رقومی گردید. سپس لایه فاصله از گسل با استفاده از ابزار تحلیلگر بافر در محیط نرم‌افزار ArcGIS 10.3 تولید شد. همچنین شبکه‌های جاده و رودخانه منطقه بر اساس نقشه‌های توپوگرافی ۱:۲۵۰۰۰ سازمان نقشه برداری کشور مربوط به سال ۱۳۸۰ تهیه شدند و برای ساخت لایه‌های رستریِ فاصله از آن‌ها، تابع فاصله اقلیدسی به کار گرفته شد. نقشه اولیه کاربری اراضی منطقه مورد مطالعه نیز از داده‌های مکانی اداره کل منابع طبیعی استان مازندران استخراج گردید. این نقشه‌های پایه مبتنی بر پردازش و طبقه‌بندی تصاویر ماهواره‌ای Landsat 8 (سنجنده OLI) با تفکیک مکانی ۳۰ متر تهیه شده‌اند. به منظور تدقیق کلاس‌های کاربری، به‌روزرسانی مرز عوارض از پردازش بصری و تفسیر تصاویر ماهواره‌ای با تفکیک مکانی بالا در محیط Google Earth Pro (شامل تصاویر ارائه‌شده توسط Maxar Technologies و Airbus مربوط به بازه زمانی ۲۰۱۹ تا ۲۰۲۱) استفاده شد. در نهایت، با انجام بررسی‌های میدانی در سطح منطقه، اصلاحات لازم بر روی لایه نهایی اعمال گردید. لازم به ذکر است که متغیرهای مستقل این پژوهش بر اساس ویژگی ذاتی آنها برای ورود به مدل‌ها آماده‌سازی شدند. متغیرهای پیوسته شامل ارتفاع، شیب، انحنای دامنه، فاصله از گسل، فاصله از جاده و فاصله از رودخانه به‌صورت مقادیر پیوسته و بدون طبقه‌بندی مجدد وارد مدل‌ها شدند. متغیرهای طبقه‌ای شامل جهت دامنه، جنس سنگ و کاربری اراضی نیز به‌عنوان متغیرهای اسمی در نظر گرفته شدند و به‌منظور جلوگیری از ایجاد رابطه ترتیبی مصنوعی میان طبقات، با استفاده از روش One-Hot Encoding به متغیرهای دودویی تبدیل شدند. بدین ترتیب، هر سه مدل رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی بر اساس مجموعه یکسانی از متغیرهای پیوسته و متغیرهای طبقه‌ای کدگذاری‌شده اجرا شدند. اما ابتدا به منظور بررسی تأثیر هر طبقه از عوامل به صورت مستقل از روش نسبت فراوانی استفاده شده است. وزن نسبی طبقات عوامل با استفاده از رابطه (۱) محاسبه گردید.

$$FR = A / B \qquad (۱)$$

که در آن A درصد پیکسل‌های ریزش در یک طبقه، B درصد پیکسل‌های فاقد ریزش در همان طبقه و FR نسبت فراوانی است. چنانچه مقادیر نسبت فراوانی برای هر طبقه بیشتر از یک باشد، حساسیت آن طبقه به وقوع ریزش بالا می‌باشد (سپه‌وند و بیرانوند، ۱۴۰۳).

از آنجا که وجود همبستگی بالا میان پارامترها و متغیرهای مستقل می‌تواند بر دقت و اعتبار نتایج مدل‌ها تأثیر منفی بگذارد؛ پیش از اجرای مدل‌ها، تحلیل هم‌خطی چندگانه انجام شد. برای این منظور از شاخص‌های آماری ضریب تحمل[2] (TOL) و عامل تورم واریانس[3] (VIF) استفاده شد. چنانچه مقدار ضریب تحمل کوچک‌تر از ۰/۱ و مقدار

آماده انتشار (بدون شماره مجله)

1 - Bilinear
2 - Tolerance
3 – Variance Inflation Factor

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

عامل تورم واریانس (VIF) بزرگ‌تر از ۵ باشد، نشان‌دهنده وجود هم‌خطی شدید بین متغیرهاست (چن[1] و همکاران، ۲۰۱۷).

در مرحله مدل‌سازی و پهنه بندی حساسیت ریزش سنگ ابتدا از مدل آماری رگرسیون لجستیک استفاده شده است. مدل رگرسیون لجستیک یک روش آماری چندمتغیره است که برای توصیف رابطه بین یک متغیر وابسته دودویی (وقوع یا عدم وقوع ریزش سنگ) و چندین متغیر مستقل (X1, X2, ..., Xn) به کار می‌رود (ژانگ و همکاران، ۲۰۲۳). در این مدل، رابطه بین متغیرها به صورت یک منحنی S شکل یا لجستیک بیان می‌شود و برآوردها را در بازه صفر تا یک (احتمال وقوع) قرار می‌دهد؛ به طوری که مقادیر نزدیک به یک نشان‌دهنده احتمال وقوع بیشتر و مقادیر نزدیک به صفر نشان‌دهنده احتمال کمتر وقوع ریزش هستند. احتمال وقوع ریزش سنگ در این مدل بر اساس روابط (۲) و (۳) تعریف می‌شود (اصغری سراسکانرود و همکاران، ۱۳۹۹).

$$p = \frac{1}{1+e^{-z}} \quad (۲)$$

که در آن p احتمال وقوع ریزش سنگ، e عدد نپر با مقدار تقریبی ۲/۷۱۸ و z متغیر خطی مدل که از رابطه ۳ به دست می‌آید:

$$z = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \quad (۳)$$

در این رابطه، b_0 عرض از مبدأ مدل، $x_1 \dots x_n$ متغیرهای مستقل و $b_1 \dots b_n$ ضرایب متناظر با متغیرها هستند. لازم به ذکر است که برای اطمینان از صحت مقایسه نتایج، مدل رگرسیون لجستیک و سایر الگوریتم‌های یادگیری ماشین با استفاده از کتابخانه Scikit-learn در محیط پایتون و با مجموعه داده‌های آموزشی یکسان اجرا شدند. در پیکربندی نهایی مدل رگرسیون لجستیک، حداکثر تعداد تکرارها روی ۱۰۰۰ دوره[2] تنظیم شد و برای جلوگیری از هرگونه سوگیری، وزن کلاس‌ها به‌صورت متوازن[3] اعمال گردید.همان‌طورکه گفته شد، به‌منظور مقایسه عملکرد مدل‌ها و دستیابی به مدلی با دقت بالاتر، از الگوریتم‌های یادگیری ماشین استفاده گردید. شبکه عصبی مصنوعی یکی از پرکاربردترین الگوریتم‌های یادگیری نظارت‌شده است که توانایی بالایی در استخراج روابط پیچیده و غیرخطی میان متغیرهای مستقل و وقوع پدیده‌های ژئومورفولوژیک دارد (یوسف[4] و پورقاسمی، ۲۰۲۱). این الگوریتم که بر پایه ساختار بیولوژیک مغز انسان طراحی شده، بدون نیاز به پیش‌فرض‌های سخت‌گیرانه درباره توزیع داده‌ها، قادر است الگوهای پنهان در مجموعه‌داده‌های چندبعدی را شناسایی کند (مختاری و همکاران، ۲۰۱۹). در این پژوهش، از شبکه عصبی مصنوعی از نوع پرسپترون چندلایه[5] (MLP) استفاده شده است. پیش از ورود داده‌ها به شبکه، مقیاس‌بندی متغیرهای مستقل با استفاده از روش استانداردسازی[6] در محیط پایتون انجام گرفت. این مدل اطلاعات را از لایه های ورودی دریافت کرده و با وزن‌های مختلف به هم ارتباط می‌دهد. عملیات‌های پیچیده‌تر توسط لایه میانی انجام می‌شود و نتیجه این پردازش به صورت لایه خروجی در دسترس قرار می‌گیرد (مائو[7] و همکاران، ۲۰۲۳). این مدل نیز با استفاده از کتابخانه Scikit-learn در محیط برنامه‌نویسی پایتون اجرا شد. ساختار شبکه شامل یک

آماده انتشار (بدون شماره مجله)

1 - Chen
2 - max_iter=1000
3 - class_weight='balanced'
4 - Youssef
5 - Multi-Layer Perceptron
6 - StandardScaler
7 - Mao

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

لایه ورودی، سه لایه پنهان و یک لایه خروجی بوده و تعداد نورون‌های لایه‌های پنهان به‌ترتیب ۱۰۰، ۵۰ و ۲۵ در نظر گرفته شد. این ساختار به‌صورت آزمایشی انتخاب شد تا ضمن افزایش توانایی شبکه در یادگیری الگوهای غیرخطی، از پیچیدگی بیش از حد مدل نیز جلوگیری شود. در تمامی لایه‌های پنهان از تابع فعال‌سازی ReLU[1] استفاده شد. این تابع به دلیل سرعت یادگیری بالا و عملکرد مناسب در شبکه‌های چندلایه، یکی از رایج‌ترین توابع فعال‌سازی در کاربردهای یادگیری عمیق و شبکه‌های عصبی محسوب می‌شود. تابع فعال‌سازی ReLU موجب یادگیری سریع‌تر و عملکرد بهتر نسبت به توابع سیگموئید و تانژانت هیپربولیک می‌شود (نایر[2] و هینتون[3]، ۲۰۱۰). فرآیند آموزش شبکه با استفاده از بهینه‌ساز Adam انجام گرفت. همچنین برای بهبود روند آموزش، نرخ یادگیری تطبیقی به‌کار گرفته شد تا در صورت کاهش سرعت بهبود مدل، نرخ یادگیری به‌صورت خودکار تعدیل شود. حداکثر تعداد تکرار آموزش شبکه ۲۰۰۰ دوره تعیین شد تا فرصت کافی برای همگرایی مدل فراهم گردد. همچنین، از سازوکار توقف زودهنگام استفاده شد. در این روش، بخشی از داده‌های آموزشی به‌عنوان مجموعه اعتبارسنجی داخلی در نظر گرفته شده تا چنانچه عملکرد شبکه طی چند تکرار متوالی بهبود نیابد، فرآیند آموزش به‌طور خودکار متوقف شود. این راهکار از بیش‌برازش جلوگیری کرده و توان تعمیم‌پذیری مدل را افزایش می‌دهد. در این پژوهش، متغیرها و عوامل موثر به‌عنوان ورودی شبکه و وضعیت وقوع یا عدم وقوع ریزش به‌عنوان خروجی مدل در نظر گرفته شد.

الگوریتم جنگل تصادفی یکی دیگر از روش‌های پرکاربرد یادگیری ماشین است که با ایجاد تعداد زیادی درخت تصمیم و تجمیع نتایج آن‌ها، به پیش‌بینی دقیق دست می‌یابد. این الگوریتم قادر به مدل‌سازی روابط غیرخطی پیچیده بین وقوع ریزش سنگ و عوامل موثر است و علاوه بر پیش‌بینی، امکان اولویت بندی اهمیت متغیرها و عوامل موثر را نیز فراهم می‌سازد (رودریگز[4] و همکاران، ۲۰۱۲). برای اجرای مدل نیز از همان مجموعه داده‌های آموزشی و اعتبارسنجی استفاده شد و به‌منظور جلوگیری از سوگیری مدل، وزن کلاس‌ها به‌صورت متوازن تنظیم گردید. در این مرحله، ابتدا پارامترهای اصلی مدل تعریف شد و سپس عملکرد آن با شاخص امتیاز خارج از کیسه[5] (OOB Score) مورد ارزیابی قرار گرفت. در الگوریتم جنگل تصادفی، هر درخت تصمیم با استفاده از نمونه‌گیری تصادفی با جایگذاری[6] از داده‌های آموزشی ساخته شده و به‌طور میانگین حدود ۳۳ درصد از داده‌ها در فرآیند آموزش هر درخت شرکت ندارند. این داده‌ها که به‌عنوان داده‌های خارج از کیسه شناخته می‌شوند، برای ارزیابی داخلی دقت مدل مورد استفاده قرار می‌گیرند. امتیاز OOB به‌صورت نسبت پیش‌بینی‌های صحیح به کل نمونه‌های OOB تعریف شده و می‌توان آن را به‌صورت رابطه (۴) بیان نمود (بریمن[7]، ۲۰۰۱):

$$OOB\ core = \frac{1}{N}\sum_{i=1}^{N} I\left(y_i, \hat{y}_i^{OOB}\right) \qquad (۴)$$

که در آن، N تعداد کل نمونه‌ها در مجموعه داده آموزشی، y_i مقدار واقعی، $\hat{y}_i^{OOB}$ مقدار پیش‌بینی‌شده با استفاده از درخت‌هایی است که نمونه i در آموزش آن‌ها حضور نداشته و I تابع شاخص می‌باشد. در این پژوهش، با تغییر تعداد درختان[8] و محاسبه مقدار OOB Score برای هر حالت، نمودار تغییرات دقت مدل ترسیم شد تا مقدار بهینه این

آماده انتشار (بدون شماره مجله)

1 - Rectified Linear Unit
2 - Nair
3 - Hinton
4 - Rodriguez
5 - Out of Bag Score
6 - Bootstrap
7 - Breiman
8 - n_estimators

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

پارامتر تعیین گردد. بر این اساس و جهت دستیابی به حداکثر پایداری، تنظیمات نهایی مدل شامل ۷۰ درخت، حداکثر عمق نامحدود[1] و تعداد ویژگی‌های بررسی‌شده معادل جذر کل متغیرها[2] تعیین گردید. این فرآیند ارزیابی داخلی، به مدل اجازه می‌دهد تا بدون نیاز به یک مجموعه داده اعتبارسنجی جداگانه، دقت خود را تخمین بزند. برای رسیدن به پیش‌بینی نهایی، فرآیند ساخت و ارزیابی درختان چندین بار تکرار شده تا در نهایت، پیش‌بینی نهایی مدل از طریق میانگین خروجی تمامی درخت‌های تصمیم به‌دست آمد. لازم به ذکر است که کلیه فرآیندهای تنظیم هایپرپارامترها در تمامی مدل‌های یادشده (رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی) صرفاً بر مبنای داده‌های آموزش و اعتبارسنجی داخلی صورت گرفته و به هیچ عنوان از مجموعه داده‌های آزمون نهایی برای بهینه‌سازی پارامترها استفاده نشده است.یکی از قابلیت‌های مهم الگوریتم جنگل تصادفی، امکان تعیین رتبه و اولویت‌بندی اهمیت عوامل مؤثر بر وقوع ریزش است. برای این منظور از شاخص اهمیت پارامتر[3] استفاده شد که در این پژوهش بر اساس میانگین کاهش ناخالصی[4] (MDI) محاسبه گردید. در این روش، پس از آموزش نهایی مدل با پارامترهای بهینه، میزان مشارکت هر یک از پارامترهای ورودی در کاهش ناخالصی گره‌های درختان تصمیم محاسبه می‌گردد. هر بار که یک پارامتر مثل شیب برای تقسیم یک گره استفاده می‌شود، مقدار ناخالصی گره کاهش پیدا می‌کند. الگوریتم، مقدار کاهش ناخالصی ناشی از هر پارامتر را ثبت کرده و این کاهش‌ها را در تمام گره‌ها و تمام درخت‌ها جمع می‌زند. در نهایت، مقادیر به دست آمده نرمال‌سازی شده تا مجموع اهمیت پارامترها برابر با یک شود. اهمیت پارامتر X_j به‌صورت رابطه (۵) تعریف می‌شود (نیکودموس[5]، ۲۰۱۱).

$$FI(X_j) = \sum_{t=1}^{T} \sum_{n \in N_j^t} \frac{N_n}{N} [I(n) - I(n_L) - I(n_R)] \quad (۵)$$

که در آن T تعداد درخت‌ها، N_j^t مجموعه گره هایی در درخت t که با متغیر X_j تقسیم شده‌اند؛ N_n تعداد نمونه ها در گره n، $I(n)$ مقدار ناخالصی گره و n_L, n_R گره های منشعب‌شده چپ و راست می‌باشد. نتایج اهمیت پارامترها به صورت نمودار میله‌ای ارائه گردید تا وزن نسبی هر عامل در وقوع ریزش مشخص شود. علاوه بر شاخصMDI ، به منظور رفع سوگیری‌های احتمالی این شاخص نسبت به متغیرهای پیوسته و همچنین ارائه تفسیر دقیق‌تر از نحوه اثرگذاری متغیرها، از روش پیشرفته مقادیر شاپ[6] (SHAP) نیز در الگوریتم جنگل تصادفی استفاده شد. این رویکرد که مبتنی بر نظریه بازی‌هاست، سهم حاشیه‌ای هر متغیر را در خروجی نهایی مدل محاسبه می‌کند. در نهایت، نتایج اهمیت پارامترها بر اساس دو رویکرد شاخص جینی و مقادیر SHAP به صورت نمودار میله‌ای مقایسه‌ای ارائه گردید تا وزن نسبی و قطعی هر عامل در وقوع ریزش مشخص شود.پس از آموزش و اجرای مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی در محیط پایتون، احتمال وقوع ریزش سنگ برای کل منطقه مورد مطالعه محاسبه شد. خروجی مدل‌ها به محیط GIS منتقل و به‌عنوان لایه‌های رستری نمایش داده شد. به‌منظور امکان مقایسه دقیق میان مدل‌ها، نقشه نهایی احتمال وقوع ریزش سنگ برای هر سه مدل با استفاده از روش فواصل

آماده انتشار (بدون شماره مجله)

1 - max_depth=None
2 - max_features='sqrt'
3 – Feature Importance
4 – Mean Decrease in Impurity
5 - Nicodemus
6 - Shapley Additive exPlanations

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

مساوی[1] و اعمال آستانه‌های مشترک به پنج کلاس حساسیت خیلی کم (۰ تا ۰/۲)، کم (۰/۲ تا ۰/۴)، متوسط (۰/۴ تا ۰/۶)، زیاد (۰/۶ تا ۰/۸) و بسیار زیاد (۰/۸ تا ۱) طبقه‌بندی گردید.

به‌منظور ارزیابی دقت و کارایی مدل‌ها در این پژوهش از شاخص‌های آماری مبتنی بر ماتریس خطا[2] و منحنی ROC استفاده شده است. این ماتریس شامل چهار مؤلفه اصلی مثبت واقعی(TP)، منفی واقعی(TN)، مثبت کاذب (FP) و منفی کاذب (FN) است که مبنای محاسبه شاخص‌های ارزیابی دقت قرار گرفتند. این شاخص ها بر اساس روابط زیر محاسبه شدند (چانگ[3] و همکاران، ۲۰۲۳):

شاخص دقت کلی (Accuracy) که به صورت رابطه (۶) بوده و بیانگر نسبت پیش‌بینی‌های صحیح به کل نمونه‌هاست:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (۶)$$

شاخص دقت پیش‌بینی مثبت (Precision) به صورت رابطه (۷) بوده که نسبت پیش‌بینی‌های صحیح مثبت به کل پیش‌بینی‌های مثبت را نشان می‌دهد و بیانگر میزان خطای پیش‌بینی مثبت کاذب می‌باشد:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (۷)$$

شاخص حساسیت یا بازیابی (Recall / Sensitivity) که به صورت رابطه (۸) بوده و توانایی مدل در شناسایی صحیح رخدادهای واقعی ریزش سنگ را نشان می‌دهد:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (۸)$$

شاخص F1-score که میانگین هارمونیک Precision و Recall و به‌صورت رابطه (۹) بوده و تعادلی بین این دو معیار ایجاد می‌کند:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (۹)$$

منحنی ROC و سطح زیر منحنی(AUC) که رابطه بین نرخ مثبت واقعی و نرخ مثبت کاذب را در آستانه‌های تصمیم مختلف نمایش می‌دهد. سطح زیر این منحنی که با شاخص AUC بیان می‌شود، معیار کلی قدرت تفکیک مدل ها بین کلاس‌های دارای ریزش و بدون ریزش است. هرچه مقدار AUC به عدد یک نزدیک‌تر باشد، نشان‌دهنده عملکرد بهتر مدل در تفکیک کلاس ها و پیش‌بینی صحیح مناطق با حساسیت بالا به ریزش است. نرخ مثبت کاذب به‌صورت رابطه (۱۰) محاسبه می‌شود:

$$\text{FPR} = \frac{FP}{FP+TN} \quad (۱۰)$$

در نهایت، به منظور مقایسه آماری و ارزیابی معناداری اختلاف بین سطح زیر منحنی (AUC) در مدل‌های مختلف از فواصل اطمینان ۹۵ درصد (مبتنی بر روش نمونه‌برداری مجدد بوت‌استرپ با ۵۰۰۰ تکرار) و آزمون آماری ناپارامتریک دلانگ[4] استفاده شد.

آماده انتشار (بدون شماره مجله)

[1] - Equal Interval
[2] - Confusion Matrix
[3] - Chang
[4] - DeLong

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

یافته‌های تحقیق

در این پژوهش، ابتدا به‌منظور بررسی اثر مستقل هر یک از متغیرها و عوامل بر وقوع ریزش سنگ، نسبت فراوانی ریزش در طبقات مختلف هر عامل محاسبه شد. در شکل (۲)، طبقات عوامل موثر بر وقوع ریزش سنگ در منطقه و در جدول (۱) ضرایب نسبت فراوانی در هر یک از طبقات آنها ارائه شده است. هر جا مقدار نسبت فراوانی بیشتر از یک باشد، نشان دهنده نقش تقویت کننده آن طبقه در وقوع ریزش بوده و در نتیجه، حساسیت آن طبقه به ریزش سنگ بالا می باشد. نتایج بررسی واحدهای زمین‌شناسی نشان داد که منطقه مورد مطالعه عمدتاً از سه سازند لار، لالون و شمشک تشکیل شده است. بیشتر سطح منطقه به سازند شمشک تعلق دارد؛ اما بالاترین نسبت فراوانی با مقدار ۱/۹۵ مربوط به سازند لار است که نشان‌دهنده حساسیت بیشتر این سازند به وقوع ریزش سنگ است. مشاهده نتایج سایر متغیرها در جدول (۱) نشان می‌دهد که فاصله ۲۰۰ تا ۳۰۰ متری از گسل، طبقات ارتفاعی ۱۲۰۰ تا ۱۸۰۰ متر از سطح دریا، شیب های بالاتر از ۷۰ درصد، دامنه‌های با جهت جنوب شرقی، شکل دامنه محدب، فاصله ۱۰۰ تا ۱۵۰ متری از رودخانه، فاصله ۵۰ تا ۱۰۰ متری از جاده و کاربری اراضی مرتع فقیر بیشترین مقدار نسبت فراوانی ریزش را به خود اختصاص داده‌اند و از حساسیت بالاتری برخوردار هستند.

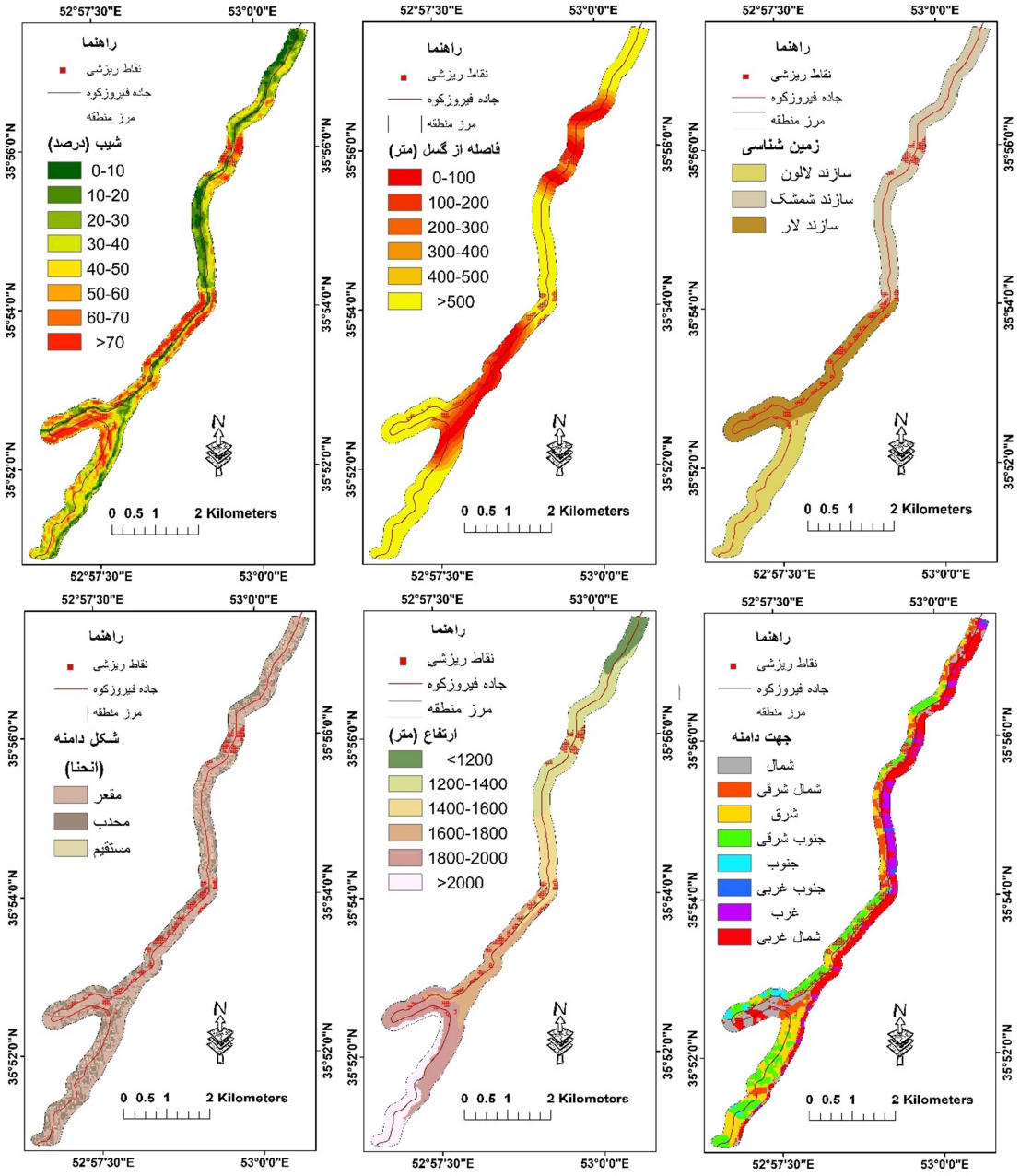
جدول ۱: نسبت فراوانی ریزش در طبقات مختلف عوامل موثر در جاده فیروزکوه (پل سفید- گدوک). *منبع: نگارندگان*

عوامل موثر	طبقات عوامل	درصد پیکسل‌های ریزش	درصد یکسل‌های فاقد ریزش	نسبت فراوانی ریزش
جنس سنگ	سازند لالون	۲/۴۶	۲۴/۸۶	۰/۱۰
	سازند شمشک	۴۳/۳۳	۴۷/۳۴	۰/۹۲
	سازند لار	۵۴/۲۱	۲۷/۸۰	۱/۹۵
فاصله از گسل (متر)	۰-۱۰۰	۲۳/۴۰	۱۳/۱۸	۱/۷۸
	۱۰۰-۲۰۰	۴/۸۹	۱۱/۴۲	۰/۴۳
	۲۰۰-۳۰۰	۱۹/۵۷	۸/۶۶	۲/۲۶
	۳۰۰-۴۰۰	۱۲/۹۸	۶/۵۸	۱/۹۷
	۴۰۰-۵۰۰	۷/۰۲	۵/۲۰	۱/۳۵
	> ۵۰۰	۳۲/۱۳	۵۴/۹۷	۰/۵۸
شیب (درصد)	۰-۱۰	۲/۶۴	۱۰/۷۹	۰/۲۴
	۱۰-۲۰	۵/۹۹	۱۷/۴۹	۰/۳۴
	۲۰-۳۰	۱۱/۰۹	۱۷/۳۸	۰/۶۴
	۳۰-۴۰	۱۲/۶۸	۱۸/۰۰	۰/۷۰
	۴۰-۵۰	۱۱/۶۲	۱۳/۶۶	۰/۸۵
	۵۰-۶۰	۱۴/۴۴	۹/۶۵	۱/۵۰
	۶۰-۷۰	۱۴/۹۶	۵/۸۹	۲/۵۴
	> ۷۰	۲۶/۵۸	۷/۱۵	۳/۷۲
جهت دامنه	شمال	۷/۵۷	۱۱/۱۶	۰/۳۲
	شمال شرق	۵/۸۱	۸/۹۸	۰/۳۰
	شرق	۱۸/۸۴	۲۰/۹۶	۰/۴۲
	جنوب شرق	۴۰/۴۹	۱۸/۹۸	۱/۰۰
	جنوب	۱/۲۳	۲/۳۲	۰/۲۵
	جنوب غرب	۰/۰۰	۰/۷۹	۰/۰۰
	غرب	۱۷/۴۳	۹/۰۷	۰/۹۰

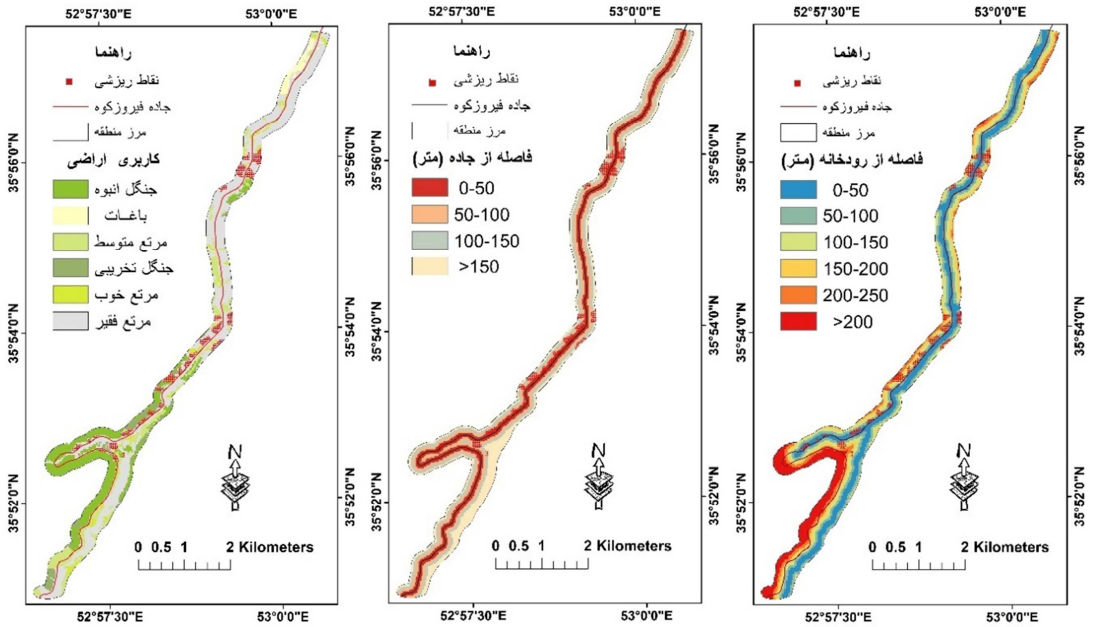
این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

عوامل موثر	طبقات عوامل	درصد پیکسل‌های ریزش	درصد یکسل‌های فاقد ریزش	نسبت فراوانی ریزش
	شمال غرب	۸/۶۳	۲۷/۷۲	۰/۱۵
ارتفاع (متر)	۱۰۰۰-۱۲۰۰	۰/۰۰	۱۸/۵۴	۰/۰۰
	۱۲۰۰-۱۴۰۰	۲۴/۶۵	۱۸/۳۳	۱/۳۴
	۱۴۰۰-۱۶۰۰	۲۳/۹۴	۱۲/۲۴	۱/۹۶
	۱۶۰۰-۱۸۰۰	۳۹/۲۶	۱۶/۲۰	۲/۴۲
	۱۸۰۰-۲۰۰۰	۱۲/۱۵	۲۱/۰۶	۰/۵۸
	> ۲۰۰۰	۰/۰۰	۱۳/۶۲	۰/۰۰
انحنای دامنه	مقعر	۴۹/۳۰	۵۹/۵۲	۰/۸۳
	محدب	۴۴/۳۹	۳۱/۹۵	۱/۳۹
	مستقیم	۶/۳۲	۸/۵۳	۰/۷۴
فاصله از رودخانه (متر)	۰-۵۰	۱۳/۵۱	۲۷/۳	۰/۵۰
	۵۰-۱۰۰	۳۰/۸۸	۲۲/۲۳	۱/۳۹
	۱۰۰-۱۵۰	۳۳/۵۱	۱۸/۸۹	۱/۷۷
	۱۵۰-۲۰۰	۱۳/۶۸	۱۱/۹۳	۱/۱۵
	۲۰۰-۲۵۰	۵/۴۴	۸/۸۱	۰/۶۲
	> ۲۵۰	۲/۹۸	۱۱/۱	۰/۲۷
فاصله از جاده (متر)	۰-۵۰	۲۸/۶۰	۲۶/۸۶	۱/۰۶
	۵۰-۱۰۰	۲۸/۹۵	۲۴/۰۸	۱/۲۰
	۱۰۰-۱۵۰	۲۷/۸۹	۲۵/۱۹	۱/۱۱
	>۱۵۰	۱۴/۵۶	۲۳/۸۸	۰/۶۱
کاربری اراضی	باغ و کشاورزی	۰/۰۰	۵/۳۵	۰/۰۰
	جنگل انبوه	۰/۰۰	۱۸/۶۹	۰/۰۰
	جنگل تخریبی	۵/۷۹	۴/۸۳	۱/۲۰
	مرتع فقیر	۸۴/۳۹	۵۵/۶۳	۱/۵۲
	مرتع متوسط	۹/۳۰	۱۰/۹۱	۰/۸۵
	مرتع خوب	۰/۵۳	۴/۵۸	۰/۱۱

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۲: نقشه‌های متغیرها و عوامل موثر بر وقوع ریزش در جاده فیروزکوه (گدوک- پل سفید). منبع: نگارندگان

به منظور بررسی وجود هم خطی چندگانه میان عوامل و متغیرهای مستقل، شاخص‌های عامل تورم واریانس (VIF) و ضریب تحمل (TOL) محاسبه شد که نتایج در جدول (۲) ارائه شده است. همان‌طور که در این جدول مشاهده می‌شود، هیچ‌یک از متغیرهای مورد استفاده در این پژوهش با مشکل هم‌خطی مواجه نیستند. مقادیر VIF برای تمامی متغیرها کمتر از آستانه بحرانی ۵ بوده و مقادیر ضریب تحمل نیز در همه موارد بیشتر از ۰/۱ هستند که بیانگر استقلال مطلوب متغیرها برای ورود به مدل‌هاست. بیشترین مقدار VIF مربوط به متغیر ارتفاع است که بسیار پایین‌تر از این آستانه بوده و نشان‌دهنده فقدان همبستگی مشکل‌ساز و شرایط کاملاً مناسب و ایده‌آل داده‌ها برای مدل‌سازی است. در مقابل، متغیر جهت دامنه با VIF برابر با ۱/۰۵۳ و ضریب تحمل ۰/۹۰۲ کمترین میزان هم‌خطی را نشان داده است. بررسی این شاخص‌ها پیش از مدل‌سازی موجب افزایش اعتبار آماری نتایج این پژوهش می‌شود.

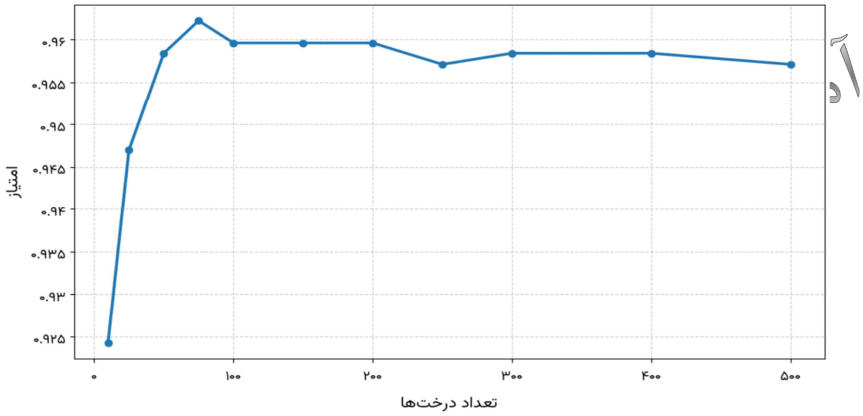
آماده انتشار (بدون شماره مجله)

جدول ۲: نتایج تحلیل هم‌خطی چندگانه میان متغیرهای مستقل با شاخص‌های عامل تورم واریانس و ضریب تحمل. منبع: نگارندگان

ردیف	متغیرهای مستقل	عامل تورم واریانس	ضریب تحمل
۱	ارتفاع	۲/۵۳۷	۰/۱۵۵
۲	فاصله از رودخانه	۱/۵۸۴	۰/۳۹۹
۳	شیب	۱/۴۳۹	۰/۴۸۳
۴	جنس سنگ	۱/۳۷۰	۰/۵۳۳
۵	فاصله از گسل	۱/۳۵۷	۰/۵۴۳
۶	فاصله از جاده	۱/۲۲۳	۰/۶۶۹
۷	کاربری اراضی	۱/۰۷۷	۰/۸۶۳
۸	انحنای دامنه	۱/۰۶۵	۰/۸۸۲
۹	جهت دامنه	۱/۰۵۳	۰/۹۰۲

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

با توجه به تفاوت در ساختار مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی، فرآیند تنظیم پارامترها متناسب با ویژگی‌های هر مدل انجام شد. از آنجا که مدل جنگل تصادفی امکان ارزیابی داخلی پارامترها را از طریق امتیاز خارج از کیسه (OOB Score) فراهم می‌کند، تغییرات این شاخص برای انتخاب مقادیر مناسب پارامترها مورد بررسی قرار گرفت. براین‌اساس، جهت دستیابی به بالاترین کارایی مدل جنگل تصادفی، ترکیبات مختلفی از تعداد درختان و حداکثر عمق درخت مورد ارزیابی قرار گرفت. نتایج تغییرات امتیاز خارج از کیسه (شکل ۳) نشان می‌دهد که با افزایش تعداد درختان تصمیم، دقت پیش‌بینی مدل در ابتدا به سرعت افزایش یافته و سپس به یک وضعیت پایدار می‌رسد. مقدار OOB Score در تعداد کم درختان پایین‌تر بوده و پس از رسیدن تعداد آنها به حدود ۷۰ تا ۱۰۰، امتیاز OOB Score به حداکثر مقدار خود رسیده و در حوالی ۰/۹۶ تغییرات آن ناچیز می‌شود. در این بازه، مدل از پایداری مناسبی برخوردار شده و افزودن درختان بیشتر تأثیر معناداری بر بهبود دقت ندارد. بر همین اساس و جهت اطمینان از تعمیم‌پذیری و پایداری کامل مدل، تعداد نهایی درختان بر روی ۷۰ تنظیم گردید. این تنظیمات که بر اساس حداکثر OOB Score انتخاب شد؛ از بروز بیش‌برازشی و افزایش غیرضروری هزینه محاسباتی جلوگیری می‌کند.



شکل ۳: نمودار تغییرات امتیاز خارج از کیسه در برابر تعداد درخت‌ها. *منبع: نگارندگان*

پس از تنظیم دقیق پارامترها و اجرای مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی، خروجی هر سه مدل به‌صورت نقشه احتمال وقوع ریزش سنگ برای منطقه مورد مطالعه تهیه شده است. همچنین عملکرد مدل‌ها با استفاده از شاخص‌های آماری و منحنی تشخیص عملکرد نسبی (ROC) مورد ارزیابی قرار گرفت و اهمیت نسبی عوامل مؤثر بر وقوع ریزش سنگ نیز با استفاده از شاخص میانگین کاهش ناخالصی و مقادیرSHAP مدل جنگل تصادفی تعیین شد.

نتایج و بحث

نتایج ارزیابی عملکرد مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی نشان می‌دهد که هر سه مدل استفاده شده در این پژوهش از توانایی بسیار بالایی در پهنه‌بندی حساسیت وقوع ریزش سنگ برخوردارند (جدول ۳). مدل جنگل تصادفی با مقدار Accuracy برابر با ۰/۹۴۷ و مقدار AUC برابر با ۰/۹۹۳عملکرد بسیار عالی ارائه داده است. مقدار AUC بالاتر از ۰/۹ نشان دهنده قدرت تفکیک و توان تمایز بسیار خوب بین مناطق دارای

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

ریزش و بدون ریزش است. همچنین مقدار Recall برابر با ۰/۹۵۹ نشان می دهد که این مدل توانست بخش عمده‌ای از رخدادهای واقعی ریزش در منطقه را به‌درستی شناسایی کند و میزان خطای نوع دوم (FN=7) را به حداقل برساند. مدل شبکه عصبی مصنوعی نیز عملکرد بسیار مطلوبی از خود نشان داد؛ به‌طوری‌که با مقدار Accuracy برابر با ۰/۹۴۱ و AUC برابر با ۰/۹۷۶ با رقابتی بسیار نزدیک پس از جنگل تصادفی در رتبه دوم قرار گرفت. مقدار Recallاین مدل برابر با ۰/۹۴۷ بود که بیانگر توانایی بسیار بالای آن در شناسایی رخدادهای واقعی ریزش در منطقه مورد مطالعه است؛ به‌طوری‌که تنها ۹ مورد از نقاط وقوع ریزش (FN=9) توسط مدل شناسایی نشد. همچنین، مقدار Precision این مدل با رقم ۰/۹۳۶ نشان می‌دهد که در کنار حساسیت بالا، تعداد مثبت‌های کاذب آن (FP=11) نیز در سطح حداقل و کاملاً مشابه با الگوریتم جنگل تصادفی بوده است. در مقابل، مدل رگرسیون لجستیک با مقدار Accuracy برابر با ۰/۹۰۹ و مقدار AUC برابر با ۰/۹۵۷ اگرچه کارایی بسیار خوبی داشت، اما عملکرد پایین‌تری نسبت به دو مدل یادگیری ماشین نشان داد. مقدار Recall این مدل با رقم ۰/۹۱۸ نشان می‌دهد که ۱۴ مورد از رخدادهای واقعی ریزش (FN=14) توسط مدل شناسایی نشده‌اند. این تفاوت عملکرد را می‌توان به ماهیت خطی مدل رگرسیون لجستیک نسبت داد؛ به‌گونه‌ای که این روش پارامتریک قادر به مدل‌سازی کامل روابط غیرخطی و پیچیده میان متغیرها و عوامل موثر در منطقه مورد مطالعه نیست.

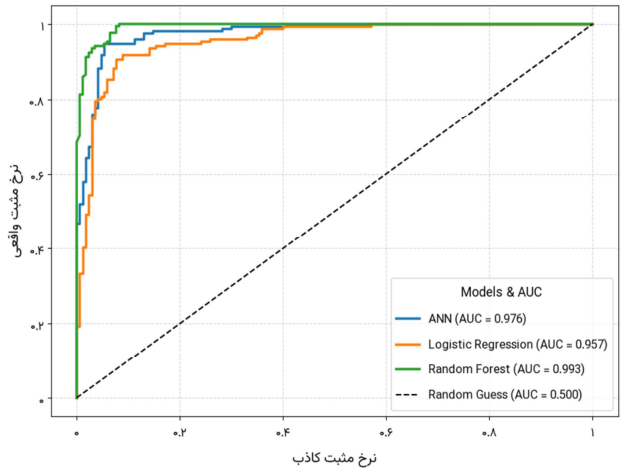
جدول ۳: ارزیابی عملکرد مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی در پهنه‌بندی حساسیت وقوع ریزش سنگ با استفاده از شاخص‌های آماری مختلف. *منبع: نگارندگان*

آماده انتشار (بدون شماره مجله)

مدل	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-score	AUC
رگرسیون لجستیک(LR)	۱۵۷	۱۵۳	۱۷	۱۴	۰/۹۰۹	۰/۹۰۲	۰/۹۱۸	۰/۹۱۰	۰/۹۵۷
شبکه عصبی مصنوعی(ANN)	۱۶۲	۱۵۹	۱۱	۹	۰/۹۴۱	۰/۹۳۶	۰/۹۴۷	۰/۹۴۲	۰/۹۷۶
جنگل تصادفی(RF)	۱۶۴	۱۵۹	۱۱	۷	۰/۹۴۷	۰/۹۳۷	۰/۹۵۹	۰/۹۴۸	۰/۹۹۳

تحلیل منحنی تشخیص عملکرد نسبی (ROC) نشان می‌دهد که مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی از توان تفکیک قابل قبولی برای تشخیص نواحی مستعد وقوع ریزش سنگ برخوردار هستند (شکل ۴). با این حال، منحنی ROC مربوط به مدل جنگل تصادفی در بیشتر مقادیر نرخ مثبت کاذب (FPR) بالاتر از دو مدل دیگر قرار گرفته که بیانگر دستیابی به نرخ مثبت حقیقی (TPR) بیشتر در شرایط خطای مشابه است. همچنین، منحنی مدل شبکه عصبی مصنوعی در بیشتر دامنه‌ها بالاتر از منحنی رگرسیون لجستیک و نزدیک به منحنی جنگل تصادفی قرار دارد که نشان‌دهنده عملکرد مناسب و قدرت تفکیک بالای این مدل می‌باشد. در مقابل، مدل رگرسیون لجستیک پایین‌ترین مقدار AUC را داشته و منحنی ROC آن در بیشتر دامنه‌ها پایین‌تر از دو مدل یادگیری ماشین قرار گرفته است. به این ترتیب، نتایج منحنی ROC عملکرد نسبتا بهتر مدل جنگل تصادفی، رقابت بسیار نزدیک آن با شبکه عصبی مصنوعی و عملکرد ضعیف‌تر رگرسیون لجستیک را در تفکیک بین کلاس‌های وقوع و عدم وقوع ریزش سنگ تأیید می‌کند. این نتایج با یافته‌های فانوس و همکاران (۲۰۱۸)، فانوس و همکاران (۲۰۲۰)، بهاگات و همکاران (۲۰۲۲)، محمودزاده و همکاران (۲۰۲۳) و هی و همکاران (۲۰۲۵) در زمینه پهنه‌بندی حساسیت ریزش سنگ هم‌خوانی دارد؛ زیرا آنها نیز در پژوهش‌های خود دریافتند که مدل جنگل تصادفی در مقایسه با مدل‌های آماری کلاسیک و دیگر مدل‌ها، عملکرد بهتری از خود نشان می‌دهد.

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۴: منحنی تشخیص عملکرد نسبی و سطح زیر منحنی نقشه‌های پهنه‌بندی حساسیت ریزش در منطقه مطالعه. *منبع: نگارندگان*

به منظور اثبات معناداری تفاوت عملکرد مدل‌ها و پرهیز از نتیجه‌گیری صرفاً توصیفی، فواصل اطمینان ۹۵ درصد (مبتنی بر روش بوت‌استرپ با ۵۰۰۰ تکرار) و آزمون معناداری آماری دلانگ برای شاخص AUC محاسبه شد (جداول ۴ و ۵). نتایج نشان داد که مدل جنگل تصادفی با فاصله اطمینان (۱-۰/۹۸۲) پایدارترین عملکرد را داراست. به علاوه، نتایج آزمون دلانگ ثابت کرد که برتری مدل جنگل تصادفی نسبت به شبکه عصبی مصنوعی (Z = -3.306, p = 0.0009) و همچنین نسبت به رگرسیون لجستیک (Z = -4.315, p < 0.001) در سطح خطای کمتر از ۰/۰۰۱ از نظر آماری معنادار است که نشان‌دهنده توانمندی ذاتی این الگوریتم در مدل‌سازی روابط پیچیده و غیرخطی ریزش سنگ در محور مورد مطالعه می‌باشد.

آماده انتشار (بدون شماره مجله)

جدول ۴: مقایسه شاخص AUC مدل‌ها به همراه فواصل اطمینان ۹۵ درصد (روش بوت‌استرپ با ۵۰۰۰ تکرار). *منبع: نگارندگان*

مدل ارزیابی شده	AUC برآورد نقطه‌ای	حد پایین (۲/۵٪)	حد بالا (۹۷/۵٪)	فاصله اطمینان ۹۵٪
رگرسیون لجستیک (LR)	۰/۹۵۷	۰/۹۱۷	۰/۹۹۴	۰/۹۵۷ (۰/۹۱۷-۰/۹۹۴)
شبکه عصبی مصنوعی (ANN)	۰/۹۷۶	۰/۹۴۶	۰/۹۹۹	۰/۹۷۶ (۰/۹۴۶-۰/۹۹۹)
جنگل تصادفی (RF)	۰/۹۹۳	۰/۹۸۲	۱	۰/۹۹۳ (۰/۹۸۲-۱)

جدول ۵: نتایج آزمون آماری دلانگ برای مقایسه معناداری زوجی AUC بین مدل‌ها. *منبع: نگارندگان*

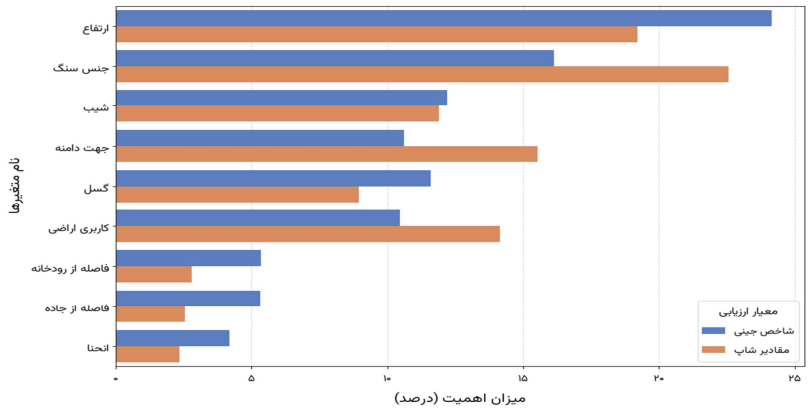
مقایسه زوجی مدل‌ها	AUC مدل اول	AUC مدل دوم	اختلاف کارایی (ΔAUC)	آماره Z آزمون	مقدار P-Value
ANN در برابر LR	۰/۹۷۶۲	۰/۹۵۶۶	+۰/۰۱۹۶	+۴/۰۸۳	۰/۰۰۰ (p < 0.001)
RF در برابر LR	۰/۹۹۳۰	۰/۹۵۶۶	+۰/۰۳۶۴	-۴/۳۱۵	۰/۰۰۰ (p < 0.001)
RF در برابر ANN	۰/۹۹۳۰	۰/۹۷۶۲	+۰/۰۱۶۸	-۳/۳۰۶	۰/۰۰۰۹ (p < 0.001)

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

اولویت‌بندی متغیرها و عوامل مؤثر بر وقوع ریزش سنگ بر اساس دو شاخص میانگین کاهش ناخالصی (MDI) یا جینی و مقادیر پیشرفته SHAP در مدل جنگل تصادفی انجام شد. نتایج این دو معیار ارزیابی به صورت اهمیت یا رتبه هر عامل مشخص شده که در شکل ۵ ارائه شده است. همان‌طور که نتایج مدل نشان می‌دهند، عوامل ارتفاع و زمین‌شناسی بیشترین سهم را در تبیین الگوی مکانی وقوع ریزش در منطقه مورد مطالعه داشته‌اند. بر مبنای شاخص جینی، عامل ارتفاع با بیش از ۲۴ درصد در رتبه اول و جنس‌سنگ با حدود ۱۶ درصد در رتبه دوم قرار گرفت. در مقابل، تحلیل مقادیر SHAP نشان داد که عامل جنس‌سنگ (با بیش از ۲۲ درصد اهمیت) تأثیرگذارترین عامل است و پس از آن ارتفاع (با حدود ۱۹ درصد) قرار دارد. از منظر ژئومورفولوژی می‌توان استنباط کرد که این درجه از اهمیت برای عامل جنس‌سنگ کاملاً توجیه‌پذیر است، زیرا لیتولوژی منطقه نقش بنیادی در تأمین مصالح و مستعد بودن دامنه‌ها برای ریزش دارد. در تفسیر این نتایج آماری می‌توان به یافته‌های تحلیل نسبت فراوانی (جدول ۱) رجوع کرد که در آن سازند لار با بالاترین نسبت فراوانی (۱/۹۵)، بیشترین حساسیت را در میان واحدهای سنگ‌شناسی نشان داد. این امر احتمالاً ناشی از ویژگی‌های ساختاری این سازند نظیر مقاومت کمتر در برابر هوازدگی فیزیکی و وجود درز و شکاف‌های فراوان در ارتفاعات منطقه مورد مطالعه است. پس از این دو عامل کلیدی، متغیرهای جهت دامنه، شیب، گسل و کاربری اراضی در رتبه‌های بعدی اهمیت قرار دارند. اگرچه رتبه شیب نسبت به مدل‌سازی‌های اولیه تعدیل شده است، اما تأثیر آماری نشان‌داده‌شده توسط مدل، با تحلیل نسبت فراوانی نیز همخوانی دارد؛ به‌گونه‌ای که طبقات شیب بالاتر از ۷۰ درصد بیشترین مقدار نسبت فراوانی را نشان دادند. در تفسیر این ارتباط می‌توان افزود که با افزایش شیب، نیروهای رانشی بیشتر شده و پتانسیل ناپایداری دامنه و تنش برشی افزایش می‌یابد. همچنین اهمیت بالای جهت دامنه و کاربری اراضی در این نمودار نشان‌دهنده اهمیت هم‌زمان شرایط پوشش سطحی و تفاوت دریافت انرژی تابشی خورشید در کنترل ناپایداری دامنه‌هاست. در طبقات مرتع فقیر که بالاترین نسبت فراوانی ریزش را در میان انواع کاربری اراضی در منطقه نشان دادند، ضعف پوشش گیاهی و کاهش نقش تقویت‌کننده ریشه‌ها می‌تواند حساسیت دامنه را افزایش دهد. به طور مشابه در دامنه‌های جنوب‌شرقی نیز بیشترین نسبت فراوانی ریزش مشاهده شد که احتمالا ناشی از اثر تابش مستقیم خورشید، تغییرات درجه حرارت روزانه و فرایندهای فیزیکی تخریب سنگ است. در مقابل، نتایج مستقیم هر دو شاخص ارزیابی نشان می‌دهند که متغیرهای فاصله از رودخانه با حدود ۵ درصد و فاصله از جاده با ارقامی مشابه نقش و اهمیت کمتری را نشان داده‌اند. این موضوع نشان می‌دهد که اگرچه فرایندهای فرسایش پای دامنه و برش‌های مصنوعی ناشی از احداث جاده در برخی نقاط می‌توانند محرک ناپایداری دامنه باشند؛ اما در مقیاس کل منطقه، عوامل ارتفاع و جنس‌سنگ بیشترین تأثیر را بر احتمال وقوع ریزش دارند. در نهایت، کمترین میزان اهمیت به شکل و انحنای دامنه با مقادیر کمتر از ۴ درصد در هر دو شاخص اختصاص پیدا کرده است که بیانگر نقش ثانویه این عامل در مقایسه با عوامل موثر دیگر در منطقه مورد مطالعه است. همخوانی کلی نتایج این دو معیار ارزیابی در شناسایی مؤثرترین و کم‌اثرترین متغیرها، اعتبار تحلیل‌های این پژوهش را دوچندان می‌سازد.

آماده انتشار (بدون شماره مجله)

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

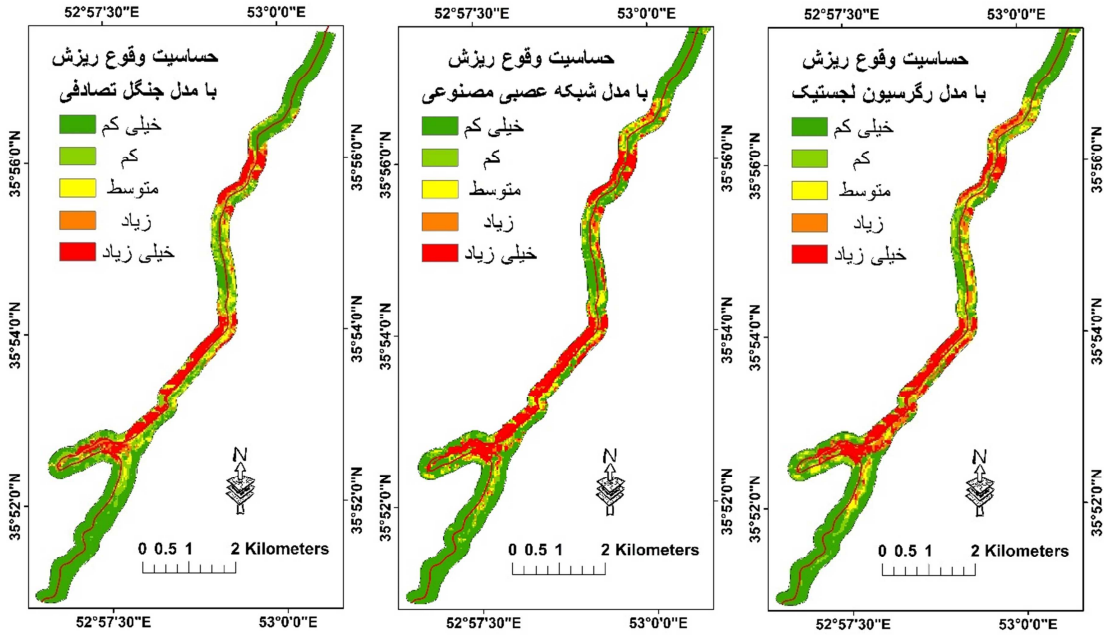


شکل ۵ : میزان اهمیت عوامل موثر در وقوع ریزش بر اساس میانگین کاهش ناخالصی و مقادیر SHAP. *منبع: نگارندگان*

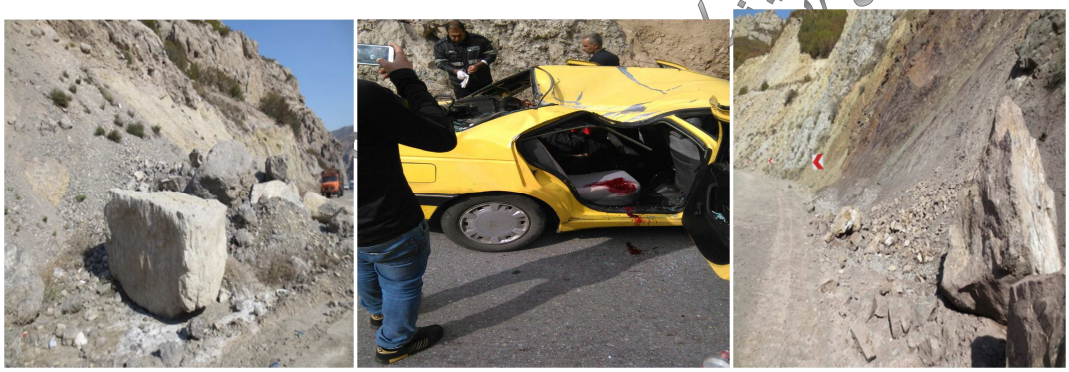
در این پژوهش، مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی با استفاده از ۷۰ درصد داده‌های آموزشی توسعه یافته و خروجی هر سه مدل به‌صورت احتمال وقوع ریزش برای هر پیکسل تولید شد. مقادیر پیش‌بینی‌شده که در بازه صفر تا یک قرار دارند، شدت حساسیت هر پیکسل به وقوع ریزش سنگ در محور فیروزکوه را نشان می‌دهند. به‌منظور تهیه نقشه‌های نهایی پهنه‌بندی حساسیت، مدل‌های آموزش‌دیده برای کل منطقه مورد مطالعه اعمال شده و نتایج با استفاده از روش فواصل مساوی با آستانه های یکسان برای تمامی مدل‌ها به پنج کلاس حساسیت (خیلی کم تا خیلی زیاد) طبقه‌بندی گردید (شکل ۶). بر اساس نتایج کمی حاصل از مدل رگرسیون لجستیک، حدود ۲۶/۶۴ درصد از مساحت منطقه در کلاس‌های حساسیت زیاد و خیلی زیاد قرار دارد. این مقدار در مدل شبکه عصبی مصنوعی به حدود ۲۴/۱۸ درصد کاهش یافت؛ در حالی که مدل جنگل تصادفی تنها حدود ۱۹/۲۴ درصد از منطقه را در این کلاس‌ها قرار داد. این تفاوت بیانگر آن است که مدل‌های رگرسیون لجستیک و شبکه عصبی مصنوعی پهنه‌های حساس را با گستردگی بیشتری برآورد نمود؛ اما جنگل تصادفی با وجود داشتن دقت و قدرت تفکیک بالاتر، نواحی با حساسیت زیاد را به‌صورت محدود و متمرکز شناسایی کرده است. تطابق این پهنه‌های حساس با واقعیت‌های میدانی، ضرورت مدل‌سازی دقیق را بیش از پیش نمایان می‌سازد. همان‌طور که در مشاهدات میدانی از رخدادهای منطقه (شکل ۷) مشخص است، وقوع ریزش در این نواحیِ مستعد صرفاً به جابه‌جایی واریزه‌های کوچک محدود نمی‌شود؛ بلکه شامل سقوط بلوک‌های سنگی با ابعاد و وزن بسیار زیاد است. این پدیده به دلیل انرژی جنبشی بالا در زمان سقوط، نه‌تنها شانه خاکی و زیرساخت جاده را تخریب می‌کند، بلکه عامل اصلی سوانح جاده‌ای و وارد آمدن خسارات شدید کالبدی و جانی به وسایل نقلیه عبوری در این محور می‌باشد.

آماده انتشار (بدون شماره و محله)

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۶: نقشه‌های پهنه‌بندی حساسیت وقوع ریزش سنگ در منطقه مورد مطالعه بر اساس مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی. *منبع: نگارندگان*



شکل ۷: شواهد میدانی از شدت خسارات در پهنه‌های با حساسیت بالا در محور پل‌سفید-گدوک؛ تصاویر، ابعاد بزرگ بلوک‌های سنگی و وقوع ریزش در حاشیه جاده و تخریب شدید وسایل نقلیه عبوری را نشان می‌دهند. *منبع: نگارندگان*

نتیجه‌گیری

ریزش سنگ یکی از فرآیندهای مهم و مخرب دامنه‌ای در مناطق کوهستانی است که جان انسان‌ها، اموال، زیرساخت‌ها و وسایل حمل‌ونقل را تهدید می‌کند. جاده فیروزکوه که در بستر زمین‌شناسی پیچیده و دامنه‌های شیب‌دار البرز مرکزی قرار گرفته است؛ به‌طور ویژه در محدوده پل سفید تا گدوک در معرض وقوع این نوع از ناپایداری‌ها قرار دارد. بنابراین، یک نقشه دقیق از حساسیت وقوع ریزش سنگ برای مدیران و برنامه‌ریزان استفاده از زمین ضروری است. بدین‌منظور، در این پژوهش به پهنه‌بندی حساسیت وقوع ریزش سنگ در محدوده یاد شده با بهره‌گیری از یک مدل آماری و دو الگوریتم یادگیری ماشین پرداخته شد و کارایی آنها با یکدیگر مقایسه و ارزیابی گردید. نتایج حاصل از تحلیل‌های پیشرفته MDI و SHAP نشان داد که وقوع ریزش در منطقه عمدتاً تحت تأثیر

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

عوامل مورفومتریک، به‌ویژه ارتفاع و جنس زمین قرار دارد. مقایسه کارایی مدل‌های رگرسیون لجستیک، شبکه عصبی مصنوعی و جنگل تصادفی با استفاده از شاخص‌های آماری مختلف و همچنین تحلیل منحنی ROC نشان داد که اگرچه هر سه مدل از قابلیت پیش‌بینی مناسبی برخوردار بودند؛ اما الگوریتم یادگیری ماشین جنگل تصادفی به دلیل توانایی در مدل‌سازی روابط غیرخطی میان متغیرها، عملکرد دقیق‌تر و تفکیک‌پذیری مکانی بهتری ارائه کرده است. علاوه بر این، محاسبه فواصل اطمینان ۹۵ درصد (مبتنی بر روش بوت‌استرپ) و نتایج آزمون معناداری آماری دلانگ ثابت کرد که برتری عملکرد مدل جنگل تصادفی نسبت به شبکه عصبی مصنوعی ($p = 0.0009$) و رگرسیون لجستیک ($p < 0.001$) از نظر آماری معنادار و قطعی است. این یافته اهمیت به‌کارگیری الگوریتم‌های یادگیری ماشین در مطالعات پهنه‌بندی حساسیت ناپایداری‌های دامنه‌ای را در محیط‌های پیچیده ژئومورفولوژیک نشان می‌دهد؛ به طوری که می‌توان به پهنه‌های حساسیت مشخص شده در منطقه اعتماد کرده و نقشه حاصل از آن را ملاک قرار داد. بر اساس نقشه نهایی پهنه بندی تهیه شده، حدود ۱۶۰/۷ هکتار از مساحت منطقه مورد مطالعه در محور پل‌سفید تا گدوک در کلاس‌های حساسیت زیاد و خیلی زیاد وقوع ریزش سنگ قرار دارد. بنابراین، توجه به این پدیده بسیار ضروری است و پیشنهاد می‌شود از نقشه حساسیت تهیه‌شده به‌عنوان ابزار پشتیبان تصمیم‌گیری برای شناسایی نواحی مستعد، اولویت‌بندی اقدامات پایدارسازی و برنامه‌ریزی زیرساختی در منطقه استفاده شود. با این همه، نتایج این پژوهش تحت تأثیر کیفیت و قدرت تفکیک داده‌های ورودی و همچنین شکل و ساختار فضایی منطقه مورد مطالعه بوده است. یکی از محدودیت‌های این پژوهش، عدم دسترسی به یک پایگاه داده کاملاً مستقل جهت اعتبارسنجی خارجی مدل‌ها بود. همچنین، اگرچه در این مطالعه برای کنترل اثرات مجاورت از روش نمونه‌برداری طبقه‌بندی‌شده مکانی استفاده شد، اما به دلیل ساختار خطی جاده و توزیع متمرکز و ناهمگن برخی متغیرهای مهم (نظیر جنس سنگ و زون‌های گسلی)، اجرای اعتبارسنجی متقاطع بلوک مکانی[1] به طور کامل امکان‌پذیر نبود. از این‌رو، برای اثبات قطعی قابلیت تعمیم‌پذیری مدل‌ها پیشنهاد می‌شود در مطالعات آتی (به ویژه در مقیاس حوضه‌ای) ضمن استفاده از داده‌های زمینی کاملاً مستقل از روش ارزیابی یادشده بهره گرفته شود.

آماده انتشار (بدون شماره مجله)

منابع

اسفندیاری درآباد، فریبا، رحیمی، مسعود، وهاب زاده، مهرداد. (۱۴۰۳). پهنه‌بندی مخاطره ریزش سنگ در جاده خلخال به شاهرود با الگوریتم پرسپترون چندلایه . *جغرافیا و مخاطرات محیطی*، ۱۳(۴)، ۲۶-۱. https://doi.org/10.22067/geoeh.2024.87847.1482

اصغری سراسکانرود، صیاد، حسن زاده، رسول، رئوفی، سهیل. (۱۳۹۹). بررسی عوامل موثر بر ریزش سنگی و پهنه‌بندی خطر آن با رگرسیون لجستیک در حوضه آبریز علی‌آباد چای هوراند. *هیدروژئومورفولوژی*، ۷(۲۳)، ۳۸-۲۱. https://doi.org/10.22034/hyd.2020.11004

اصغری سراسکانرود، صیاد، مظفری، حسن. (۱۳۹۸). ارزیابی و مقایسه مدل های ضریب نسبت فراوانی و تحلیل شبکه در پهنه بندی ریزش سنگ. *تحلیل فضایی مخاطرات محیطی*، ۶(۴)، ۱۲۳-۱۴۲. https://doi.org/10.29252/jsaeh.6.4.123

پاک‌نژاد فریبا، احمدآبادی علی، قنواتی عزت اله، زحمتکش حسن. (۱۴۰۲). پهنه‌بندی مناطق مستعد دارای جریان‌های واریزه‌ای با استفاده از الگوریتم‌های یادگیری ماشین (مطالعه موردی: حوضه تنگراه استان گلستان)، *پژوهش های فرسایش محیطی*، ۱۳ (۱)، ۱-۲۸. http://magazine.hormozgan.ac.ir/article-1-702-fa.html

جوکار سرهنگی، عیسی، علیزاده، عبداله. (۱۳۹۳). ارزیابی کارایی مدل های نسبت فراوانی و ارزش اطلاعات در پهنه بندی خطر ریزش سنگ. مطالعه موردی دره *هراز*. *مجله آمایش جغرافیایی فضا*، ۴(۱۱)، ۱۶۳-۱۵۱. https://gps.gu.ac.ir/article_7402.html?lang=en

[1]- Spatial Block Cross-Validation

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

رحیمی، مسعود، اسفندیاری، فریبا، زکی زاده، امید، وهاب زاده، زرگری، مهرداد. (۱۴۰۴). شناسایی پهنه‌های مستعد خطر ریزش‌سنگ در جاده ارتباطی مشگین‌شهر- اهر با استفاده از شبکه عصبی مصنوعی. *مخاطرات محیط طبیعی*، ۱۴(۴۶)، ۲۳-۴۲.
https://doi.org/10.22111/jneh.2025.50028.2072

روستایی، شهرام، مختاری، داود، خدایی، فاطمه. (۱۳۹۴). کاربرد روش رگرسیون لجستیک در پهنه‌بندی خطر وقوع ناپایداری‌های دامنه‌ای در جاده‌های کوهستانی (تنگه دره دیز). *مخاطرات محیط طبیعی*، ۴(۶)، ۸۹-۱۰۳.https://doi.org/10.22111/jneh.2016.2525

سپه وند، علیرضا، بیرانوند، نسرین. (۱۴۰۳). پهنه‌بندی حساسیت وقوع زمین‌لغزش با استفاده از الگوریتم‌های یادگیری ماشین (بخشی از حوزهٔ آبخیز هراز). *مدل سازی و مدیریت آب و خاک*، ۴(۲)، ۲۶۱-۲۷۸. https://doi.org/10.22098/mmws.2023.12678.1263

شیرزادی عطاالله، سلیمانی کریم، حبیب نژاد روشن محمود، کاویان عطاالله، قاسمیان بهاره.(۱۳۹۴). مقایسه روش های رگرسیون لجستیک، نسبت فراوانی و تحلیل سلسله مراتبی در تهیه ی نقشه ی حساسیت ریزش‌سنگ (مطالعه موردی: استان کردستان، گردنه صلوات آباد) *پژوهشنامه مدیریت حوزه آبخیز*، ۶ (۱۲)،۲۰۴-۱۹۳. https://jwmr.sanru.ac.ir/article-1-570-en.html

عابدینی، موسی، مظفری، حسن، فعال نذیری، مهدی. (۱۴۰۱). بررسی و مقایسه کارایی مدل‌های ارزش اطلاعات و ضریب نسبت فراوانی و آنتروپی شانون در پهنه‌بندی خطر ریزش سنگ (مطالعه موردی:جاده زنجان ـ تهم ـ طارم). *فصلنامه مطالعات جغرافیایی مناطق کوهستانی*، ۳(۹)، ۵۷-۷۷. https://doi.org/10.52547/gsma.3.1.55

مومنی، علی اکبر، حیدری، مجتبی. (۱۴۰۰). آنالیز خطر سقوط سنگ در منطقه توریستی گنجنامه، غرب ایران. *مخاطرات محیط طبیعی*، ۱۰(۲۸)، ۱-۱۴. https://doi.org/10.22111/jneh.2021.31304.1553

Abedini, M., Mozafari, H., & Faal naziri, M. (2022). Investigating and comparing the effectiveness of information value models and frequency ratio coefficient and Shannon's entropy in zoning rock fall risk (case study: Zanjan-Teham-Taram Road). *Geographical Studies of Mountainous Regions*, 3(1), 55-75. https://doi.org/10.52547/gsma.3.1.55. [In Persian]

Asghari Saraskanrood, S., Hasan zadeh, R. and Raoofi, S. (2020). Investigation of Factors Influencing rock fall and Its Zoning with Logistic Regression in Ali Bayad Basin of Horand Abstract. *Hydrogeomorphology*, 7(23), 38-21. https://doi.org/10.22034/hyd.2020.11004. [In Persian]

Asghari Saraskanrood, S., & Mozafari, H. (2020). Estimate and Comparison of Frequency Ratio and Network Analysis Models in Rockfall Zoning (Case study Zanjan road to Taham and Tarom). *Journal of Spatial Analysis, Environmental Hazards*, 6(4), 123-142. https://doi.org/10.29252/jsaeh.6.4.123. [In Persian]

Bhagat, N. K., Mishra, A. K., Singh, R. K., Sawmliana, C., & Singh, P. K. (2022). Application of logistic regression, CART, and random forest techniques in prediction of blast-induced slope failure during reconstruction of railway rock-cut slopes. *Engineering Failure Analysis*, 137, 106230. https://doi.org/10.1016/j.engfailanal.2022.106230

Breiman, L., 2001. Random forests. *Machine Learning*, 45(1), 5-32. https://doi.org/10.1023/a:1010933404324

Chang, Z., Huang, F., Huang, J., Jiang, Sh.H., Yuting Liu, Y., Meena, S.R., & Catani, F. (2023). An update of landslide susceptibility prediction from the perspective of space and time. Geoscience Frontiers, 14(5), 101619. https://doi.org/10.1016/j.gsf.2023.101619

Chen W, Xie X, Wang J, Pradhan B, Hong H, Tien Bui D, Duan Z, Ma J. 2017. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena. 151, 147-160. https://doi.org/10.1016/j.catena.2016.11.032

Esfandyari Darabad, F., Rahimi, M., & Vahabzadeh, M. (2024). Rockfall risk zoning on the Khalkhal to Shahroud road using multilayer perceptron algorithm. *Journal of Geography and Environmental Hazards*, 13(4), 1-26. https://doi.org/10.22067/geoeh.2024.87847.1482. [In Persian]

Fanos, A. M., & Pradhan, B. (2019). A Spatial Ensemble Model for Rockfall Source Identification from High-Resolution LiDAR Data and GIS. *IEEE Access*, 7, 74570–74585. https://doi.org/10.1109/ACCESS.2019.2919977

Fanos, A. M., Pradhan, B., Alamri, A., & Lee, C.W. (2020). Machine Learning-Based and 3D Kinematic Models for Rockfall Hazard Assessment Using LiDAR Data and GIS. Remote Sensing, 12(11), 1755. https://doi.org/10.3390/rs12111755

Fanos, A. M., Pradhan, B., Mansor, S., Yusoff, Z. M., & Abdullah, A. F. B. (2018). A hybrid model using machine learning methods and GIS for potential rockfall source identification from airborne laser scanning data. Landslides, 15(9), 1833-1850. https://doi.org/10.1007/s10346-018-0990-4

Ferrari, F., Giacomini, A., & Thoeni, K. (2016). Qualitative rockfall hazard assessment: a comprehensive review of current practices. *Rock Mechanics and Rock Engineering*, 49, 2865-2922. https://doi.org/10.1007/s00603-016-0918-z

He, X., Lang, Q., Zhang, J., Zhang, Y., Jin, Q., & Xu, J. (2025). Interpretable Machine Learning for Explaining and Predicting Collapse Hazards in the Changbai Mountain Region. Sensors (Basel, Switzerland), 25(5), 1512. https://doi.org/10.3390/s25051512

Jokar sarhangi, E., & Alizadeh, A. (2014). An Investigation into the Efficiency of Frequency Ratio and Information Value Models in Rockfall Hazard Zonation (Case Study: Haraz Valley). *Geographical planning of space quarterly journal*, 4(11), 151- 163. https://gps.gu.ac.ir/article_7402.html?lang=en. [In Persian]

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

Mahmoodzadeh, A., Alanazi, A., Mohammed, A., Ibrahim, H., Alqahtani, A., Alsubai, S., & Elhag, A. (2023). Comprehensive analysis of multiple machine learning techniques for rock slope failure prediction. Journal of Rock Mechanics and Geotechnical Engineering, 16(11), 4386-4398. https://doi.org/10.1016/j.jrmge.2023.08.023

Mao, K., Zhu, J., Su, L., Cai, G., Li, Y., & Dong, Z. (2023). FinalMLP: An Enhanced Two-Stream MLP Model for CTR Prediction. Proceedings of the AAAI Conference on Artificial Intelligence, 37(4), 4552–4560. https://doi.org/10.1609/aaai.v37i4.25577

Mokhtari, M., Abedian, S. and Almasi, S. A. (2020). Rockfall Susceptibility Mapping Using Artificial Neural Network, Frequency Ratio, and Logistic Regression: A Case Study in Central Iran, Taft County. *AUT Journal of Civil Engineering*, 4(1), 63-80. https://doi.org/10.22060/ajce.2019.15368.5538

Momeni, A. and Heidari, M. (2021). Rockfall hazard analysis in the Ganjnameh tourist area, west of Iran. *Journal of Natural Environmental Hazards*, 10(28), 1-14. https://doi.org/10.22111/jneh.2021.31304.1553. [In Persian]

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In Proceedings of *the 27th International Conference on Machine Learning* (ICML-10), 807–814. Omnipress. https://www.cs.toronto.edu/~hinton/absps/reluICML.pdf

Nanehkaran, Y. A., Licai, Z., Chen, J., Azarafza, M., & Yimin, M. (2022). Application of artificial neural networks and geographic information system to provide hazard susceptibility maps for rockfall failures. Environmental Earth Sciences, 81(19), 475. https://doi.org/10.1007/s12665-022-10603-6

Nguyen, K. A., Jiang, Y.J., Huang, C.S., Kuo, M.H., & Chen, W. (2024). Leveraging Internet News-Based Data for Rockfall Hazard Susceptibility Assessment on Highways. *Future Internet.* 16(8), 299. https://doi.org/10.3390/fi16080299

Paknejad F., Ahmadabadi A., Qanavati E., & zahmatkesh H. (2023). Zoning of Susceptible Areas with Debris Flows Using Machine Learning Algorithms (Case study of Tangarah Basin -Golestan Province). *Environmental Erosion Research Journal*, 13 (1), 1-28. http://magazine.hormozgan.ac.ir/article-1-702-fa.html. [In Persian]

Rahimi, M., Esfandiari, F., Zakizadeh, O., & Vahabzadeh Zargari, M. (2025). Identification of Rockfall Hazard Zones Along the Meshginshahr-Ahar Road Using Artificial Neural Networks. *Journal of Natural Environmental Hazards*, 14(46), 23–42. https://doi.org/10.22111/jneh.2025.50028.2072. [In Persian]

Rodriguez, V., Ghimire, B., Rogan, J., Chica Olmo, M., & Rigol Sánchez, J.P. (2012). An assessment of the effectiveness of a Random Forest classifier for land cover classification. *ISPRS Journal of Photogram Remote Sens*, 67, 9-104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Roostaei, S., Mokhtari, D. and Khodaiae, F. (2015). The application of Logistic Regression in the zoning of slope instabilities on the mountain roads (case study: Dare Diz Strait). *Journal of Natural Environmental Hazards*, 4(6), 89-103. https://doi.org/10.22111/jneh.2016.2525. [In Persian]

Sepahvand, A., & Beiranvand, N. (2024). Landslide susceptibility mapping using various soft computing techniques (Case study: A part of Haraz Watershed). *Water and Soil Management and Modelling*, 4(2), 261-278. https://doi.org/10.22098/mmws.2023.12678.1263. [In Persian]

Shirzadi, A., Solaimani, K., Habibnejad Roshan, M., Kavian, A., & Ghasemian, B. (2016). Comparison of Logistic Regression, Frequency Ratio, and AHP in Rockfall Susceptibility Mapping. *Journal of Watershed Management Research*, 6(12), 193– 204. https://jwmr.sanru.ac.ir/article-1-570-en.html. [In Persian]

Youssef, A. M., & Pourghasemi, H. R. (2021). Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geoscience Frontiers*, 12(2), 639-655. https://doi.org/10.1016/j.gsf.2020.05.010

Zhang, S., Tan, S., Liu, L., Ding, D., Sun, Y., & Li, J. (2023). Slope Rock and Soil Mass Movement Geological Hazards Susceptibility Evaluation Using Information Quantity, Deterministic Coefficient, and Logistic Regression Models and Their Comparison at Xuanwei, China. *Sustainability*, 15(13), 10466. https://doi.org/10.3390/su151310466.

آماده انتشار (بدون شماره مجله)