

## Original Article



# Evaluation of Machine Learning Techniques: Artificial Neural Network, Random Forest, and Support Vector Machine in Prediction of Alborz-Azerbaijan Earthquakes

Received: 2024.07.06

Accepted: 2024.11.05

Salma Ommi\*

Department of  
Environmental Planning  
and Design, Environmental  
Science Research Institute,  
Shahid Beheshti University,  
Tehran, Iran

## EXTENDED ABSTRACT

**Introduction:** One of the most destructive environmental hazards is earthquakes. Therefore, predicting this hazard to reduce its consequences and improve crisis management is one of the most important goals for researchers. By using seismicity indices and applying machine learning techniques, researchers can reveal seismic behavior patterns in a region. These methods have proven especially effective at modeling the nonlinear behavior of seismic data and have thus become important tools for understanding natural phenomena.

**Materials and methods:** In the present study, a catalog of earthquakes from the Alborz-Azerbaijan seismotectonic province covering the period from January 1, 1995, to January 23, 2024, has been compiled. Earthquake magnitudes have been converted to torque magnitudes to homogenize the catalog. Subsequently, based on the temporal variations in the completeness magnitude, the threshold magnitude for the study in the Alborz-Azerbaijan seismotectonic province has been determined to ensure the necessary accuracy for analysis. Three machine learning methods—Artificial Neural Network, Random Forest, and Support Vector Machine—were selected to predict the time and magnitude of earthquakes. Recognizing that some machine learning methods require feature definition, nine representative indices of seismic behavior were estimated for the Alborz-Azerbaijan earthquake catalog to serve as input for the chosen methods. Following the implementation of these techniques, the estimation error rate was calculated and reported using four types of error metrics: F1 Score, Recall, Precision, and Accuracy.

**Results and discussion:** Machine learning in this study was conducted using 245 vectors formed by 9 indicators. These indices are stored in corresponding two-dimensional arrays, with each column representing a specific set of indices. Each data vector is associated with a binary label of 1 or 0; the label "1" indicates the occurrence of at least one earthquake with a magnitude equal to or greater than the moment magnitude threshold of 5.5, while the label "0" indicates the absence of seismic activity for earthquakes with magnitudes less than 5.5. In this research, 80% of the data vectors were used for model training, and 20% were used for testing. The findings, regarding the estimated true and false alarm error values for each of the machine learning techniques applied to the seismic data of Alborz-Azerbaijan, demonstrate the success of all three techniques in predicting events recorded in the seismic catalog of Alborz-Azerbaijan. Generally, an accuracy exceeding 95% was achieved for all three methods.

**Conclusion:** The findings indicate the success of these techniques in estimating the cycle of stress accumulation and release associated with seismic activity in the Alborz-Azerbaijan geotechnical province. The accuracy of all three methods shows only a small difference, reflecting the high performance of machine learning techniques. For the seismic data of Alborz-Azerbaijan, the Random Forest method exhibits slightly higher accuracy. The accuracy values obtained from the selected methods in this research suggest that the optimal machine learning method depends on the diversity and quantity of the data. In the context of natural hazard data, particularly geophysical hazards, the differences in success levels among machine learning methods are influenced by the tectonic and geological characteristics of the environment. Furthermore, the results of this study demonstrate that utilizing machine learning techniques for preparedness and mitigation of environmental consequences, as well as for earthquake crisis management, is promising.

**Keywords:** Artificial Neural Network, Random Forest, Support Vector Machine, Alborz-Azerbaijan seismotectonic province

**How to cite this article:**  
Ommi, S., 2025. Evaluation of machine learning techniques: artificial neural network, random forest, and support vector machine in prediction of Alborz-Azerbaijan earthquakes. *Adva. Environ. Sci.* 23(1):259-274.

\* Corresponding Author Email Address: s\_ommi@sbu.ac.ir

DOI: 10.48308/envs.2025.236179.1412



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license <https://creativecommons.org/licenses/by/4.0/>.

## ارزیابی تکنیک‌های یادگیری عمیق ماشینی در تخمین خطر لرزه‌خیزی استان زمین‌ساختی البرز- آذربایجان

تاریخ دریافت: ۱۴۰۳/۰۴/۱۶

تاریخ پذیرش: ۱۴۰۳/۰۸/۱۵

سلما امی\*

گروه برنامه ریزی و طراحی محیط زیست، پژوهشکده علوم محیطی، دانشگاه شهید بهشتی، تهران، ایران

### چکیده مبسوط

**سابقه و هدف:** یکی از مخرب‌ترین مخاطرات محیطی، زلزله است. از اینرو پیش بینی و پیش‌یابی این مخاطره برای کاهش پیامدهای آن و بهبود امر مدیریت بحران در زمره مهم‌ترین اهداف پژوهشگران محسوب می‌شود. با استفاده از شاخص‌های لرزه‌خیزی و بکارگیری تکنیک‌های یادگیری ماشینی، الگوهای رفتار لرزه‌خیزی در یک منطقه لرزه‌زا آشکار می‌شود. این روش‌ها توانمندی ویژه‌ای در مدل‌سازی رفتارهای غیرخطی داده‌های لرزه‌ای نشان داده‌اند. از اینرو امروزه به ابزاری برای درک پدیده‌های طبیعی تبدیل شده‌اند.

**مواد و روش‌ها:** در مطالعه حاضر کاتالوگ زلزله‌های استان زمین‌ساختی البرز-آذربایجان در بازه زمانی اول ژانویه ۱۹۹۵ تا ۲۳ ژانویه ۲۰۲۴ تدوین شده است. تبدیلات بزرگای زلزله‌ها به بزرگای گشتاوری به منظور همگن نمودن کاتالوگ انجام شده است. سپس براساس تغییرات زمانی مقدار بزرگای کامل بودن، آستانه بزرگای مطالعه در استان زمین‌ساختی البرز-آذربایجان تعیین شده است تا در ادامه دقت لازم برای تجزیه و تحلیل‌ها فراهم گردد. سه روش یادگیری ماشینی شامل شبکه عصبی مصنوعی (Artificial Neural Network)، جنگل تصادفی (Random Forest) و ماشین بردار پشتیبان (Support Vector Machine) برای پیش بینی زمان و بزرگای زلزله‌ها انتخاب شدند، نه شاخص معرف رفتار لرزه‌خیزی برای کاتالوگ زلزله‌های البرز- آذربایجان برآورد شد تا پس از اجرای تکنیک‌های مذکور میزان خطای برآوردها در قالب چهار نوع برآورد خطا (F1 Score, Recall, Precision, Accuracy) محاسبه و گزارش گردد.

**نتایج و بحث:** یادگیری ماشینی در این مطالعه با استفاده از ۲۴۵ بردار در قالب ۹ شاخص انجام شده است. برچسب "۱" نشان دهنده وقوع حداقل یک زمین لرزه با بزرگی برابر یا بالاتر از آستانه بزرگای گشتاوری ۵.۵ است، در حالی که برچسب "۰" نشان دهنده عدم وجود فعالیت لرزه‌ای زمین لرزه‌هایی با بزرگای کمتر از ۵.۵ است. در این پژوهش ۸۰ درصد از بردارهای داده برای آموزش مدل بکارگرفته می‌شود و ۲۰٪ از بردارهای داده به عنوان آزمون استفاده می‌شود. یافته‌ها از مقادیر حاصل از خطای هشدارهای درست و غلط تخمین زده شده برای هر یک از تکنیک‌های یادگیری ماشینی برای داده‌های لرزه‌ای البرز- آذربایجان، حاکی از موفقیت هر سه تکنیک یادگیری ماشینی در پیش بینی رویدادهای کاتالوگ لرزه‌ای البرز- آذربایجان است. بطور کلی، دقت بالای ۹۵٪ برای هر سه تکنیک حاصل شده است.

**نتیجه‌گیری:** یافته‌ها معرف موفقیت این تکنیک‌ها در تخمین چرخه انباشت تنش و رهاسازی مرتبط با فعالیت لرزه‌ای در استان زمین‌ساختی البرز- آذربایجان است. با اختلاف اندک برای داده‌های لرزه‌ای البرز- آذربایجان، دقت روش جنگل تصادفی (Random Forest)، بالاتر حاصل شده است. مقادیر دقت حاصل از بکارگیری روش‌های منتخب در این پژوهش حاکی از آن است روش بهینه در تکنیک یادگیری ماشینی به تنوع داده‌ها و تعداد آنها بستگی دارد. در مورد داده‌های مخاطرات طبیعی، به ویژه مخاطرات ژئوفیزیکی، تفاوت در سطح موفقیت روش‌های یادگیری ماشینی تابعی از ویژگی‌های تکنونیک و زمین‌شناسی محیط است. بکارگیری تکنیک‌های یادگیری ماشینی در جهت آمادگی و کاهش پیامدهای محیطی و مدیریت بحران زلزله نوید بخش است.

**واژه‌های کلیدی:** یادگیری ماشینی، شبکه عصبی مصنوعی (Artificial Neural Network)، جنگل تصادفی (Random Forest)، ماشین بردار پشتیبان (Support Vector Machine)، استان زمین‌ساختی البرز-آذربایجان.

استناد به این مقاله: امی، س. ۱۴۰۴. ارزیابی تکنیک‌های یادگیری عمیق ماشینی در تخمین خطر لرزه‌خیزی استان زمین‌ساختی البرز- آذربایجان. فصلنامه علوم محیطی نوین. ۲۳(۱): ۲۵۹-۲۷۴.

\* Corresponding Author Email Address: s\_ommi@sbu.ac.ir

DOI: 10.48308/envs.2025.236179.1412



## مقدمه

Lomax, Michellini ) هشدار زلزله (Marfurt 2017 and Jozinović 2019)، نظارت بر آتشفشان (Lomax et al., 2019; Ren et al., 2020b)، مورد توجه قرار گرفته است.

موفقیت در این تحقیقات، ایده پیش بینی کنون بینی (Nowcasting Prediction) زلزله را توسعه داده است. در این روش‌ها از ویژگی‌های لرزه‌ای برای استنتاج وضعیت فعلی چرخه زلزله استفاده می‌شود (Asim et al., 2017; Pasari and Mehta 2018; Rundle et al., 2022). در این رویکرد از مشاهدات لرزه خیزی زلزله‌های اخیر برای تخمین احتمال شرطی زمین لرزه بزرگ، با این فرض که پس از وقوع تعدادی زمین لرزه کوچک و یا متوسط، ممکن است یک زلزله بزرگ رخ دهد، استفاده می‌شود. روش‌های پیش بینی "کنون بینی" براساس توصیف و برآورد متغیرهایی برای تامین هدف پیش بینی زلزله در مناطق لرزه خیز کالیفرنیا (Rundle et al., 2022)، ژاپن (Nanjo, 2020)، هند (Pasari, 2019) و زاگرس شمالی ایران (Omni and Hashemi, 2024) بکار گرفته شده است.

ایده مدل‌های پیش بینی کنون بینی متکی بر داده‌های واقعی و زمان کنونی هستند که می‌توان آنها را به به روش‌های پیشیابی تعمیم داد. نتایج تحقیقات آزمایشگاهی اخیر موفقیت پیش بینی رخدادهای لرزش و لغزش را با استفاده از تکنیک‌های یادگیری ماشین متکی بر سری زمانی تایید کرده است (Rouet-Leduc et al., 2017). مطالعات دیگری نیز موفقیت سری‌های زمانی را در پیش بینی گسیختگی‌های کوچک در کالیفرنیا آزمودند، نتایج این پژوهش‌ها بر وجود همبستگی میان رخدادهای کوچک و بزرگ تکیه دارند (Rundle et al., 2021). الگوهای رخداد توالی زمانی وقوع زلزله‌ها را براساس تنوع رفتار رژیم لرزه‌ای متفاوت ترسیم می‌نماید (Rundle et al., 2019, Rundle et al.,

در امر مدیریت بحران و خدمات اورژانسی و در فاز آمادگی قبل از زلزله‌های بزرگ، تخمین پتاسیل لرزه‌خیزی از اهمیت ویژه برخوردار است. مطالعه روند رخدادها در یک منطقه لرزه‌ای، درک خوبی از فرایند وقوع زلزله و وضعیت کنونی ناحیه لرزه‌زا به همراه دارد. این مهم اطلاعات اولیه را برای پیش‌بینی زلزله‌های بزرگ فراهم می‌کند. روند رخداد زلزله در فرضیه بازگشتی، پس از زلزله سانفرانسیسکو که با بزرگای ۷.۹ در سال ۱۹۰۶ رخ داد، ارائه شد (Lawson and Reid, 1910). بر اساس این نظریه، مسئله پیش بینی رخداد این پدیده طرح گردید. یکی از عمده‌ترین معضلات پیش‌بینی زمین‌لرزه این است که علی‌رغم اینکه فرآیندهای اساسی که باعث ایجاد آن‌ها می‌شوند یکسان هستند، الگوهای زلزله بسته به مکان می‌توانند بسیار متفاوت باشند. این مهم بر میزان توانمندی روش‌های تخمین پتاسیل لرزه‌خیزی تاثیرگذار است. بنابراین، بسیاری از مطالعات، همبستگی زمین‌لرزه‌ها در فضا و زمان را برای تخمین آستانه مقاومت یک ناحیه لرزه‌زا، تحلیل کرده‌اند، تا بر اساس آن زمان شکست با تکیه بر ویژگی‌های فیزیکی و زمین‌شناختی هر منطقه برآورد گردد (Kagan, 1982; Brehm and Braile, 1998). گرچه همچنان تاکید می‌شود که فاصله زیادی تا پیش بینی زلزله وجود دارد، اما نتایج پژوهش‌های علمی اخیر متمرکز بر زمین‌لرزه‌های آزمایشگاهی بارقه امید کاربردهای تکنیک‌های یادگیری ماشینی را در علوم زمین ایجاد کرده است (Johnson et al., 2021). در دو دهه گذشته تلاقی الگوریتم‌های جدید یادگیری ماشینی و واحدهای پردازش و همچنین در دسترس بودن مجموعه داده‌های گسترده و اغلب پیوسته، انقلابی در تجزیه و تحلیل داده محور ایجاد کرده است. این نتایج موفقیت آمیز در مسائل زمین شناسی (Ren et al., 2020a)، تشخیص و شناسایی فاز و مکان شناسایی سازند زمین شناسی (Zhao, Li and

## مواد و روش‌ها

### منطقه مورد مطالعه

جنابایی گسل‌های استان لرزه‌زمین ساختی البرز-آذربایجان، این منطقه را به عنوان یکی از فعال‌ترین ایالات زمین‌ساختی فلات ایران معرفی می‌کند. کوتاه شدگی پوسته در شمال ایران، محدوده چین خورده رشته کوه البرز را به طول تقریبی ۶۰۰ کیلومتر و عرض ۱۰۰ کیلومتر بوجود آورده است. نرخ این کوتاه شدگی در بخش غربی استان زمین‌ساختی البرز-آذربایجان معادل  $9 \pm 2$  میلی متر سال است (Hessami et al., 2003, Masson et al., 2005). این استان زمین‌ساختی که رشته کوه البرز را در بر گرفته است از شمال به گسل کاسپین، از جنوب و جنوب خاوری به دشت کویر ایران مرکزی و گسل تروند، و از خاور نیز به کوه‌های کپه داغ محدود می‌شود (Mirzaei, Mengtan and Yuntai 1998). در باختر البرز نیز می‌توان به گسل رودبار- منجیل اشاره کرد که مسبب زلزله فاجعه‌بار ۱۹۹۰ رودبار ( $M_w = 7.3$ ) است (Berberian and Walker, 2010). کاتالوگ زلزله‌های فلات ایران حاکی از وقوع زمین‌لرزه‌های بزرگ دیگر پیش از وقوع این زمین لرزه در رشته کوه‌های البرز غربی است. گسل‌های استان لرزه زمین‌ساخت البرز- آذربایجان با سازوکار بیشتر راستالغز چپ گرد شناخته می‌شوند. گرچه برای گستره جنوبی البرز خاوری گسل تروند با سازوکار راندگی گزارش شده است. از گسل‌های معکوس به سمت جنوب به عنوان مثال می‌توان به گسل شمال تهران، شمال قزوین و گسل مشا نیز در این محدوده اشاره نمود. شواهد موجود در این منطقه نشان می‌دهد که تغییر شکل پوسته در این ناحیه عمدتاً به شکل رخداد لرزه‌ای است (Jackson and McKenzie, 1988; Masson et al. 2005). البرز خاوری یکی از گستره‌های جنبا است که نقش مهمی در برهمکنش با محیط اطراف، به ویژه با پوسته کهن دریای کاسپین دارد (Hessami et al., 2003) مطالعه تغییرات پتانسیل لرزه‌ای در این منطقه از اهمیت ویژه‌ای برخوردار

(2022). براساس تدقیق این روابط ریاضی نرخ زمین لرزه‌ها تعیین می‌شود تا بتوانند ماهیت غیرخطی رخداد زلزله را توصیف کنند. تکنیک‌های یادگیری ماشینی که نوعی هوش مصنوعی است از طریق توسعه الگوریتم‌ها به درک الگوها و ارتباطات در داده‌ها کمک می‌کند. اساس این روش یادگیری از تجربیات به دست آمده از داده‌ها است. در واقع از طریق این تکنیک می‌توان رفتار آینده داده‌ها را بدون تکیه بر یک معادله از پیش تعیین شده به عنوان مدل تخمین زد. این ویژگی منجر به موفقیت قابل توجهی در موضوع پیش بینی مخاطرت محیطی از جمله زلزله شده است که به قضاوت در مورد اینکه آیا پس از زلزله رخ داده، وقوع زلزله بزرگ دیگر محتمل است یا خیر، کمک می‌کند. به عنوان نمونه در پژوهشی انجام شده بر داده‌های لرزه‌ای هندوکش، نه شاخص لرزه خیزی تعریف شده و سپس بر اساس مدل شبکه عصبی مصنوعی پتانسیل لرزه‌خیزی یک منطقه تعیین شده است، نتایج موفقیت این شیوه‌ها را تایید نمودند (Asim et al., 2017).

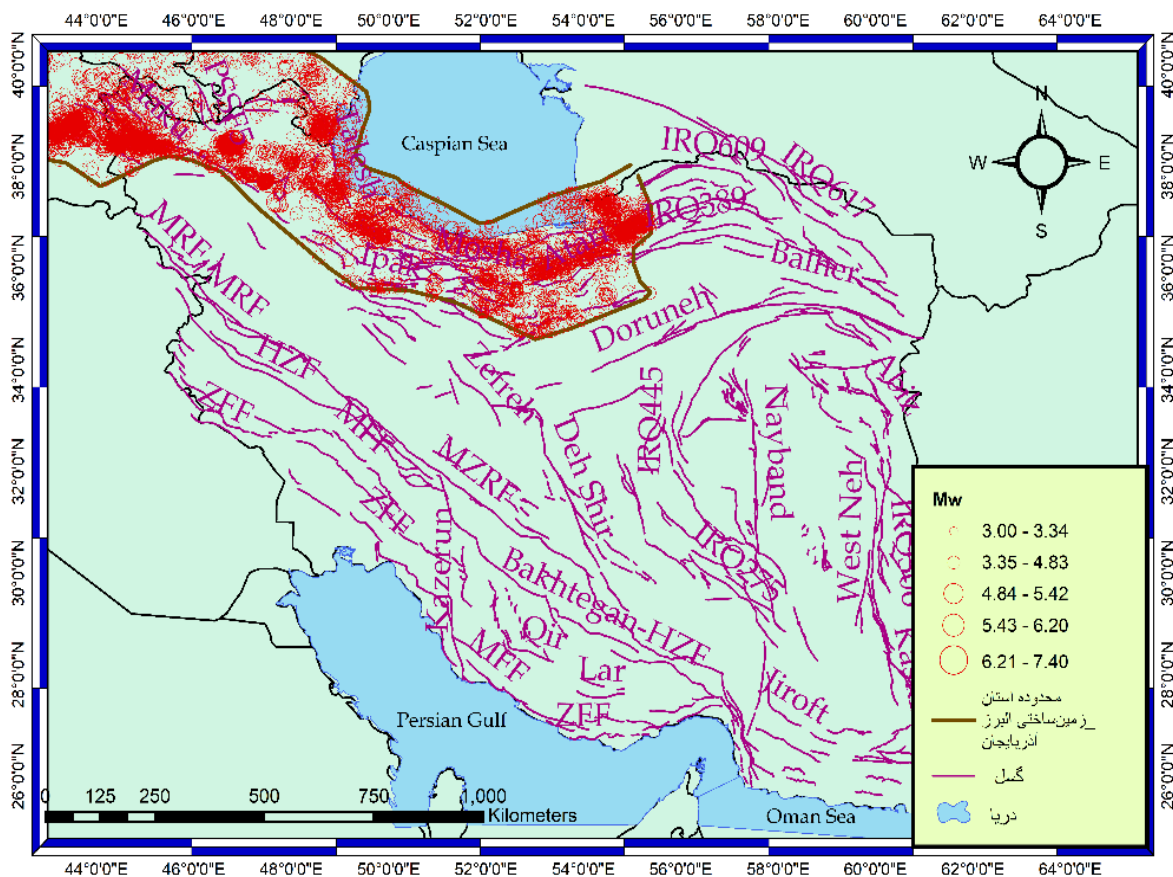
بر مبنای تکنیک‌های یادگیری ماشینی شبکه عصبی مصنوعی (Artificial Neural Network)، جنگل تصادفی (Random Forest)، و ماشین بردار پشتیبان (Support Vector Machine) پژوهشی با هدف سنجش میزان موفقیت تکنیک‌های یادگیری عمیق ماشینی در پیش‌بینی زمان و بزرگی زمین‌لرزه‌ها بر اساس شاخص‌های لرزه‌خیزی در استان لرزه‌ساخت زمین‌ساخت زاگرس شمالی ایران (Ommi and Hashemi, 2024)، صورت گرفت، در این مطالعه برای تکنیک شبکه عصبی عصبی بالاترین میزان دقت ۹۶٪ برای پیش‌بینی زلزله‌های بزرگتر از ۵.۵ حاصل شد. این پژوهش با هدف ارزیابی تکنیک‌های یادگیری ماشینی در استان زمین‌ساختی البرز- آذربایجان و مقایسه نتایج با سایر مطالعات، انجام شده است.

زمین‌ساختی البرز-آذربایجان و تعیین محدوده مطالعاتی بهره گرفته شده است. شکل ۱ توزیع زلزله‌های رخ داده در بازه زمانی ۱۹۹۵-۲۰۲۴ را در این استان زمین‌ساختی البرز-آذربایجان نشان می‌دهد. زلزله‌های بزرگتر از ۵.۵ کاتالوگ لرزه‌ای در جدول ۱ فهرست شده است.

تغییرات فراوانی رویدادهای لرزه‌ای براساس بزرگی آنها از جمله ویژگی‌های رفتار لرزه‌خیزی محسوب می‌شود. تغییرات فراوانی رویدادهای البرز در شکل ۲ نمایش داده شده است. همانطور که از شکل مشاهده می‌شود استان زمین‌ساختی البرز-آذربایجان دارای حوادث نادر و بزرگ‌های مهم است. این ویژگی می‌تواند بر عملکرد روش‌های یادگیری ماشینی اثرگذار باشد.

است از اینرو، تحقیق امکان پیش بینی فعالیت‌های لرزه‌ای در شکل زلزله‌های بزرگ موضوع کنجکاوای این پژوهش است.

در این مطالعه برای تدوین داده‌های زمین لرزه از ۱۹۹۵ تا سال ۲۰۱۴ از کاتالوگ منتشر شده موسوی (Mousavi-Bafrouei and Mahani, 2020) و برای تکمیل بازه زمانی داده‌ها تا ۲۳ ژانویه ۱۴۰۲۴ داده‌های لرزه‌ای شبکه لرزه نگاری مؤسسه ژئوفیزیک دانشگاه تهران (Iranian Seismological Center) (<http://irsc.ut.ac.ir/bulletin.phb>) استفاده شده است. سپس با تبدیل بزرگی زلزله‌ها به بزرگی گشتاوری ( $M_w$ ) کاتالوگ یکنواخت شده است. در این مطالعه از تقسیمات ایالات لرزه زمین‌ساختی میرزایی (۱۹۹۸) برای تفکیک زلزله‌های استان

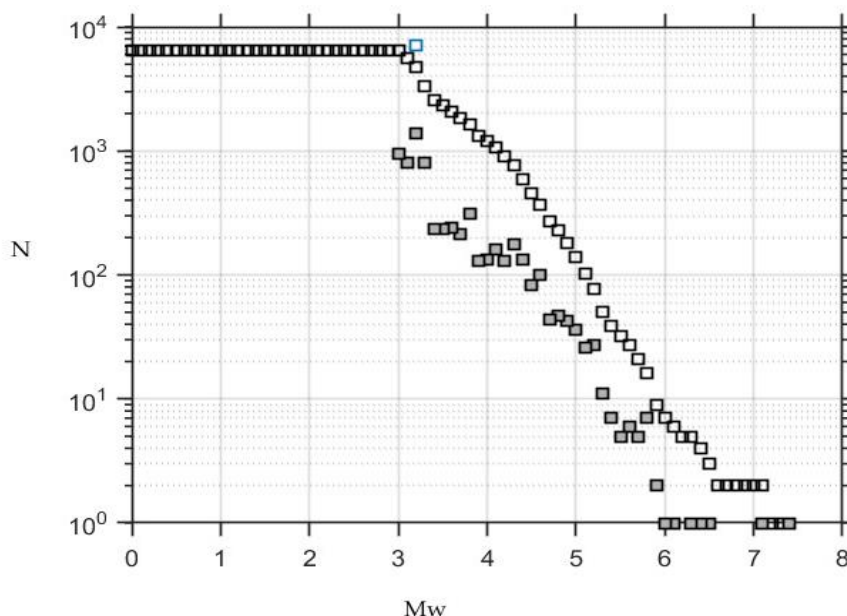


شکل ۱- توزیع زمین لرزه‌های با بزرگی  $M_w \geq 3$  مورد مطالعه و موقعیت گسل‌های اصلی فعال.

Fig. 1- Distribution of earthquakes with magnitudes  $M_w \geq 3$  under study and the locations of major active faults

جدول ۱- زمین لرزه‌های با بزرگای گشتاوری ( $M_W$ ) برابر و بیشتر از ۵.۵ استان لرزه زمین‌ساخت البرز از ۱ ژانویه ۱۹۹۵ تا ۲۳ ژانویه ۲۰۲۴.  
 Tabel 1. Earthquakes with a moment magnitude ( $M_W$ ) equal to or greater than 5.5 in the Alborz tectonic province from January 1, 1995, to 23.

Date	Time	Latitude	longitude	Depth	MW
1997-02-28	12:57:22	38.12	48.08	9	6.1
1998-07-09	2:19:21	38.72	48.53	26	5.9
1999-12-03	5:06:54	40.49	42.35	6	5.7
2002-11-11	12:13:41	35.87	52.26	15	5.8
2004-05-28	12:38:43	36.25	51.57	27	6.3
2004-11-07	9:46:19	37.15	54.43	30	5.6
2010-08-27	7:23:46	35.49	54.47	6.7	5.8
2011-11-09	7:23:33	38.43	43.23	5	5.7
2011-11-23	10:41:23	38.72	43.51	16	7.1
2011-11-23	10:48:17	38.75	43.60	9	5.8
2011-11-23	10:56:49	38.82	43.43	18	5.8
2011-11-23	11:32:41	38.80	43.43	18	5.9
2011-11-23	8:45:35	38.63	43.08	5	6
2011-11-25	2:55:08	38.81	43.62	14	5.6
2012-08-11	12:23:20	38.43	46.81	9	6.5
2012-08-11	12:34:39	38.40	46.80	10	6.4
2012-11-07	6:26:33	38.46	46.57	10	5.6
2019-11-07	10:47:04	37.71	47.52	8.5	5.83
2020-02-23	5:52:59	38.45	44.52	6	5.66
2020-02-23	4:00:30	38.49	44.56	6	5.83
2023-01-28	6:14:45	38.53	44.84	11	5.83

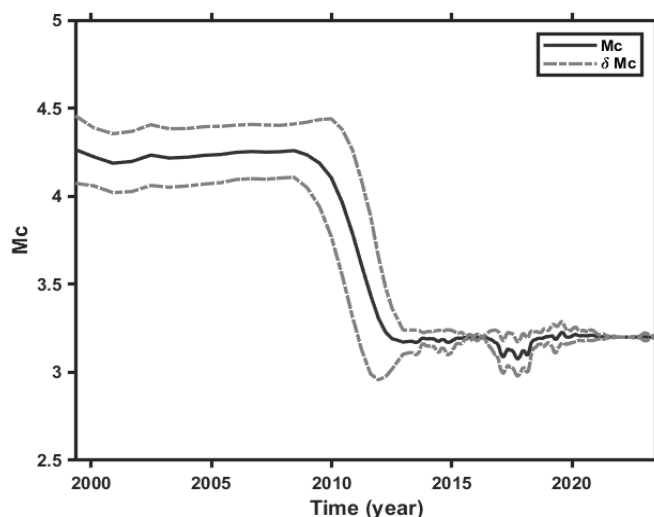


شکل ۲- فراوانی تجمعی رویدادهای لرزه‌ای  $N$  در بزرگای  $M_W$  در بازه زمانی ژانویه ۱۹۹۵ تا ۲۳ ژانویه ۲۰۲۴ برای کاتالوگ لرزه‌ای استان زمین‌ساختی البرز-آذربایجان.

Fig. 2- Cumulative frequency of seismic events  $N$  with magnitudes  $M_W$  during the period from January 1995 to January 23, 2024, for the seismic catalog of the Alborz-Azerbaijan tectonic province.

از روش بیشینه انحنای (Maximum curvature (MAXC) (Woessner and Wiemer, 2005) برای تخمین بزرگای کامل بودن  $M_c$  استفاده و تغییرات نسبت به زمان آن در نظر گرفته شده است (شکل ۳). برای افزایش کیفیت داده‌ها، رویدادهای کوچکتر از  $M_c$  حذف شدند.

بزرگای کامل بودن ( $M_c$ )، به عنوان حداقل بزرگای قابل تشخیص برای تعیین دقت شبکه لرزه‌نگاری شناخته می‌شود. بر اساس روش‌های پردازش داده‌ها، متغیر  $M_c$  که متضمن نتایج با کیفیت بالا در آنالیزهای داده‌های لرزه‌ای است (Wiemer and Wyss, 2002)، برای کاتالوگ البرز-آذربایجان برآورد شده است (جدول ۲). در این تحقیق



شکل ۳- تغییرات بزرگای کامل بودن  $M_c$  کاتالوگ لرزه‌ای استان زمین‌ساختی البرز-آذربایجان در بازه زمانی ژانویه ۱۹۹۵ تا ۲۳ ژانویه ۲۰۲۴.  
Fig. 3- Changes in the completeness magnitude  $M_c$  of the seismic catalog for the Alborz-Azerbaijan tectonic province during the period from January 1995 to January 23, 2024.

جدول ۲- تغییرات بزرگای کامل بودن در بازه زمانی ۲۰۲۴-۱۹۹۵ به روش بیشینه انحنای (MAXC) برای کاتالوگ لرزه‌ای استان زمین‌ساختی البرز-آذربایجان.

Table 2. Changes in the completeness magnitude using the maximum curvature method for the period from 1995 to 2024 for the seismic catalog of the Alborz-Azerbaijan tectonic province.

$M_c$	1995-2012	2013-2024
Alborz -Azerbaijan	4	3.6

به مانند مطالعات دیگر (Panakkat and Adeli 2007, Ommi and Hashemi 2024) مقدار  $n=100$  انتخاب شده است.

$$M_{mean} \quad (2)$$

میانگین بزرگای آخرین  $n$  زمین لرزه که در بازه زمانی رخ می‌دهد، دومین شاخص لرزه‌ای است.

$$\Delta M = \frac{\sum_i M}{n} \quad (2)$$

### شاخص‌های لرزه خیزی

ویژگی‌های لرزه‌ای به صورت ۹ شاخص در نظر گرفته می‌شود تا تکنیک‌های یادگیری عمیق ماشینی بر اساس اطلاعات آنها اجرا شود. این شاخص‌ها به شکل ذیل در نظر گرفته شده است:

$$(1) \text{ بازه زمانی } T$$

$T$  زمان سپری شده بین  $n$  امین زلزله و زلزله نخست در کاتالوگ انتخاب شده:

$$T = T_n - T_1 \quad (1)$$

لرزه خیزی در منطقه مورد مطالعه محاسبه می‌شود (معادله ۵).

$$\Delta M = M_{\max}^{\text{obs}} - M_{\max}^{\text{exp}} \quad (5)$$

مقدار مورد انتظار حداکثر بزرگای در آخرین رویداد  $n$  در کاتالوگ لرزه‌ای با استفاده از قانون گوتنبرگ-ریشتر با جایگزینی  $N = 1$  (معادله ۶) برآورد می‌شود:

$$M_{\max}^{\text{exp}} = \frac{a}{b} \quad (6)$$

(۶) مقدار  $\mu$

با در نظر گرفتن فرض بازگشت الاستیک، وقوع زمان در میان زلزله‌های بزرگ نسبتاً ثابت است. براین اساس فاصله زمانی بین زمین لرزه‌های بزرگ به عنوان زمین لرزه سرشتی شناخته می‌شود به عنوان یک ویژگی لحاظ می‌گردد. این مقدار که از عدم قطعیت تصادفی و شناختی برخوردار است معیار مهمی برای مطالعه پتانسیل لرزه‌خیزی است. میانگین زمان زمین لرزه‌های سرشتی از رابطه (۷) حاصل می‌شود که مانند سایر مطالعات (Asim et al., 2017) به عنوان یک ویژگی در نظر گرفته شده است.

$$\mu = \frac{t_i^{\text{characteristic}}}{n_i^{\text{characteristic}}} \quad (7)$$

(۷) مقدار  $c$

نتیجه نسبت انحراف معیار زمان‌های مشاهده شده یکی دیگر از شاخص‌های لرزه‌خیزی در این تحقیق است که بر اساس فواصل زمانی بین رویدادهای سرشتی تخمین زده شده است (معادله ۸).

$$c = \frac{\sigma_t}{\mu} \quad (8)$$

مقدار  $c$  نشان دهنده تفاوت بین میانگین زمان محاسبه شده و میانگین زمان مشاهده شده است. در بین زمین لرزه‌های مشخصه، یک مقدار کم  $c$  نشان‌دهنده تطابق مقادیر مشاهده شده با مقادیر تخمینی است

(۳) (b) شیب منحنی گوتنبرگ ریشتر

قانون کلاسیک گوتنبرگ-ریشتر (Gutenberg and Richter 1944) (معادله ۳) فراوانی زمین لرزه‌ها را بر اساس بزرگی آن‌ها در یک منطقه خاص و در طول دوره زمانی معین  $T$  تخمین می‌زند.

$$\log N(M) = a - bM \quad (3)$$

$N$  تعداد زمین لرزه‌هایی با بزرگی  $M$  در یک منطقه خاص رخ داده است. ضرایب  $a$  و  $b$  پارامترهای لرزه‌خیزی هستند، به نحوی که  $a$  معرف زلزله‌های با بزرگای بیشتر از صفر در منطقه و  $b$  شیب خط در نمودار  $\log N$  و معرف نسبت زلزله‌های کوچک به بزرگ است. بنابراین، ضریب  $b$  یکی از پارامترهای بنیادی به ویژه در محاسبه خطر و خطرپذیری لرزه‌ای محسوب می‌شود. این مقدار نشان‌دهنده سطح لرزه‌خیزی در منطقه است و معمولاً عددی نزدیک به ۱ است. این پارامتر توسط محققان برای تجزیه و تحلیل لرزه‌خیزی (فعال شدن یا سکون) استفاده شده است (Enescu and Ito 2001, Mignan and Broccardo 2020, Wang et al., 2021) و به عنوان پیش‌نشانگر زلزله قریب الوقوع معرفی شده است.

(۴) مقدار  $\eta$

فراوانی زلزله‌های کوچک و بزرگ معرف رفتار لرزه خیزی است. جمع میانگین انحراف مربع از خط رگرسیون با استفاده از قانون توان معکوس گوتنبرگ ریشتر با مشاهدات ثبت شده در کاتالوگ لرزه‌خیزی مقدار  $\eta$  نامیده می‌شود که به صورت زیر نشان داده شده است:

$$\eta = \frac{\sum (\log_{10} N_i - (a - bM_i))^2}{(n-1)} \quad (4)$$

(۵) مقدار  $\Delta M$

یکی از شاخص‌های عدم قطعیت، تفاوت بین حداکثر بزرگای مشاهده شده و حداکثر بزرگای مورد انتظار است که از طریق قانون گوتنبرگ-ریشتر بر اساس پارامترهای



گره‌های ورودی نشان دهنده تکامل شاخص‌های لرزه‌خیزی در زمان (۲) و (۳) رمزگذاری می‌کنند. لایه ورودی: اولین لایه داده‌های ورودی را دریافت می‌کند که می‌تواند یک بردار ویژگی باشد که نشان دهنده ویژگی‌های نقاط داده‌هایی است که باید طبقه‌بندی شوند. هر نورون در لایه ورودی مربوط به یک ویژگی است. در این مطالعه، لایه ورودی مدل شبکه عصبی مصنوعی پیشنهادی دارای ۹ نورون است که تعداد شاخص‌های لرزه‌ای را بیان می‌کنند. لایه‌های پنهان: در بین لایه‌های ورودی و خروجی، یک یا چند لایه مخفی وجود دارد. این لایه‌ها مسئول یادگیری الگوها و نمایش‌های پیچیده از داده‌های ورودی هستند. هر نورون در یک لایه پنهان، ورودی‌های لایه قبلی را می‌گیرد و وزن‌ها و بایاس‌ها را برای محاسبه خروجی اعمال می‌کند. لایه خروجی از تابع ذیل برآورد می‌گردد.

$$f(x) = \frac{1}{1+e^{-x}} \quad (10)$$

برای طبقه‌بندی باینری، لایه خروجی معمولاً از یک نورون منفرد با تابع فعال سازی سیگموئید (sigmoid)  $f(x)$  تشکیل شده است. خروجی این نورون مقداری بین ۰ و ۱ است که نشان دهنده احتمال تعلق نقطه داده به کلاس مثبت است. معمولاً یک آستانه اعمال می‌شود (مثلاً ۰.۵)، که در آن مقادیر بیشتر از آستانه به عنوان مثبت و مقادیر کمتر یا مساوی از آستانه به عنوان منفی طبقه‌بندی می‌شوند. در این مطالعه، با اجرای آزمایش‌های مختلف، تعیین شده است. معماری کلی شبکه‌های عصبی مصنوعی پیشنهادی در شکل ۴ نشان داده شده است.

#### ماشین بردار پشتیبانی

ماشین بردار پشتیبان یک الگوریتم یادگیری ماشین قدرتمند و پرکاربرد است که برای کارهای طبقه‌بندی باینری طراحی شده است. این روش به دلیل توانایی آن در مدیریت مجموعه داده‌های پیچیده و تولید نتایج دقیق محبوبیت پیدا کرده است (Abdullah and Abdulazeez

#### (۸) جذر مجموع انرژی لرزه‌ای $dE^{1/2}$

مجموع انرژی انباشته در منطقه به عنوان یک شاخص مهم از سطح لرزه‌ای است که بیشتر جهت برآورد بزرگا و زمان زلزله قریب الوقوع مورد استفاده قرار گرفته است (Brehm and Braile 1998, Ommi and Smirnov 2024). این شاخص توسط معادله ۹ (بر حسب  $erg$ ) برآورد شده است.

$$dE^{1/2} = \frac{\sum (10^{11.8+1.5M})^{1/2}}{T} \quad (9)$$

افزایش تصاعدی در انرژی لرزه‌ای قبل از گسیختگی فرایند فاز گسیختگی را توصیف می‌کند. در طول این مدت، میزان لغزش و اندازه گسیختگی هر دو به‌طور تصاعدی افزایش می‌یابند تا زمانی که گسیختگی به حدی برسد که امکان انتشار پویای گسیختگی را فراهم کند. مرحله هسته‌زایی ممکن است بین چند هفته تا چند ماه طول بکشد (Hulbert et al., 2020).

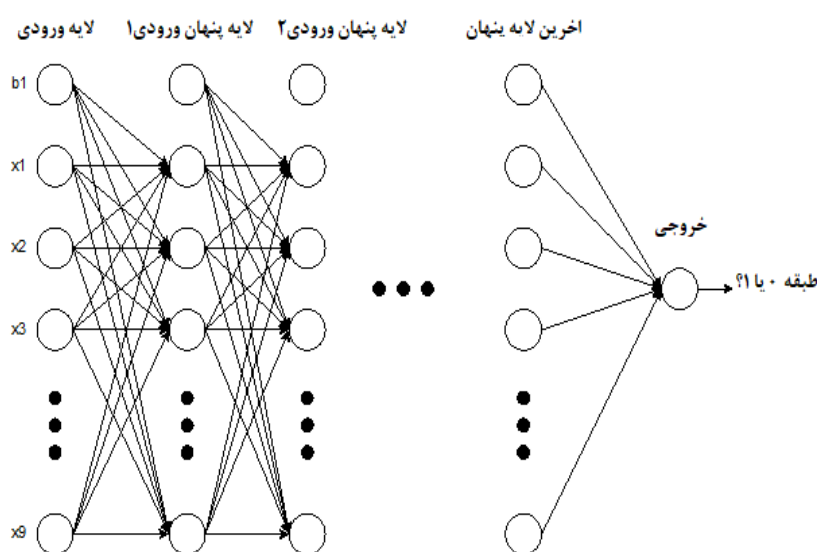
#### تکنیک‌های یادگیری ماشینی

##### شبک عصبی مصنوعی

در علم داده، یادگیری عمیق به عنوان یکی از قوی‌ترین تکنیک‌های برای حل مسائل سخت و حل نشدنی به سرعت در حال افزایش است و در امر پیش بینی مخاطرات محیطی از جمله پیش بینی زلزله مورد استفاده قرار می‌گیرد (Kong et al. 2019, Bergen et al. 2019). در این روش بردارهای داده را به یکی از دو کلاس یا دسته طبقه بندی می‌شود که معمولاً به صورت "۱" و "۰" نشان داده می‌شود. انواع مختلفی از شبکه عصبی مصنوعی در ادبیات وجود دارد. با این حال، در این مطالعه، از شبکه عصبی مصنوعی پیش‌خور استفاده می‌شود که متشکل از نورون‌های مصنوعی به هم پیوسته متعدد است که در لایه‌ها، سازماندهی شده است. برای اتصال کامل شبکه‌های پیش‌خور، سه لایه ورودی عبارتند از (۱) پارامترها یا شاخص‌های لرزه‌خیزی، در یک پنجره زمانی منحصر به فرد، (۲) مقادیر یک نشانگر در پنجره‌های زمانی متوالی، که (۳)

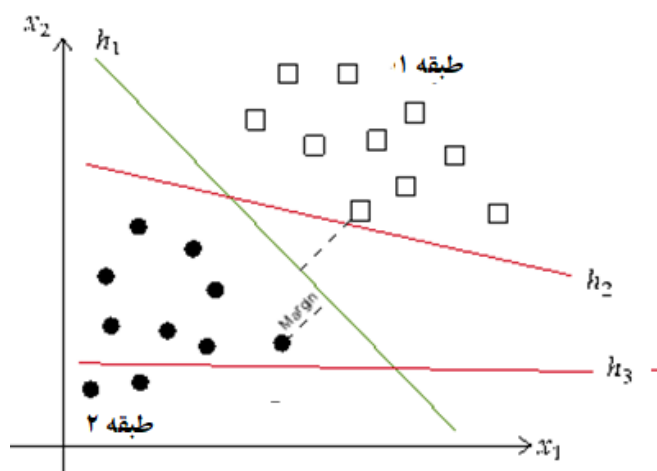
استحکام آن را در برابر نویز و موارد پرت نیز افزایش می‌دهد (Pisner and Schnyer 2020). هنگامی که مدل ماشین بردار پشتیبان با استفاده از ۹ پارامتر لرزه‌ای مربوط به ۱ ماه آموزش داده شد، می‌توان از آن برای پیش‌بینی کلاس نقاط داده جدید و نادیده بر اساس موقعیت آنها نسبت به مرز تصمیم استفاده کرد. در شکل ۵، مدل ماشین بردار پشتیبان نشان داده شده است که از تعریف حاشیه استفاده می‌کند تا مرز تصمیم‌گیری قوی‌تر را پیش‌بینی کند.

هدف اصلی ماشین بردار پشتیبان یافتن ابر صفحه‌ای بهینه است که نقاط داده کلاس‌های مختلف را در فضایی با ابعاد بالا به بهترین نحو از هم جدا می‌کند. ایده اصلی ماشین بردار پشتیبان، شناسایی یک مرز تصمیم است که حاشیه بین دو کلاس را به حداکثر می‌رساند. حاشیه به عنوان فاصله بین مرز تصمیم و نزدیکترین نقاط داده از هر کلاس تعریف می‌شود. با به حداکثر رساندن این حاشیه، ماشین بردار پشتیبان نه تنها توانایی مدل را برای تعمیم به داده‌های غیرقابل مشاهده بهبود می‌بخشد، بلکه



شکل ۴. ساختار شبکه عصبی مصنوعی در پیش‌بینی زلزله.

Fig. 4- Structure of the artificial neural network for earthquake prediction.



شکل ۵- شمای ماشین بردار پشتیبان برای پیش‌بینی زلزله.

Fig. 5- Schematic of the Support Vector Machine for earthquake prediction

## جنگل تصادفی

الگوریتم جنگل تصادفی یک شیوه یادگیری است که برای کارهای طبقه‌بندی باینری استفاده می‌شود. ایده اصلی جنگل تصادفی ایجاد تعداد زیادی درخت تصمیم در طول مرحله آموزش است. هر درخت با استفاده از یک زیر مجموعه تصادفی از ویژگی‌های اصلی و یک زیر مجموعه تصادفی از داده‌های آموزشی ساخته می‌شود. این فرآیند باعث ایجاد تنوع در میان درختان می‌شود و خطر حساسیت بیش از حد درختان به ویژگی‌ها یا نقاط داده خاص را کاهش می‌دهد. در طول مرحله پیش‌بینی، هر درخت در جنگل تصادفی به طور مستقل داده‌های ورودی را طبقه‌بندی و تصمیم می‌کند. بنابراین پیش‌بینی نهایی کمتر مستعد خطا است و دقت بالاتری نسبت به درخت‌های تصمیم فردی ارائه می‌دهد. تعداد درختان تصمیم در مجموعه، از طریق آزمون بر روی مجموعه داده تعیین می‌شود. در این مطالعه نتایج بر اساس ۳۰ درخت برای گروه انتخاب شده است.

در این پژوهش بر مبنای محاسبه شاخص‌های (ویژگی‌های) لرزه‌خیزی توصیف شده سه تکنیک یادگیری ماشینی شبکه عصبی مصنوعی، جنگل تصادفی، و ماشین بردار پشتیبان برای پیش‌بینی زلزله‌های کاتالوگ البرز- آذربایجان مورد آزمون قرار گرفت.

## نتایج و بحث

برای آموزش داده‌ها در یادگیری ماشین، شاخص‌های لرزه‌خیزی برای هر ماه بین ژانویه ۱۹۹۵ تا ۲۳ ژانویه ۲۰۲۴ محاسبه شده و براساس آن برچسب دهی انجام شده است. اندازه پنجره به منظور استخراج شاخص‌های لرزه‌ای از صد رویداد آخر از بازه زمانی یک ماهه استفاده می‌کند. تصمیم‌گیری برای پیش‌بینی‌های ماهانه پس از تجزیه و تحلیل جامع کاتالوگ زلزله بر اساس ویژگی‌های مبتنی بر فیزیک مسئله گسیختگی صورت می‌گیرد. این منجر به آموزش تمام تکنیک‌های یادگیری ماشینی فوق با استفاده از ۲۴۵ بردار داده می‌شود. شاخص‌های لرزه‌خیزی در آرایه‌های دو بعدی مربوطه ذخیره می‌شوند، جایی که هر ستون یک مجموعه از شاخص‌ها را نشان می‌دهد. هر بردار داده با یک برچسب باینری ۱ یا ۰ مرتبط است. برچسب "۱" نشان دهنده وقوع حداقل یک زمین لرزه با بزرگی برابر یا بالاتر از آستانه  $\tau = 5.5$  است، در حالی که برچسب "۰" نشان دهنده عدم وجود فعالیت لرزه‌ای با بزرگی  $\tau = 5.5$  و بیشتر است. جدول ۳ یک مجموعه نمونه از داده آموزشی نمونه را برای ۱۲ ماه از ژانویه تا دسامبر ۲۰۱۳ ارائه می‌دهد. این نمونه شامل بردارهای ورودی ۹ عنصری به همراه خروجی‌های دلخواه متناظر آنها است که نشان دهنده وقوع زلزله‌هایی با بزرگی برابر یا بیشتر از آستانه  $\tau$  است.

جدول ۳- نمونه داده آموزشی برای ۱۲ ماه بین ژانویه و دسامبر ۲۰۱۳ که مقادیر حاصل از نه ویژگی و خروجی‌های مورد نظر مربوطه را نشان می‌دهد، مقدار آستانه بزرگا  $\tau = 5.5$  است.

Table 3. Sample training data for 12 months between January and December 2013, showing the values obtained from nine features and the corresponding desired outputs, with the threshold magnitude value  $\tau = 5.5$ .

Date	$\Delta T$ (Days)	$M_{mean}$	$dE^{1/2} \times 10^{18}$	$M_{expected}$	$b$	$\Delta M$	$\mu$	$c$	$\eta$	label
January2013	168	3.91	0.007	6.25	1.51	1.21	0.65	0	1.51	1
February2013	93	3.48	0.003	5.29	1.91	1.91	0.19	0	1.91	0
March2013	51	3.43	0.004	4.99	1.94	2.32	0.08	0	1.94	0
April2013	55	3.48	0.004	5.21	1.67	2.03	0.21	27	1.67	0
May2013	57	3.42	0.004	5.07	1.80	2.19	0.07	27	1.80	0
June2013	63	3.39	0.003	5.05	1.81	2.20	0.05	55	1.81	0
July2013	64	3.39	0.003	4.97	1.95	2.36	0.30	0	1.95	0
August2013	59	3.44	0.004	5.25	1.63	1.98	0.58	51	1.63	0
September2013	83	3.37	0.002	4.86	2.07	2.52	0.19	0	2.07	0
October2013	93	3.42	0.002	4.91	2.04	2.47	0.24	0	2.04	0
November2013	93	3.41	0.002	4.95	1.98	2.39	0.28	0	1.98	0
Decber2013	84	3.45	0.002	5.05	1.90	2.28	0.38	0	1.51	0

پیش‌بینی‌های مثبت واقعی نسبت به تمام نمونه‌های مثبت واقعی در مجموعه داده Recall است. توانایی مدل را برای شناسایی صحیح همه موارد مثبت تعیین می‌کند (معادله ۱۳). امتیاز F1 (F1-score) امتیاز F1 میانگین هارمونیک Recall و precision است. تعادلی بین precision و Recall ایجاد می‌کند و زمانی مفید است که توزیع کلاسی نابرابر وجود داشته باشد. به عنوان مثال، در مجموعه داده‌های این مطالعه، حدود ۴۶٪ از بردارهای داده حاوی ۱ برچسب هستند در حالی که ۵۵٪ بقیه حاوی ۰ برچسب هستند که کل مجموعه داده را نسبتاً متعادل معرفی می‌کند (معادله ۱۴).

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (11)$$

$$precision = \frac{TP}{TP + FP} \quad (12)$$

$$recall = \frac{TP}{TP + FN} \quad (13)$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (14)$$

مقادیر حاصل از معیارهای دقت معرفی شده برای هر یک از تکنیک‌های یادگیری ماشینی روی داده‌های لرزه‌ای البرز - آذربایجان در جدول ۴ گزارش شده است. همانطور که از جدول ۴ مشخص است، دقت‌ها معرف موفقیت هر سه تکنیک یادگیری ماشینی در پیش‌بینی زلزله‌های کاتالوگ لرزه‌ای است. در میان این سه روش بکارگرفته شده، برای روش جنگل تصادفی بیشترین توفیق حاصل شده است. همانطور که جدول ۴ نشان می‌دهد، دقت بالای ۹۵٪ برای هر سه تکنیک حاصل شده است که میزان موفقیت بیشتر را در این مطالعه نسبت به مطالعه قبلی در شمال زاگرس (Omami and Hashemi 2024) نشان می‌دهد. این نتیجه را به ماهیت داده‌های کاتالوگ مرتبط است که حاکی از دقت بالاتر شبکه لرزه‌نگاری در برخی از ادوار در استان

جهت ارزیابی مدل‌ها از روش اعتبارسنجی متقاطع استفاده می‌شود. این یک تکنیک پرکاربرد در یادگیری ماشینی برای ارزیابی عملکرد و تعمیم یک مدل است. یکی از روش‌های رایج، اعتبارسنجی متقابل ۵ برابری است که شامل تقسیم مجموعه داده‌ها به پنج زیرمجموعه با اندازه تقریباً برابر است. در هر تکرار، چهار زیرمجموعه از این زیرمجموعه (۸۰ درصد از بردارهای داده) برای آموزش مدل استفاده می‌شود، در حالی که زیرمجموعه باقیمانده (۲۰ درصد از بردارهای داده) به عنوان مجموعه آزمون استفاده می‌شود. مزیت استفاده از اعتبارسنجی متقاطع این است که به تخمین قوی‌تری از عملکرد مدل کمک می‌کند. با میانگین‌گیری نتایج حاصل از تکرارهای متعدد، ارزیابی کمتر به تقسیم‌بندی داده‌های خاص وابسته می‌شود و نشانه قابل اعتمادتری از عملکرد مدل یادگیری ماشینی روی داده‌ها ارائه می‌دهد.

ارزش  $b$  و انرژی  $dE^{1/2}$  (بر حسب  $erg$ ) در جدول ۳ آمده است. در این جدول، مقدار  $T$  زمان سپری شده از  $n$  رویداد گذشته) برای سال‌های ابتدایی کاتالوگ لرزه‌ای به دست آمده است. این می‌تواند با گسترش شبکه لرزه‌ای مرتبط باشد. ویژگی  $\mu$  و  $C$  در برخی ردیف‌ها مقدار صفر به دست آمده است، به این معنی که تعداد زلزله‌های با بزرگای سرشتی در آن دسته ۱ یا کمتر بوده است. برای ارزیابی و عملکرد مدل‌ها در تخمین احتمال رخداد زلزله‌ها از چهارمعیار سنجش دقت، استفاده شده است. دقت یا Accuracy، نسبت نقاط داده به درستی پیش‌بینی شده به تعداد کل نقاط داده است (معادله ۱۱). این یک معیار کلی از عملکرد مدل پیش‌بینی زلزله در هر دو کلاس ۱ و ۰ ارائه می‌دهد. معیار دیگر معرف دقت Precision است که نسبت پیش‌بینی‌های مثبت واقعی (نمونه‌های مثبت پیش‌بینی شده درست) را به همه پیش‌بینی‌های مثبت (هم مثبت واقعی و هم مثبت کاذب) نشان می‌دهد (معادله ۱۲). این معیار توانایی مدل پیش‌بینی زلزله در شناسایی صحیح موارد مثبت را اندازه‌گیری می‌کند. همچنین Recall: نسبت

شمال استان زمین‌ساختی زاگرس (Ommi and Hashemi, 2024)، بر افزایش دقت روش درخت تصادفی برای استان زمین‌ساختی البرز-آذربایجان دلالت دارد. مطالعه مشابه در شمال زاگرس برای سطح بزرگای آستانه ۵.۵، مقدار دقت را برای روش شبکه عصبی مصنوعی به ترتیب  $Accuracy=96.27\%$ ،  $Precision=85\%$ ،  $Recall=97.5\%$  و  $F1-score=90\%$  گزارش کرده است. در مطالعه حاضر برای آستانه بزرگای مشابه براساس کاتالوگ لرزه‌ای استان زمین‌ساختی البرز-آذربایجان روش شبکه عصبی مصنوعی به سطح دقت  $Accuracy=97.93\%$ ،  $Precision=98.17\%$ ،  $Recall=97.31\%$  و  $F1-score=97.71\%$  دست یافته است و روش درخت تصادفی در بالاترین سطح دقت  $Accuracy=98.78\%$ ،  $Precision=97.26\%$ ،  $Recall=99.9\%$  و  $F1-score=97.71\%$  قرار گرفته است. گرچه میزان دقت روش شبکه عصبی نیز افزایش را نسبت به مطالعه قبلی نشان می‌دهد اما روش درخت تصادفی پیشی گرفته است با توجه به این مهم که در مطالعه استان زمین‌ساختی زاگرس شمالی بخشی از زاگرس مورد مطالعه قرار گرفته بود، می‌توان افزایش دقت روش شبکه عصبی در این مطالعه را با افزایش تعداد داده‌ها مرتبط دانست. از اینرو مقایسه نتایج این مطالعه با نتایج سایر استان‌های زمین‌ساختی در پژوهش‌های آتی پیشنهاد می‌شود تا میزان اثر فراوانی مجموعه داده و ویژگی‌های تکتونیکی بر نتایج آشکار شود.

زمین‌ساختی البرز-آذربایجان نسبت به زاگرس است (Mousavi-Bafrouei and Mahani, 2020). این بهبود در کیفیت و کمیت اطلاعات نتایج بهتری با دقت بالاتر در تکنیک‌های یادگیری ماشینی برای استان زمین‌ساختی البرز-آذربایجان به همراه داشته است. توجه به این مهم ضروری است که عملکرد هر یک از روش‌های یادگیری ماشینی به ماهیت داده‌ها بستگی دارد برخی روش‌ها به ویژه روش شبکه عصبی مصنوعی متکی بر تعداد داده‌ها هستند و با افزایش فراوانی مجموعه داده، مقدار دقت کاهش می‌یابد (Yousefzadeh, Hosseini and Farnaghi 2021). برخی مطالعات توانایی روش‌های بردار پشتیبان و درخت تصادفی را در تخمین رویدادهای نادر و کم تعداد برجسته نمودند (Galkina and Grafeeva, 2019). از سوی دیگر، تنوع داده، به معنای تنوع در کلاس طبقات، در دقت این روش‌ها موثر هستند با تکیه بر فراوانی داده‌ها و تنوع رخدادهای لرزه‌ای در تمام بزرگاها، که شکل ۲ نیز آن را تایید می‌کند، استان لرزه‌زمین‌ساختی البرز و آذربایجان دارای رخدادهایی با توان لرزه‌ای بالا است. در مقایسه با این استان زمین‌ساختی، زاگرس زلزله‌های متوسط را با فراوانی بالاتر تجربه می‌کند. کاتالوگ لرزه‌ای شواهد اندکی دال بر رخداد رویدادهای با بزرگای ۷ و بیشتر، در استان زمین‌ساختی زاگرس نشان می‌دهد (Mirzaei et al., 1998). مقایسه میزان دقت حاصل از تخمین‌های روش‌های یادگیری ماشینی در این مطالعه و مطالعه انجام شده در

جدول ۴- مقادیر دقت برآورد شده در سه روش یادگیری ماشینی: شبکه عصبی (Artificial Neural Network)، جنگل تصادفی (Random Forest)، و ماشین بردار پشتیبان (Support Vector Machine) برای کاتالوگ زلزله‌های البرز-آذربایجان.

Table 4. Estimated accuracy values using three machine learning methods: Artificial Neural Network, Random Forest, and Support Vector Machine for the Alborz-Azerbaijan earthquake catalog.

ML model	Support Vector Machine	Random Forest	Artificial Neural Network
Accuracy	97.50	98.76	97.93
Precision	96.39	97.26	98.17
Recall	98.18	99.9	97.31
F1-score	97.23	98.60	97.71

زمین‌ساختی زاگرس بر نتایج متفاوتی تأکید می‌کند. بطور کلی نتایج دقت بالا ۹۵٪ در چهار متریک دقت Accuracy، Precision، Recall و F1-score، در هر سه روش حاکی از موفقیت روش‌های یادگیری ماشینی در تخمین خطر لرزه‌ای این استان زمین‌ساختی است. روش جنگل تصادفی موفقیت بالایی را در تشخیص الگوی رفتار رژیم لرزه‌ای این استان زمین‌ساختی نشان داده است و کمترین آلام‌های نادرست را تولید نموده است که آن را می‌توان با تنوع بالای داده‌های لرزه‌ای در این منطقه و وقوع رخداد‌های نادر برای این استان زمین‌ساختی مرتبط دانست. این مطالعه نشان می‌دهد، اگرچه وقوع زلزله یک پدیده تصادفی است، اما مدل سازی پیچیده و رویکردهای یادگیری عمیق ماشینی می‌تواند برای توصیف رفتار پیشرو منطقه لرزه‌زا نویدبخش باشد. نتایج این مطالعه و مطالعه مشابه در زاگرس (Omami and Hashemi, 2024) بر اهمیت غنای کاتالوگ در نیکویی تخمین‌ها دلالت دارد. همچنین تجزیه و تحلیل‌های انجام شده حاکی از آن است که میزان این موفقیت همچنین مرهون شاخص‌های لرزه‌خیزی است که منعکس‌کننده تغییرات در پتانسیل لرزه‌خیزی منطقه هستند. افزایش کیفیت و کمیت داده‌های لرزه‌خیزی و افزودن سایر ویژگی‌های محیطی، مانند مکانیسم کانونی رخدادها، می‌تواند نقش مهمی در موفقیت آموزش داده‌ها در تکنیک‌های یادگیری ماشینی ایفا کند که تحقیق آن برای مطالعات آتی پیشنهاد می‌شود.

### سپاسگزاری

بدین وسیله از شبکه لرزه‌نگاری مؤسسه ژئوفیزیک دانشگاه تهران (Iranian Seismological Center) که اطلاعات لازم را برای این تحقیق تأمین کردند، صمیمانه سپاسگزاری می‌نمایم. همچنین از اساتید فرهیخته دانشکده کامپیوتر دانشگاه شهید بهشتی به خاطر راهنمایی‌ها و حمایت‌های ارزشمندشان در طول این مطالعه، تشکر می‌کنم.

انتخاب شاخص‌های لرزه‌خیزی توصیف شده در این مطالعه نیز در موفقیت تخمین‌ها از عوامل مهم موفقیت تخمین‌ها است. در این پژوهش پیش‌بینی بزرگا و زمان زلزله‌ها براساس مجموعه‌ای از پارامترهای ورودی صورت می‌گیرد. این پارامترهای استخراج شده از توزیع زمانی زمین لرزه‌های گذشته و فراوانی وقوع رویدادهای لرزه‌ای که به عنوان تابعی از بزرگای زلزله است، بهره گرفته است. این پارامترها روابط زیربنایی را نشان می‌دهد که حقایق از سکون لرزه‌ای (Hainzl et al., 2000)، قانون گوتنبرگ-ریشتر (Rundle and Donnellan, 2020) و فراوانی پیش لرزه‌ها (McGuire, Boettcher and Jordan 2005) را روشن و رفتار فعالیت‌های لرزه‌ای را مدل‌سازی می‌کند. همچنین ماهیت غیرخطی بین این روابط نیز آشکار می‌گردد. وقوع زمان زلزله آتی از رصد توالی رخداد‌های لرزه‌ای (فرآیند سکون و فعالیت لرزه‌ای)، ممکن است، برآورد گردد. همچنین از مقدار انرژی انباشته در ناحیه گسل که منجر به وقوع زلزله می‌شود تخمین بزرگی ممکن می‌شود. به طور مشابه، پیش لرزه به عنوان نشانه‌ای از یک زلزله بزرگ در نظر گرفته می‌شود. همچنین بکارگیری قانون گوتنبرگ-ریشتر که رابطه رگرسیون بین بزرگای زلزله و فراوانی آنها است به عنوان یک شاخص مهم لرزه‌خیزی به خوبی در توصیف رفتار لرزه‌زایی منطقه در این محاسبات مداخله کرده است.

### نتیجه‌گیری

این پژوهش با هدف مطالعه میزان موفقیت تکنیک یادگیری ماشینی در تخمین خطر لرزه‌ای و رصد تغییرات پتانسیل لرزه‌خیزی در منطقه البرز-آذربایجان انجام شده است این منطقه که از فعال‌ترین مناطق لرزه خیز فلات ایران است، برای ارزیابی سه تکنیک یادگیری ماشینی شامل شبکه عصبی مصنوعی، جنگل تصادفی، و ماشین بردار پشتیبان در نظر گرفته شد. مقایسه یافته‌های حاصل از این مطالعه با پژوهش مشابه در شمال استان

## References

- Abdullah, D. M. & A. M. Abdulazeez (2021) Machine learning applications based on SVM classification a review. *Qubahan Academic Journal*, 1, 81-90. <https://zenodo.org/records/5079050>
- Asim, K., F. Martínez-Álvarez, A. Basit & T. Iqbal (2017) Earthquake magnitude prediction in Hindukush region using machine learning techniques. *Natural Hazards*, 85, 471-486. <https://doi.org/10.1007/s11069-016-2483-2>
- Berberian, M. & R. Walker (2010) The Rudbār M w 7.3 earthquake of 1990 June 20; seismotectonics, coseismic and geomorphic displacements, and historic earthquakes of the western 'High-Alborz', Iran. *Geophysical Journal International*, 182, 1577-1602. <https://academic.oup.com/gji/article/182/2/613/568363>
- Bergen, K. J., P. A. Johnson, M. V. de Hoop & G. C. Beroza (2019) Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363, eaau0323. <https://www.science.org/doi/10.1126/science.aau0323>
- Brehm, D. J. & L. W. Braile (1998) Intermediate-term earthquake prediction using precursory events in the New Madrid seismic zone. *Bulletin of the Seismological Society of America*, 88, 564-580. <https://doi.org/10.1785/BSSA088003564>
- Enescu, B. & K. Ito (2001) Some premonitory phenomena of the 1995 Hyogo-Ken Nanbu (Kobe) earthquake: seismicity, b-value and fractal dimension. *Tectonophysics*, 338, 297-314. [https://doi.org/10.1016/S0040-1951\(01\)00194-5](https://doi.org/10.1016/S0040-1951(01)00194-5)
- Galkina, A. & N. Grafeeva. 2019. Machine learning methods for earthquake prediction: A survey. In *Proceedings of the fourth conference on software engineering and information management (sEIM-2019)*, Saint Petersburg, Russia, 25. <https://doi.org/10.1145/3341608.3341626>
- Gutenberg, B. & C. F. Richter (1944) Frequency of earthquakes in California. *Bulletin of the Seismological society of America*, 34, 185-188. <https://doi.org/10.1785/BSSA034002185>
- Hainzl, S., G. Zöller, J. Kurths & J. Zs. <https://doi.org/10.1785/BSSA034002185chau>
- (2000) Seismic quiescence as an indicator for large earthquakes in a system of self-organized criticality. *Geophysical research letters*, 27, 597-600. <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/1999GL010810>
- Hessami, K., D. Pantosti, H. Tabassi, E. Shabanian, M. R. Abbassi, K. Fegghi & S. Solaymani (2003) Paleearthquakes and slip rates of the North Tabriz Fault, NW Iran: preliminary results. *Annals of Geophysics*, 46. <https://www.annalsofgeophysics.eu/index.php/annals/article/view/3471>
- Hulbert, C., B. Rouet-Leduc, R. Jolivet & P. A. Johnson (2020) An exponential build-up in seismic energy suggests a months-long nucleation of slow slip in Cascadia. *Nature communications*, 11, 4139. <https://doi.org/10.1038/s41467-020-17949-1>
- Jackson, J. & D. McKenzie (1988) The relationship between plate motions and seismic moment tensors, and the rates of active deformation in the Mediterranean and Middle East. *Geophysical Journal International*, 93, 45-73. <https://doi.org/10.1111/j.1365-246X.1988.tb01712.x>
- Johnson, P. A., B. Rouet-Leduc, L. J. Pyrak-Nolte, G. C. Beroza, C. J. Marone, C. Hulbert, A. Howard, P. Singer, D. Gordeev & D. Karaflos (2021) Laboratory earthquake forecasting: A machine learning competition. *Proceedings of the national academy of sciences*, 118, e2011362118. <https://www.pnas.org/doi/abs/10.1073/pnas.2011362118>
- Kagan, Y. (1982) Stochastic model of earthquake fault geometry. *Geophysical Journal International*, 71, 659-691. <https://doi.org/10.1111/j.1365-246X.1982.tb02791.x>
- Kong, Q., D. T. Trugman, Z. E. Ross, M. J. Bianco, B. J. Meade & P. Gerstoft (2019) Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, 90, 3-14. <https://doi.org/10.1785/0220180259>
- Lawson, A. C. & H. F. Reid. 1910. The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission. Carnegie institution of Washington.
- Lomax, A., A. Michelini & D. Jozinović (2019) An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network. *Seismological Research Letters*, 90, 517-529. <https://doi.org/10.1785/0220180311>
- Masson, F., J. Chéry, D. Hatzfeld, J. Martinod, P. Vernant, F. Tavakoli & M. Ghafory-Ashtiani (2005) Seismic versus aseismic deformation in Iran inferred from earthquakes and geodetic data. *Geophysical Journal International*, 160, 217-226. <https://doi.org/10.1111/j.1365-246X.2004.02465.x>
- McGuire, J. J., M. S. Boettcher & T. H. Jordan (2005) Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults. *Nature*, 434, 457-461. <https://doi.org/10.1038/nature03377>
- Mignan, A. & M. Broccardo (2020) Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations. *Seismological Research Letters*, 91, 2330-2342. <https://doi.org/10.1785/0220200021>

- Mirzaei, N., G. Mengtan & C. Yuntai (1998) Seismic source regionalization for seismic zoning of Iran: major seismotectonic provinces. *Journal of earthquake prediction research*, 7, 465-495.
- Mousavi-Bafrouei, S. H. & A. B. Mahani (2020) A comprehensive earthquake catalogue for the Iranian Plateau (400 BC to December 31, 2018). *Journal of Seismology*, 24, 709-724. <https://link.springer.com/article/10.1007/s10950-020-09923-6>
- Nanjo, K. (2020) Were changes in stress state responsible for the 2019 Ridgecrest, California, earthquakes? *Nature communications*, 11, 3082. <https://www.nature.com/articles/s41467-020-16867-5>
- Ommi, S. & M. Hashemi (2024) Machine learning technique in the north zagros earthquake prediction. *Applied Computing and Geosciences*, 22, 100163. <https://doi.org/10.1016/j.acags.2024.100163>
- Ommi, S. & V. B. Smirnov (2024) Seismicity patterns before the 2021 Fin (Iran) doublet earthquakes using the region-time-length and time-to-failure methods. *Earthquake Science*, 37, 324-336.
- Panakkat, A. & H. Adeli (2007) Neural network models for earthquake magnitude prediction using multiple seismicity indicators. *International journal of neural systems*, 17, 13-33.
- Pasari, S. (2019) Nowcasting earthquakes in the Bay of Bengal region. *Pure and Applied Geophysics*, 176, 1417-1432.
- Pasari, S. & A. Mehta (2018) Nowcasting earthquakes in the northwest Himalaya and surrounding regions. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 855-859.
- Pisner, D. A. & D. M. Schnyer. 2020. Support vector machine. In *Machine learning*, 101-121. Elsevier.
- Ren, C. X., C. Hulbert, P. A. Johnson & B. Rouet-Leduc (2020a) Machine learning and fault rupture: a review. *Advances in Geophysics*, 61, 57-107.
- Ren, C. X., A. Peltier, V. Ferrazzini, B. Rouet-Leduc, P. A. Johnson & F. Brenguier (2020b) Machine learning reveals the seismic signature of eruptive behavior at Piton de la Fournaise volcano. *Geophysical Research Letters*, 47, e2019GL085523.
- Rouet-Leduc, B., C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys & P. A. Johnson (2017) Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, 44, 9276-9282.
- Rundle, J. B. & A. Donnellan (2020) Nowcasting earthquakes in Southern California with machine learning: Bursts, swarms, and aftershocks may be related to levels of regional tectonic stress. *Earth and Space Science*, 7, e2020EA001097.
- Rundle, J. B., A. Donnellan, G. Fox & J. P. Crutchfield (2022) Nowcasting earthquakes by visualizing the earthquake cycle with machine learning: A comparison of two methods. *Surveys in Geophysics*, 43, 483-501.
- Rundle, J. B., A. Donnellan, G. Fox, J. P. Crutchfield & R. Granat (2021) Nowcasting earthquakes: Imaging the earthquake cycle in California with machine learning. *Earth and Space Science*, 8, e2021EA001757.
- Rundle, J. B., A. Giguere, D. L. Turcotte, J. P. Crutchfield & A. Donnellan (2019) Global seismic nowcasting with Shannon information entropy. *Earth and Space Science*, 6, 191-197.
- Wang, R., Y. Chang, M. Miao, Z. Zeng, H. Chen, H. Shi, D. Li, L. Liu, Y. Su & P. Han (2021) Assessing earthquake forecast performance based on b value in Yunnan Province, China. *Entropy*, 23, 730.
- Wiemer, S. & M. Wyss. 2002. Mapping spatial variability of the frequency-magnitude distribution of earthquakes. In *Advances in geophysics*, 259-V. Elsevier.
- Woessner, J. & S. Wiemer (2005) Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, 95, 684-698.
- Yousefzadeh, M., S. A. Hosseini & M. Farnaghi (2021) Spatiotemporally explicit earthquake prediction using deep neural network. *Soil Dynamics and Earthquake Engineering*, 144, 106663.
- Zhao, T., F. Li & K. J. Marfurt (2017) Constraining self-organizing map facies analysis with stratigraphy: An approach to increase the credibility in automatic seismic facies classification. *Interpretation*, 5, T163-T171.

