



Earthquake prediction using a hybrid deep learning model

Javad Ezzati¹, Seyed Hassan Mortazavi² , Fatemeh Zare MehrJardi³ , Mohsen Sardari Zarchi⁴

1. MSc. Student, Department of Computer Engineering, University of Meybod, Meybod, Iran .Email: javadezzati@iran.ir
2. Corresponding Author ,Assistant Prof, Department of Computer Engineering, University of Meybod, Meybod, Iran .Email: hassanmortazavi@gmail.com
3. Assistant Prof, Department of Computer Engineering, University of Meybod, Meybod, Iran .Email: fzare@meybod.ac.ir
4. Associate Prof, Department of Computer Engineering, University of Meybod, Meybod, Iran. Email: sardari@meybod.ac.ir

Article Info	ABSTRACT
Article type: Research Article Article history: Received 30 June 2025 Received in revised form 3 September 2025 Accepted 21 Octobr 2025 Published online 1 December 2025 Keywords: CNN, Deep Learning, Earthquake Prediction, LSTM, Transformer Model.	Earthquake prediction is one of the most challenging issues in geosciences, especially in seismically active regions such as the northern Zagros. In this study, advanced deep learning methods were employed to predict earthquakes with a magnitude greater than 5.5. The dataset used includes over 23,000 recorded seismic events from 1990 to 2024 in the Sarpol-e Zahab area and its surroundings. After data cleaning and preprocessing, the data were divided into three subsets: training (2012 to 2015), validation (2016 to 2018), and testing (from 2019 onward). Subsequently, several deep learning models were designed and implemented, including convolutional neural networks (CNNs), long short-term memory networks (LSTMs), transformer models, and a hybrid architecture combining these three approaches. The hybrid model, by simultaneously leveraging spatial feature extraction, temporal dependency recognition, and adaptive attention mechanisms, demonstrated superior performance compared to the other models. In the final evaluation, this model achieved an accuracy of 99.51%, a precision of 96.59%, a recall of 93.92%, and an overall F1-score of 95.24%, confirming its high efficacy in predicting potential earthquakes within a 30-day timeframe. Despite existing limitations in precisely forecasting the exact time, location, and intensity of earthquakes, the results of this study indicate that the strategic use of hybrid deep learning models can play a crucial role in the development of intelligent earthquake early warning systems and in mitigating human and financial losses.
Cite this article: Ezzati, j. & et al, (2025)., Earthquake prediction using a hybrid deep learning model. <i>Enginiiring Management and Soft Computing</i>, 11 (2). 257-281. DOI: https://doi.org/10.22091/jemsc.2025.13249.1282	
 © Ezzati et. al (2025) DOI: https://doi.org/10.22091/jemsc.2025.13249.1282 Publisher: University of Qom	

1) INTRODUCTION

The accurate prediction of strong earthquakes is of critical importance due to their destructive impacts on human life and infrastructure. Traditional seismological methods rely on geophysical indicators but face challenges in capturing complex spatio-temporal dependencies. With the advent of Artificial Intelligence (AI), data-driven methods have gained significant attention. Deep learning models, particularly CNNs, LSTMs, and Transformers, have shown promise in time-series analysis. However, each model has inherent limitations when applied individually. This research addresses these gaps by designing a hybrid architecture that leverages CNNs for spatial feature extraction, LSTMs for sequential temporal dependencies, and Transformers for adaptive long-term relations.

2) Solution Method

2-1) Equations

The proposed problem is formulated as a binary classification task:

$$Y = \begin{cases} 1 & \text{if an earthquake with } M \geq 5.5 \text{ occurs within 30 days,} \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

3) conclusion

This study introduces a hybrid deep learning model for earthquake prediction in the northern Zagros, integrating CNN, LSTM, and Transformer components. The experimental results demonstrate that the model achieves near-perfect accuracy and strong generalization, outperforming traditional approaches. Although the exact time and location of earthquakes cannot be predicted with certainty, the proposed method provides a reliable framework for intelligent early warning systems. Future work should focus on incorporating geophysical indicators (e.g., GPS, InSAR, stress models) to further improve prediction reliability.

Tables and Figures

Three baseline models (CNN, LSTM, Transformer) and the hybrid architecture were evaluated. As shown in Table 1, the hybrid model outperformed all baselines in terms of accuracy, recall, and F1-score.

Table 1. Performance comparison of models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	96.49	91.34	89	94.05
LSTM	94.83	89.12	92.87	90.95
Transformer	94.86	90.08	92.41	91.22
Hybrid	99.51	96.59	93.92	95.24

The hybrid model not only reduced false alarms but also provided robustness against noisy data. The Area Under the ROC Curve (AUC = 0.9993) confirmed its outstanding classification ability. Compared to previous studies on the Zagros dataset using classical machine learning models (e.g., ANN, SVM, RF), our hybrid model achieved superior predictive performance.

4) ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Iranian Seismological Center (IRSC) for providing seismic event catalogs.

REFERENCES

M. Bhatia, T. A. Ahanger, and A. Manocha, "Artificial intelligence based real-time earthquake prediction," Engineering Applications of Artificial Intelligence, vol. 120, p. 105856, 2023. DOI: 10.1016/j.engappai.2023.105856

- A. Bhatia, S. Pasari, and A. Mehta, "Earthquake forecasting using artificial neural networks," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, no. 5, pp. 823–827, 2018. <https://doi.org/10.5194/isprs-archives-XLII-5-823-2018>
- Z. Bao, J. Zhao, P. Huang, S. Yong, and X. Wang, "A deep learning-based electromagnetic signal for earthquake magnitude prediction," *Sensors*, vol. 21, no. 13, p. 4434, Jun. 2021. [doi: 10.3390/s21134434](https://doi.org/10.3390/s21134434).
- M. T. Green and S. K. Patel, "Machine learning approaches for earthquake risk assessment: A comprehensive survey," *IEEE Reviews in Geosciences*, vol. 29, pp. 102–118, 2022. doi.org/10.1016/j.pdisas.2024.100398
- N. B. Jarah, A. H. H. Alasadi, and K. M. Hashim, "Earthquake prediction technique: a comparative study," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 3, pp. 1026–1032, 2023. [DOI:10.11591/ijai.v12.i3.pp1026-1032](https://doi.org/10.11591/ijai.v12.i3.pp1026-1032)
- S. Ommi and M. Hashemi, "Machine learning technique in the north Zagros earthquake prediction," *Applied Computing and Geosciences*, vol. 22, p. 100163, 2024. [DOI:10.1016/j.acags.2024.100163](https://doi.org/10.1016/j.acags.2024.100163)
- A. Gupta and R. Kumar, "A REAL TIME PERFORMANCE BASED MODEL FOR PREDICTING THE EARTHQUAKE EARLY WARNING ALARMS USING A MACHINE LEARNING TECHNIQUE," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4567–4578, June 2021. [DOI: http://doi.org/10.26480/mjg.02.2024.115.120](https://doi.org/10.26480/mjg.02.2024.115.120)
- M. H. Al Banna, K. A. Taher, M. S. Kaiser, M. Mahmud, M. S. Rahman, A. S. Hosen, and G. H. Cho, "Application of artificial intelligence in predicting earthquakes: State-of-the-art and future challenges," *IEEE Access*, vol. 8, pp. 192880–192923, 2020. [DOI: 10.1109/ACCESS.2020.3029859](https://doi.org/10.1109/ACCESS.2020.3029859)
- M. Apriani, S. K. Wijaya, and Daryono, "Earthquake magnitude estimation based on machine learning: Application to earthquake early warning system," *Journal of Physics: Conference Series*, vol. 1951, no. 1, p. 012057, 2021. [DOI 10.1088/1742-6596/1951/1/012057](https://doi.org/10.1088/1742-6596/1951/1/012057)
- J. P. Huang, X. A. Wang, Y. Zhao, C. Xin, and H. Xiang, "Large earthquake magnitude prediction in Taiwan based on deep learning neural network," *Neural Network World*, vol. 28, no. 2, pp. 149–160, 2018, [DOI: 10.14311/NNW.2018.28.009](https://doi.org/10.14311/NNW.2018.28.009)
- Keogh, E., Chu, S., Hart, D., & Pazzani, M. (2001). An online algorithm for segmenting time series. *Proceedings IEEE International Conference on Data Mining*, pp. 289–296. <https://doi.org/10.1109/ICDM.2001.989531>
- T. Perol, M. Gharbi, and M. A. Denolle, "Convolutional neural network for earthquake detection and location," *Science Advances*, vol. 4, no. 2, p. e1700578, Feb. 2018, [DOI: 10.1126/sciadv.1700578](https://doi.org/10.1126/sciadv.1700578)

پیش‌بینی زلزله با استفاده از مدل ترکیبی یادگیری عمیق

جواد عزتی^۱، سیدحسن مرتضوی^۲ ^۳، فاطمه زارع مهر جردی^۳ ^۴ و محسن سرداری زارچی^۴ 

۱. دانشجوی کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه میبد، میبد، ایران. رایانامه: javadezzati@iran.ir

۲. نویسنده مسئول، استادیار، دانشکده مهندسی کامپیوتر، دانشگاه میبد، میبد، ایران. رایانامه: hassanmortazaviy@gmail.com

۳. استادیار، دانشکده مهندسی کامپیوتر، دانشگاه میبد، میبد، ایران. رایانامه: fzare@meybod.ac.ir

۴. دانشیار، دانشکده مهندسی کامپیوتر، دانشگاه میبد، میبد، ایران. رایانامه: sardari@meybod.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی تاریخ دریافت: ۱۴۰۳/۰۵/۲۹ تاریخ بازنگری: ۱۴۰۳/۱۲/۱۶ تاریخ پذیرش: ۱۴۰۴/۰۵/۱۳ تاریخ انتشار: ۱۴۰۴/۰۹/۱۰	یکی از چالش‌برانگیزترین مسائل در علوم زمین به‌ویژه در مناطق فعال لرزه‌ای مانند زاگرس شمالی، پیش‌بینی زلزله است. در این پژوهش، از روش‌های پیشرفته یادگیری عمیق برای پیش‌بینی زلزله‌هایی با بزرگی بیش از ۵.۵ ریشتر استفاده شده است. داده‌ها شامل بیش از ۲۳ هزار رویداد لرزه‌ای ثبت شده، در بازه زمانی سال ۱۹۹۰ تا ۲۰۲۴ در منطقه سرپل ذهاب و نواحی اطراف آن است. پس از پاک‌سازی و آماده‌سازی، داده‌ها به سه بخش آموزش (۲۰۱۲ تا ۲۰۱۵)، اعتبارسنجی (۲۰۱۶ تا ۲۰۱۸) و آزمون (از ۲۰۱۹ به بعد) تقسیم شدند. در ادامه، چندین مدل یادگیری عمیق طراحی و پیاده‌سازی شد که شامل شبکه‌های عصبی پیچشی، حافظه بلندمدت، مدل‌های مبدل و همچنین یک معماری ترکیبی از این سه مدل بود. مدل ترکیبی با بهره‌گیری همزمان از قابلیت‌های استخراج ویژگی‌های مکانی، شناسایی وابستگی‌های زمانی و تمرکز تطبیقی، عملکرد بهتری نسبت به سایر مدل‌ها نشان داد. در ارزیابی نهایی، این مدل دقت ۹۹.۵۱ درصد، شاخص صحت ۹۶.۵۹ درصد، بازیابی ۹۳.۹۲ درصد و امتیاز نهایی ۹۵.۲۴ درصد را ثبت کرد که کارایی بالایی آن در پیش‌بینی زلزله‌های احتمالی در بازه ۳۰ روزه را تایید می‌کند. با وجود محدودیت‌های موجود در پیش‌بینی دقیق زمان، مکان و شدت زلزله‌ها، نتایج این پژوهش نشان می‌دهد که استفاده راهبردی از مدل‌های ترکیبی یادگیری عمیق می‌تواند نقش مهمی در توسعه سامانه‌های هوشمند هشدار زلزله و کاهش خسارات انسانی و مالی ایفا کند.

استناد: عزتی، جواد و همکاران. (۱۴۰۴). «پیش‌بینی زلزله با استفاده از مدل ترکیبی یادگیری عمیق». مدیریت مهندسی و رایانش نرم، دوره ۱۱ (۲). صص:

۲۵۷-۲۸۱. <https://doi.org/10.22091/jemsc.2025.13249.1282>



(۱) مقدمه

پیش‌بینی وقوع زلزله یکی از چالش‌های مهم و پیچیده در زمینه علوم زمین است. وقوع زلزله‌های بزرگ می‌تواند خسارات جانی و مالی زیادی به همراه داشته باشد لذا توانایی پیش‌بینی دقیق این رویدادها، می‌تواند به کاهش اثرات منفی آنها کمک کند. با این حال، پیش‌بینی دقیق زمان، مکان و بزرگی زلزله‌ها همچنان یکی از مسائل چالش‌برانگیز است. درحالی‌که مدل‌های سنتی بر پایه تجزیه و تحلیل داده‌های ژئوفیزیکی و تجربی در تلاش هستند تا روندهای وقوع زلزله را شبیه‌سازی کنند، استفاده از مدل‌های جدیدتر مبتنی بر یادگیری ماشین به‌ویژه یادگیری عمیق، توانسته است توجه بیشتری جلب کند. این مدل‌ها قادرند الگوهای پیچیده و ناشناخته در داده‌های سری زمانی لرزه‌ای را شناسایی کنند. در اینجا چند سؤال اساسی مطرح می‌شود. ویژگی‌های جغرافیایی و محیطی بطور مداوم در حال تغییر هستند، بنابراین چرا یک روش یادگیری می‌تواند قابلیت پیش‌بینی زلزله‌های بالای ۵.۵ ریشتر را در یک منطقه خاص داشته باشد؟ آن هم بر روی داده‌های زمانی؟ این مسئله این سؤال را به دنبال دارد که آیا داده‌های زمانی منقضی می‌شوند؟ بطور کلی ماهیت پیش‌بینی در اینجا چیست؟ آیا منظور تخمین کلاس نمونه‌ای از یک زلزله رخ داده است؟ مدل‌های یادگیری عمیق که قادرند وقوع زلزله‌های بالای ۵.۵ ریشتر را پیش‌بینی کنند، براساس کدام ویژگی‌ها این قابلیت را به دست آورده‌اند؟ در پاسخ به این سؤالات باید گفت که داده‌های زمانی به مدل‌های یادگیری کمک می‌کنند تا الگوهای مشابه گذشته را شبیه‌سازی کنند و براساس آنها، پیش‌بینی‌هایی برای آینده ارائه دهند. این پیش‌بینی‌ها به‌ویژه زمانی موثر است که مدل بتواند ویژگی‌های ثابت و ماندگار از داده‌های گذشته را شناسایی کند. یکی دیگر از سؤالات مطرح شده در این تحقیق این است که آیا داده‌های استفاده شده در مدل ترکیبی بکارگرفته شده در پژوهش فعلی، قادر به پیش‌بینی وقوع زلزله قبل از رخداد آن هستند؟ پاسخ به این سؤال قطعاً منفی است. چراکه داده‌های لرزه‌ای استفاده شده در این تحقیق تنها شامل ویژگی‌هایی از زلزله‌های گذشته است و هیچ پارامتری از وضعیت‌های زمین‌شناسی (مانند تنش‌ها و گسل‌ها) که بطور مستقیم به وقوع زلزله‌ها مرتبط باشد، در اختیار ندارند. از آنجایی‌که داده‌های زمان‌محور می‌توانند موثر واقع شوند، زمانیکه مدل بطور آنلاین و در زمان واقعی اجرا شود، این نکته در مقدمه باید بطور کامل‌تر و دقیق‌تر بیان شود. استفاده از داده‌های زمانی در مدل‌ها به‌ویژه زمانی مفید است که مدل بطور برخط و در زمان واقعی اجرا شده و داده‌ها به‌صورت مداوم به‌روز شوند. در پژوهش‌های مشابه، این ویژگی می‌تواند به پیش‌بینی‌های دقیق‌تر و سریع‌تر منجر شود. درمجموع پیش‌بینی زلزله همچنان یکی از مهمترین چالش‌ها در علوم زمین محسوب می‌شود. اگرچه روش‌های سنتی با محدودیت‌هایی همراه هستند، مدل‌های نوین مبتنی بر یادگیری ماشینی و تحلیل داده‌های لرزه‌ای، پیشرفت‌های چشمگیری در این زمینه ایجاد کرده‌اند. توسعه مدل‌های دقیق‌تر و بهبود سیستم‌های هشداردهی زلزله می‌تواند به کاهش خسارات جانی و مالی ناشی از این پدیده طبیعی کمک کند.

(۲) پیشینه تحقیق

پیش‌بینی زلزله در منطقه شمال زاگرس با استفاده از تکنیک‌های یادگیری ماشین توسط هاشمی و همکاران مورد بررسی قرار گرفته است. در این تحقیق نشان داده شد که مدل‌های یادگیری ماشین می‌توانند با دقت بالایی زمان و مکان وقوع زلزله

را پیش‌بینی کنند. در این پژوهش، استفاده از داده‌های لرزه‌ای و تکنیک‌های پیش‌بینی زمان‌محور برای تحلیل رفتار زلزله‌ها مورد توجه قرار گرفته‌است.

در سال‌های اخیر، پیش‌بینی زلزله‌ها با استفاده از مدل‌های یادگیری ماشین و یادگیری عمیق به یکی از موضوعات پُرطرفدار در تحقیقات علمی تبدیل شده‌است. این تحقیق به بررسی مدل‌ها و تکنیک‌هایی که برای پیش‌بینی وقوع زلزله‌ها و ویژگی‌های مختلف آنها از جمله بزرگی و زمان وقوع استفاده شده‌اند، پرداخته‌است. بسیاری از تحقیقات از داده‌های زمان‌محور استفاده کرده‌اند و براساس آنها، مدل‌های مختلفی طراحی شده‌اند که توانایی پیش‌بینی دقیق‌تر وقوع زلزله‌ها را دارند. در این راستا، دو روش مختلف برای پیش‌بینی وقوع زلزله توسط Gupta و همکاران (۲۰۲۱) معرفی شده‌است. یکی از این روش‌ها، بر ارزیابی دقت شبکه عصبی چندلایه با ورودی‌های مختلف متمرکز است. در این تحقیق از تکنیک "نوک ستینگ" برای پیش‌بینی وقوع زلزله استفاده شده‌است که میزان پیشرفت در دوره وقوع زلزله‌های بزرگتر را با توجه به تعداد زلزله‌های کوچکتر در همان منطقه مشخص می‌کند. در این تحقیق یک مدل یادگیری ماشین برای سیستم‌های هشدار سریع زلزله (EEW) ارائه کردند که بر شناسایی امواج اولیه (P-waves) و تخمین بزرگی زلزله تمرکز دارد. مدل پیشنهادی آنها که از یک چارچوب اولویت‌بندی شده برای ترکیب الگوریتم‌های مختلف استفاده می‌کند، با هدف افزایش قابلیت اطمینان و کاهش زمان هشدار طراحی شده‌است. این تحقیق نشان داد که استفاده از تکنیک "نوک ستینگ" که در آن وقوع زلزله‌های بزرگ براساس تعداد زلزله‌های کوچکتر پیش‌بینی می‌شود، در کنار ساختار شبکه بازگشتی حافظه طولانی کوتاه‌مدت (LSTM)، می‌تواند به بهبود عملکرد سیستم کمک کند. نتایج شبیه‌سازی‌ها بر روی داده‌های واقعی زلزله نشان داد که مدل ترکیبی آنها در مقایسه با سیستم‌های تک‌الگوریتمی، عملکرد بهتری در تشخیص و هشدار سریع دارد. Banna و همکاران (۲۰۲۰) در یک تحقیق مروری بطور جامع به بررسی ۸۴ مقاله علمی در زمینه پیش‌بینی زلزله با استفاده از تکنیک‌های هوش مصنوعی پرداخته‌اند. این مقاله به تحلیل انواع الگوریتم‌های یادگیری ماشین و روش‌های مبتنی بر قوانین برای پیش‌بینی زلزله می‌پردازد و عملکرد آنها را از نظر معیارهای ارزیابی مختلف مقایسه می‌کند. آنها دریافتند که اکثر مدل‌ها بر پیش‌بینی زمان، مکان و بزرگی زلزله تمرکز دارند اما چالش‌هایی مانند عدم تعادل داده‌ها (تعداد کم زلزله‌های بزرگ در مقابل زلزله‌های کوچک) و نیاز به بهبود مدل‌ها برای پیش‌بینی دقیق‌تر همچنان پابرجاست. این مقاله مروری تاکید کرد که علیرغم پیشرفت‌ها، هیچ پیش‌نشانگر قطعی برای زلزله یافت نشده‌است و بسیاری از مدل‌های بررسی شده بطور خاص بر تحلیل داده‌های سری زمانی تمرکز نکرده‌اند. این نکته اهمیت پژوهش‌های جاری در این حوزه را برجسته می‌سازد. در این تحقیق، چالش‌ها و مسیرهای تحقیقات آینده نیز برجسته شده‌است. در این مقاله مروری، هیچ‌یک از مدل‌های بررسی شده بطور خاص بر داده‌های زمان‌محور متمرکز نبوده‌اند که این نکته در مقایسه با پژوهش‌های جاری و استفاده از داده‌های زمان‌محور برای پیش‌بینی زلزله‌ها اهمیت دارد.

تحقیق Aprianin و همکاران (۲۰۲۱) به بررسی استفاده از یادگیری ماشین برای پیش‌بینی بزرگی زلزله‌ها با استفاده از داده‌های سیموگرافی پرداخته‌است. در این تحقیق، استفاده از شبکه‌های عصبی عمیق^۱ برای پیش‌بینی بزرگی زلزله‌ها

نشان داده‌است که این مدل‌ها می‌توانند عملکرد بهتری نسبت به الگوریتم‌های یادگیری ماشین مانند جنگل تصادفی^۱ ارائه دهند. نتایج نشان داد که مدل DNN طبقه‌بندی‌کننده (Classification DNN) با رسیدن به میزان خطای نهایی ۰.۶۳، عملکرد بهتری نسبت به مدل جنگل تصادفی از خود نشان داد. این پژوهش نتیجه‌گیری کرد که برای تخمین بزرگی زلزله از روی داده‌های لرزه‌نگاری خام، مدل‌های DNN به‌ویژه با مجموعه داده‌های بزرگتر، گزینه مناسب‌تری هستند و مدل رگرسیون DNN نیز جایگزین مناسبی برای تخمین بزرگی زلزله محسوب می‌شود.

در پژوهش حاضر نیز هدف پیش‌بینی زلزله‌های مناطق شمال زاگرس با استفاده از انواع مدل‌های یادگیری عمیق به صورت ساده و ترکیبی است. این تحقیق بطور خاص از داده‌های زمان‌محور برای پیش‌بینی دقیق‌تر وقوع زلزله‌ها بهره برده و تلاش دارد تا روش‌های پیش‌بینی فعلی را با استفاده از تکنیک‌های پیشرفته‌تر بهبود بخشد.

(۳) روش تحقیق

در این بخش، توضیح مجموعه داده، فرآیند آماده‌سازی داده‌ها و طراحی مدل‌های یادگیری عمیق برای پیش‌بینی وقوع زلزله‌های شدید در بازه‌های زمانی مشخص شرح داده می‌شود. رویکرد اصلی مورد استفاده در این پژوهش، مبتنی بر مفهوم پنجره متحرک^۲ است که داده‌های خام سری زمانی را به شکلی ساختارمند و قابل استفاده برای آموزش مدل‌های یادگیری عمیق تبدیل می‌کند. هدف اصلی این پژوهش، پیش‌بینی وقوع زلزله‌هایی با بزرگی بیشتر از ۵.۵ ریشتر در آینده است. در این پژوهش، آستانه ۵.۵ ریشتر برای تفکیک زلزله‌ها به دو کلاس "رویداد بزرگ" و "رویداد کوچک" در نظر گرفته شده‌است. این انتخاب براساس ملاحظات لرزه‌خیزی و آماری منطقه زاگرس انجام شده‌است. زلزله‌هایی با بزرگی ۵.۵ ریشتر و بالاتر به دلیل پتانسیل تخریب بالا و تاثیرات اجتماعی و اقتصادی قابل توجه، به عنوان رویدادهای بحرانی در نظر گرفته می‌شوند. علاوه بر این، انتخاب این آستانه به تعادل نسبی در تعداد نمونه‌ها میان دو کلاس کمک کرده و از چالش‌های احتمالی ناشی از عدم تعادل شدید در داده‌ها جلوگیری می‌کند که این امر برای آموزش بهینه مدل‌های یادگیری عمیق ضروری است. این مسئله به صورت یک مسئله طبقه‌بندی دودویی^۳ فرموله شده‌است؛ به این معنا که مدل باید تعیین کند آیا در یک بازه زمانی مشخص (مثلاً ۳۰ روز آینده) زلزله‌ای شدید (کلاس ۱) رخ خواهد داد یا خیر (کلاس ۰).

۳-۱) مجموعه داده

در این پژوهش از یک پایگاه داده محلی شامل اطلاعات دقیق زمین لرزه‌های ثبت شده در منطقه سرپل ذهاب و نواحی اطراف آن استفاده شده‌است. این پایگاه داده حاوی ۲۳,۳۴۲ رخداد لرزه‌ای از تاریخ ۱۰ ژانویه ۱۹۹۰ تا آخرین رکورد موجود در پایگاه است. داده‌های لرزه‌ای مورد استفاده در این پژوهش از کاتالوگ زلزله مرکز لرزه‌نگاری ایران (IRSC) برای منطقه زاگرس استخراج شده‌اند. این پایگاه داده یکی از معتبرترین منابع داده‌های لرزه‌نگاری در کشور است که مشخصات دقیق هر رویداد زلزله را ثبت می‌کند. منظور از پایگاه داده محلی، همان مجموعه داده‌ای است که از این منبع استخراج و پس از عملیات پیش‌پردازش در قالب یک فایل CSV آماده‌سازی شده‌است و شامل ویژگی‌های زمانی (تاریخ،

1 RF

2 Sliding Window

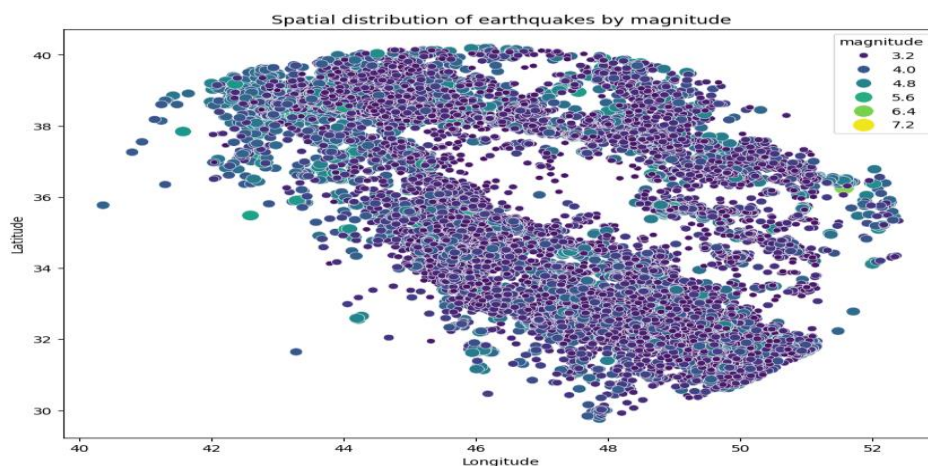
3 Binary Classification

ساعت، دقیقه، ثانیه)، موقعیت جغرافیایی (عرض و طول جغرافیایی)، عمق کانونی و بزرگی زلزله هستند. ویژگی‌های موجود در دیتاست بطور جامع در جدول ۱ آمده‌است.

جدول ۱. ویژگی‌های مجموعه داده زاگرس

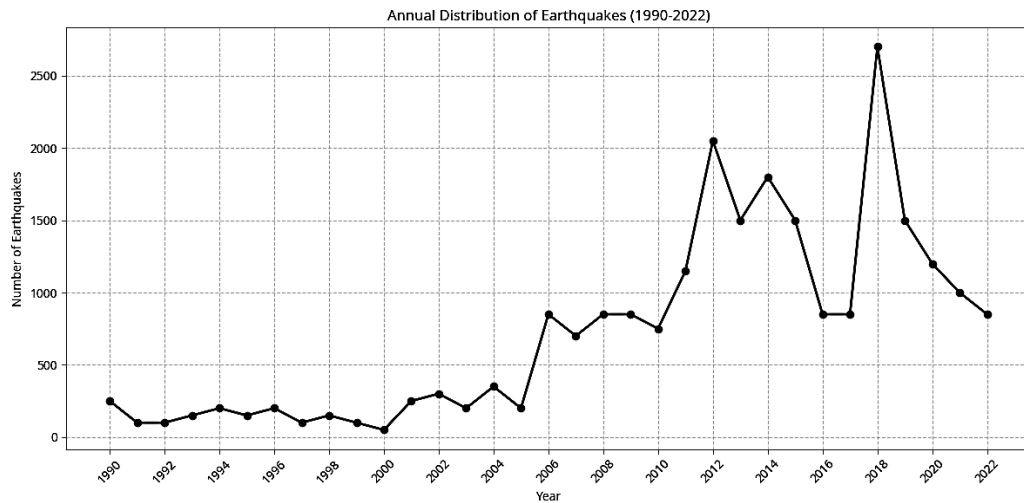
ویژگی	توضیح
زمان وقوع زلزله	شامل سال، ماه، روز، ساعت، دقیقه و ثانیه وقوع
موقعیت جغرافیایی	عرض و طول جغرافیایی وقوع زلزله در منطقه
بزرگی زلزله	بزرگی زلزله برحسب مقیاس ریشتر
عمق زلزله	عمق زلزله برحسب کیلومتر

برای بررسی دقیق‌تر ساختار آماری داده‌ها، توزیع مکانی و زمانی رخدادهای زلزله در منطقه مورد بررسی، تحلیل شده‌است. شکل ۱ توزیع مکانی زمین‌لرزه‌ها را در منطقه سرپل ذهاب به صورت نقشه حرارتی (Heatmap) نمایش می‌دهد. همانطور که در این شکل مشاهده می‌شود، بیشترین تمرکز رخدادها در ناحیه زاگرس چین خورده و بخش مرزی غرب ایران با عراق قرار دارد که با ویژگی‌های تکتونیکی منطقه همخوانی دارد.



شکل ۱. توزیع مکانی مجموعه داده

در شکل ۲، توزیع زمانی زلزله‌ها در بازه سال‌های ۱۹۹۰ تا پایان مجموعه داده نشان داده شده‌است. این تحلیل زمانی، به بررسی دوره‌های افزایش یا کاهش فعالیت لرزه‌ای و نیز روندهای بلندمدت در داده کمک می‌کند. تحلیل این روندها برای مدل‌سازی پدیده‌های زمانی غیرخطی مانند زلزله بسیار مهم است.



شکل ۲. نمودار توزیع زمانی دیتاست

۲-۳ آماده‌سازی داده‌ها

یکی از مراحل حیاتی در طراحی مدل‌های یادگیری عمیق، آماده‌سازی دقیق و هدفمند داده‌هاست. در این پژوهش، هدف پیش‌بینی احتمال وقوع زلزله‌ای با بزرگی برابر یا بیشتر از ۵.۵ ریشتر در بازه ۳۰ روز آینده پس از یک دنباله از زلزله‌های قبلی است. به همین منظور، از الگوریتم پنجره متحرک برای استخراج نمونه‌های آموزشی استفاده شد که سناریوهای واقعی از گذشته را به شکل ساختارمند برای مدل بازسازی می‌کند.

الف. تعریف ساختار ورودی‌ها (X)

داده‌های خام شامل توالی زمانی از رویدادهای لرزه‌ای ثبت‌شده در منطقه مورد مطالعه هستند. برای تبدیل این داده‌ها به نمونه‌های قابل استفاده در مدل، ساختاری به شرح زیر تعریف شد:

- هر نمونه ورودی متشکل است از دنباله‌ای شامل ۱۰۰ زلزله متوالی که دقیقاً پیش از لحظه پیش‌بینی، رخ داده‌اند.
 - ویژگی‌های هر زلزله شامل طول جغرافیایی، عرض جغرافیایی، عمق و بزرگی رویداد است.
- این ویژگی‌ها به صورت برداری در یک آرایه زمانی سازماندهی شده‌اند تا مدل بتواند وابستگی‌های زمانی و مکانی را یاد بگیرد.

ب. تعریف برجسب خروجی (Y)

برای هر نمونه ورودی، هدف پیش‌بینی وقوع یا عدم وقوع یک زلزله شدید (≥ 5.5 ریشتر) در ۳۰ روز آینده است. روند برجسب‌گذاری به صورت زیر انجام شد:

- ابتدا تاریخ آخرین زلزله در دنباله ۱۰۰ تایی مشخص می‌شود.
- سپس بازه زمانی ۳۰ روزه پس از آن مورد بررسی قرار می‌گیرد.

• اگر در این بازه، حداقل یک زلزله با بزرگی ۵.۵ یا بیشتر رخ داده باشد، برچسب خروجی برابر ۱ در نظر گرفته می‌شود.

• در غیر این صورت، برچسب خروجی ۰ (رخداد منفی) اختصاص می‌یابد.

به این ترتیب، هر جفت (X, Y) نمایانگر یک سناریوی واقعی از تاریخچه لرزه‌ای و پیامدهای آن در آینده است.

ج. فرآیند تولید نمونه‌ها با الگوریتم پنجره متحرک

بر اساس تحقیق Keogh و همکاران (۲۰۰۱)، روش پنجره متحرک یک رویکرد کارآمد برای تحلیل و بخش‌بندی داده‌های سری زمانی است که در آن، جریان پیوسته داده‌ها به بازه‌هایی با طول ثابت تقسیم می‌شود. این بازه‌ها یا «پنجره‌ها» با یک گام مشخص بر روی کل داده حرکت کرده و می‌توانند دارای هم‌پوشانی یا بدون هم‌پوشانی باشند. هر پنجره به عنوان یک نمونه مستقل برای آموزش مدل در نظر گرفته می‌شود و حاوی اطلاعات زمانی و مکانی رویدادها است. این روش با تبدیل داده‌های خام به دنباله‌های ساختاریافته، به مدل اجازه می‌دهد الگوهای محلی و وابستگی‌های زمانی بلندمدت را بطور همزمان شناسایی کند. با توجه به مقاله Bhatia و همکاران (۲۰۲۳) در حوزه پیش‌بینی زلزله، استفاده از پنجره متحرک موجب افزایش توانایی مدل در استخراج الگوهای پیش‌نشانگر رویدادهای بزرگ و بهبود دقت پیش‌بینی می‌شود.

برای تولید مجموعه کامل نمونه‌ها، از الگوریتم پنجره متحرک با گام ۱ استفاده شد:

- ابتدا یک پنجره ۱۰۰ تایی از اولین رویدادها انتخاب می‌شود.
- سپس پنجره به اندازه یک رویداد به جلو حرکت کرده و نمونه جدیدی ایجاد می‌شود.
- این فرآیند تا انتهای مجموعه داده‌ها ادامه می‌یابد.

در نتیجه مجموعه‌ای بزرگ از داده‌های آموزشی تولید می‌شود که به مدل اجازه می‌دهد الگوهای زمانی موثر در پیش‌بینی زلزله‌های شدید را یاد بگیرد. همچنین یکی از پارامترهای حیاتی در این پژوهش، انتخاب بهینه طول دنباله رویدادهای زلزله به عنوان ورودی مدل بود. این مقدار با تحلیل‌های تجربی و آزمون‌های متعدد تعیین شده است. آزمایش‌هایی با طول دنباله‌های مختلف (۵۰، ۱۰۰، ۱۵۰ و ۲۰۰ رویداد) انجام شد و نتایج نشان داد که طول دنباله ۱۰۰ رویداد بهترین عملکرد را برای پیش‌بینی زلزله‌های بزرگ ارائه می‌دهد. این انتخاب به مدل اجازه می‌دهد تا وابستگی‌های زمانی بلندمدت و الگوهای آماری مهم در تاریخچه لرزه‌خیزی منطقه را شناسایی کند درحالی‌که از پیچیدگی محاسباتی غیرضروری ناشی از داده‌های بیش‌ازحد قدیمی جلوگیری می‌شود. این رویکرد به مدل کمک می‌کند تا تعادلی موثر بین اطلاعات کافی و بار محاسباتی برقرار کند.

د. نرمال‌سازی ویژگی‌ها

نرمال‌سازی ویژگی‌ها یکی از مراحل حیاتی در آماده‌سازی داده‌ها برای مدل‌های یادگیری عمیق است که به همگرایی سریع‌تر و بهبود عملکرد مدل کمک می‌کند. در این مرحله، تمامی ویژگی‌های ورودی مدل از جمله بزرگی، عمق و مختصات جغرافیایی، به یک محدوده مقیاس یکسان تبدیل شدند. در این تحقیق، از روش RobustScaler برای نرمال‌سازی استفاده شده است. بر اساس پژوهش Ioffe و همکاران (۲۰۱۵) این روش برخلاف روش‌های مبتنی بر میانگین و انحراف معیار، از چارک‌ها و میانه برای مقیاس‌بندی استفاده می‌کند و به همین دلیل در برابر داده‌های پرت، مقاوم است. این

فرآیند به مدل اجازه می‌دهد تا الگوهای واقعی در داده‌ها را بهتر شناسایی کرده و از تاثیر منفی مقادیر پرت جلوگیری کند که در نهایت به افزایش دقت و پایداری مدل کمک می‌کند. این روش براساس میانه^۱ و فاصله بین چارکی^۲ عمل می‌کند که در برابر داده‌های پرت، بسیار مقاوم است.

- ابتدا داده‌ها به آرایه‌ای دوبعدی (Flatten) تبدیل شدند.
- سپس نرمال‌سازی روی تمام ویژگی‌های عددی اعمال شد و داده‌ها به ساختار دنباله‌ای اولیه بازگردانده شدند.

یکی از چالش‌های مهم در استفاده از داده‌های سری زمانی بلندمدت، چگونگی مقابله با اهمیت متفاوت داده‌های قدیمی و جدید است. در این پژوهش، از داده‌های آموزشی قدیمی (مانند داده‌های دهه ۹۰ میلادی) استفاده شده و وزن‌دهی متفاوتی برای آنها در نظر گرفته نشده است. این تصمیم براساس دو دلیل اصلی اتخاذ شده است: اولاً فرض پدیده حافظه زمین‌ساختی که الگوهای لرزه‌خیزی را در طول زمان پایدار می‌داند. ثانیاً به دلیل اینکه مدل‌های یادگیری عمیق به ویژه مدل‌های LSTM و Transformer، به صورت ذاتی قادر به تشخیص و وزن‌دهی مناسب به داده‌های جدیدتر هستند. لایه LSTM با استفاده از دروازه‌های فراموشی و ورودی و لایه Transformer با استفاده از مکانیزم توجه، به صورت پویا اهمیت رویدادهای اخیر را نسبت به رویدادهای قدیمی‌تر در نظر می‌گیرند. این رویکرد به مدل اجازه می‌دهد تا بدون کاهش حجم داده‌های آموزشی و با حفظ اطلاعات تاریخی ارزشمند، تعادلی موثر بین یادگیری از گذشته و سازگاری با شرایط کنونی برقرار کند و از این طریق، از کاهش دقت پیش‌بینی جلوگیری شود.

۵. تقسیم داده‌ها

پس از آماده‌سازی، مجموعه داده لرزه‌ای به سه بخش تقسیم شد: ۷۰ درصد برای آموزش، ۱۵ درصد برای اعتبارسنجی و ۱۵ درصد برای آزمون. این تقسیم‌بندی یک روش استاندارد در یادگیری عمیق است که دلایل علمی و عملی مشخصی دارد. براساس کتاب Deep Learning اثر Goodfellow (۲۰۱۶) بزرگترین بخش داده (۷۰ درصد) برای آموزش مدل استفاده شد تا الگوریتم بتواند الگوهای پیچیده و پنهان را به صورت جامع استخراج کند. بخش اعتبارسنجی (۱۵ درصد) برای تنظیم هایپرپارامترها و جلوگیری از بیش‌برازش در طول آموزش بکار رفت. در نهایت بخش آزمون (۱۵ درصد) که هرگز توسط مدل دیده نشد، برای ارزیابی نهایی و بیطرفانه عملکرد مدل و سنجش قابلیت تعمیم‌پذیری آن استفاده گردید. در این روش تقسیم‌بندی داده‌ها به صورت گسترده در منابع معتبر یادگیری عمیق مورد تاکید قرار گرفته است. بنابراین برای ارزیابی عملکرد واقعی مدل، کل مجموعه داده‌ها سه بخش مجزا دارد:

- مجموعه آموزش^۳: ۷۰ درصد مجموعه اعتبارسنجی^۴: ۱۵ درصد مجموعه آزمون^۵: ۱۵ درصد

جهت جلوگیری از عدم تعادل بین کلاس‌های مثبت (برچسب ۱) و منفی (برچسب ۰) در هر بخش، از روش Stratified Split استفاده شد. این روش اطمینان حاصل می‌کند که نسبت کلاس‌ها در همه بخش‌ها حفظ شود و در مسائل نامتوازن

1 Median
2 IQR
3 Train
4 Validation
5 Test

مانند پیش‌بینی زلزله، اهمیت زیادی دارد. وجود این تعادل مانع از بایاس مدل به سمت کلاس اکثریت می‌شود و عملکرد آن را در تشخیص رخداد‌های نادر بهبود می‌بخشد.

۵. هایپر پارامترها

انتخاب و تنظیم هایپر پارامترهای بهینه، یک مرحله حیاتی در فرآیند آموزش مدل‌های یادگیری عمیق است که تاثیر مستقیمی بر عملکرد مدل دارد. در این پژوهش، از یک رویکرد نظام‌مند برای این منظور استفاده شد. ابتدا محدوده‌ای از مقادیر منطقی برای هایپر پارامترهای کلیدی براساس کارهای پیشین در این حوزه تعیین شد سپس از جستجوی شبکه‌ای^۱ برای ارزیابی جامع ترکیب‌های مختلف هایپر پارامترها استفاده گردید. عملکرد هر ترکیب بر روی مجموعه داده اعتبارسنجی ارزیابی شد و در نهایت ترکیبی که منجر به بالاترین دقت می‌شد، به عنوان تنظیمات نهایی مدل انتخاب شد. این فرآیند اطمینان داد که هایپر پارامترها به صورت بهینه و نه به صورت تصادفی یا دلخواه انتخاب شده‌اند.

۳-۳) انتخاب مدل

هدف این پژوهش بررسی و پیش‌بینی وقوع زلزله‌های بزرگتر از ۵.۵ ریشتر در منطقه زاگرس با استفاده از مدل‌های ترکیبی یادگیری عمیق است. بدین منظور از مدل‌های مختلف یادگیری عمیق برای پیش‌بینی وقوع زلزله در این پژوهش بکار گرفته شده‌است. جدول ۲ مدل‌های مورد استفاده در این پژوهش را نشان می‌دهد. دلیل استفاده از این مدل‌ها عبارتند از:

- شبکه عصبی کانولوشنی^۲: برای شبیه‌سازی ویژگی‌های مکانی داده‌ها.
- حافظه طولانی کوتاه مدت^۳: برای پردازش وابستگی‌های زمانی در داده‌ها.
- ترنسفورمر^۴: برای مدل‌سازی وابستگی‌های بلندمدت و پیچیدگی‌های زمانی داده‌ها.
- مدل ترکیبی: ترکیب نتایج مدل‌های مختلف برای بهبود دقت پیش‌بینی.

جدول ۲. مدل‌های مورد استفاده

نقاط قوت	مدل یادگیری عمیق
استخراج ویژگی‌های مکانی و شناسایی الگوهای داده‌های ساختاریافته	CNN
مدیریت موثر وابستگی‌های زمانی، به ویژه در داده‌های دنباله‌ای	LSTM
مدل‌سازی وابستگی‌های بلندمدت و پیچیده زمانی با مکانیسم Attention	Transformer
ترکیب قدرت چندین مدل برای افزایش دقت و قابلیت اطمینان	مدل ترکیبی

همچنین در طراحی معماری مدل ترکیبی، انتخاب تعداد لایه‌ها برای هر بخش یکی از مهمترین تصمیمات بود. ما معماری‌های متنوعی را با تعداد لایه‌های متفاوت مورد آزمایش قرار دادیم و مشاهده کردیم که استفاده از لایه‌های بیشتر، اغلب به بیش‌برازش منجر می‌شود. در این حالت، مدل عملکرد بسیار خوبی بر روی داده‌های آموزشی دارد اما در مواجهه

1 Grid Search

2 CNN

3 LSTM

4 Transformer

با داده‌های جدید، دقت آن به شدت کاهش می‌یابد. از سوی دیگر، تعداد لایه‌های بسیار کمتر باعث کم‌برازش شده و مدل قادر به یادگیری الگوهای پیچیده داده‌ها نبود. در نهایت با در نظر گرفتن محدودیت‌های محاسباتی و پس از ارزیابی نتایج، معماری دولایه‌ای برای هر بخش انتخاب شد. این پیکربندی توانست بهترین تعادل را بین قدرت مدل برای یادگیری الگوها و قابلیت تعمیم‌پذیری آن به دست آورد و در آزمون‌های ما، بهترین نتایج را از خود نشان داد. در ادامه، معماری مدل‌های مختلف مورد استفاده در تحقیق شامل کانولوشنی، حافظه طولانی کوتاه مدت، ترنسفورمر و مدل ترکیبی توضیح داده می‌شود. هر معماری با هدف استخراج ویژگی‌ها و شبیه‌سازی وابستگی‌های زمانی یا مکانی طراحی شده است.

۳-۳-۱) معماری مدل کانولوشنی

به منظور استخراج ویژگی‌های مکانی و زمانی نهفته در توالی‌های زلزله، یک مدل شبکه عصبی پیچشی طراحی و پیاده‌سازی گردید. ساختار این مدل به گونه‌ای طراحی شده است که قادر است الگوهای پیچیده را از توالی‌های لرزه‌ای ورودی استخراج کرده و در عین حال از بیش‌برازش جلوگیری نماید. در این پژوهش از یک معماری ترکیبی استفاده شده است که یکی از بخش‌های اصلی آن را شبکه عصبی کانولوشنی تشکیل می‌دهد. این شبکه‌ها یک مدل یادگیری عمیق هستند که بطور خاص برای پردازش داده‌هایی با ساختار شبکه‌ای مانند سری‌های زمانی طراحی شده‌اند. CNNها با بهره‌گیری از لایه‌های کانولوشن، به صورت خودکار و سلسله‌مراتبی، الگوها و ویژگی‌های محلی را از داده‌های ورودی استخراج می‌کنند. این ویژگی، CNNها را به گزینه‌ای ایده‌آل برای تحلیل داده‌های لرزه‌ای و شناسایی الگوهای پنهان در آنها تبدیل کرده است. جدول شماره ۳ ساختار لایه‌های این مدل را به تفصیل نمایش می‌دهد.

جدول ۳. معماری مدل شبکه عصبی کانولوشنی

نوع لایه	شکل خروجی	پارامترها
conv1d (کانولوشن یک بعدی)	(None, 98, 32)	۴۱۶
max_pooling1d (پولینگ بیشینه یک بعدی)	(None, 49, 32)	۰
conv1d (کانولوشن یک بعدی)	(None, 47, 64)	۶,۲۰۸
max_pooling1d (پولینگ بیشینه یک بعدی)	(None, 23, 64)	۰
global_max_pooling1d (پولینگ بیشینه سراسری یک بعدی)	(None, 64)	۰
dense_8 (متراکم)	(None, 64)	۴,۱۶۰
dropout_6 (حذف تصادفی)	(None, 64)	۰
dense_9 (متراکم)	(None, 1)	۶۵

مدل طراحی شده شامل دو لایه Conv1D است. در لایه نخست، ۳۲ فیلتر با اندازه کرنل ۳ بر روی ورودی اعمال شده و خروجی به شکل (۳۲, ۹۸) تولید می‌گردد. پس از آن، با استفاده از لایه ادغام حداکثر^۱، ابعاد زمانی داده‌ها به نصف

1 MaxPooling1D

کاهش یافته و به خروجی (۳۲, ۴۹) منجر می شود. در ادامه، دومین لایه با ۶۴ فیلتر بکار گرفته شده است که پس از پردازش، خروجی را به شکل (۶۴, ۴۷) تبدیل می کند. برای کاهش بیشتر ویژگی های غیر ضروری و تمرکز بر قوی ترین ویژگی ها، مجدداً از لایه ادغام حداکثر استفاده شده که خروجی نهایی این مرحله به صورت (۶۴, ۲۳) خواهد بود. به منظور کاهش ابعاد و حفظ مهمترین ویژگی ها، از لایه متوسط گیری سراسری یک بعدی^۱ بهره گرفته شده که بردار خروجی ۶۴ بعدی تولید می کند. در ادامه، این بردار از طریق یک لایه تمام متصل^۲ با ۶۴ نورون و تابع فعال ساز ReLU پردازش می شود. برای مقابله با بیش برآزش، از لایه Dropout با نرخ ۲۰ درصد استفاده شده است. در نهایت خروجی مدل با استفاده از لایه تمام متصل نهایی با تنها یک نورون و تابع فعال ساز سیگموید، احتمال وقوع زلزله ای بزرگتر از ۵.۵ ریشتر در بازه ۳۰ روز آینده را پیش بینی می کند. مدل مذکور در مجموع ۱۱۰۸۴۹ پارامتر قابل آموزش دارد که بطور مناسبی میان استخراج ویژگی و پیچیدگی محاسباتی توازن ایجاد می کند. این معماری سبک، برای کاربردهای عملی با محدودیت منابع محاسباتی نیز مناسب است.

۳-۳-۲) معماری مدل حافظه کوتاه-بلندمدت^۳

در این پژوهش، برای پردازش داده های سری زمانی از یک شبکه عصبی بازگشتی مبتنی بر واحدهای حافظه طولانی - کوتاه مدت استفاده شده است. معماری مدل به صورت سلسله مراتبی و متشکل از چند لایه طراحی شده که جزئیات آن در جدول ۴ ارائه شده است:

جدول ۴. معماری مدل شبکه عصبی بازگشتی

نوع لایه	شکل خروجی	پارامترها
lstm (حافظه طولانی-کوتاه مدت)	(None, 128)	۶۸,۰۹۶
dropout (حذف تصادفی)	(None, 128)	۰
dense (متراکم)	(None, 64)	۸,۲۵۶
dropout_1 (حذف تصادفی)	(None, 64)	۰
dense_1 (متراکم)	(None, 1)	۶۵

این معماری با داشتن یک لایه LSTM با ۱۲۸ واحد حافظه، قادر به استخراج و مدل سازی وابستگی های بلندمدت و کوتاه مدت داده های سری زمانی است. تعداد پارامترهای این لایه ۱۶۵,۸۸۸ پارامتر بوده که شامل وزن های ورودی، وزن های بازگشتی و بایاس های مرتبط با دروازه های حافظه می باشد و توان بالای مدل را در یادگیری ویژگی های پیچیده و غیرخطی داده های زمانی نشان می دهد. پس از آن یک لایه Dropout بدون پارامتر برای کاهش بیش برآزش و افزایش تعمیم پذیری مدل اضافه شده است که به صورت تصادفی نورون های لایه قبلی را در زمان آموزش غیرفعال می کند و شکل خروجی را حفظ می نماید. در ادامه، یک لایه Dense با ۶۴ نورون و ۸,۲۵۶ پارامتر قرار گرفته است که وظیفه آن یادگیری ترکیب های غیرخطی ویژگی های استخراج شده توسط لایه LSTM است و شامل وزن های اتصال و بایاس ها می باشد.

1 GlobalMaxPooling

2 Dense

3 LSTM

سپس مجدداً یک لایه Dropout دیگر بدون پارامتر برای جلوگیری از بیش‌برازش استفاده شده‌است که خروجی لایه Dense را به همان شکل حفظ می‌کند. در نهایت یک لایه Dense خروجی با یک نرون و ۶۵ پارامتر قرار دارد که وزن‌ها و بایاس این نرون را شامل می‌شود و خروجی نهایی مدل را برای انجام پیش‌بینی تولید می‌کند. کل معماری دارای ۱۷۴,۲۰۹ پارامتر است که با توجه به تعداد پارامترها و ساختار لایه‌ها، مدلی با توانایی بالا در یادگیری الگوهای زمانی پیچیده به شمار می‌رود. استفاده از لایه‌های Dropout به‌عنوان تکنیکی موثر در جلوگیری از بیش‌برازش، موجب شده‌است مدل قابلیت تعمیم بهتری داشته باشد و در مسائلی همچون پیش‌بینی و طبقه‌بندی داده‌های سری زمانی عملکرد قابل قبولی ارائه دهد. این ساختار بطور متعادل بین توانایی یادگیری و جلوگیری از بیش‌برازش تنظیم شده‌است و بنابراین چارچوب مناسبی برای تحلیل داده‌های ترتیبی و استخراج الگوهای معنادار فراهم می‌آورد.

۳-۳-۳ معماری مدل ترنسفورمر^۱

در این پژوهش، معماری مدل ترنسفورمر به‌منظور پردازش داده‌های ترتیبی طراحی شده‌است که ساختاری متشکل از چندلایه کلیدی دارد و در جدول ۵ خلاصه شده‌است:

جدول ۵. معماری مدل ترنسفورمر

نوع لایه	شکل خروجی	پارامترها
لایه ورودی	(None, 100, 4)	۰
توجه چند-سر (MultiHeadAttention)	(None, 100, 4)	۳۰۸
نرمال‌سازی لایه (LayerNormalization)	(None, 100, 4)	۸
پولینگ میانگین سراسری ۱ بعدی (GlobalAveragePooling1D)	(None, 4)	۰
متراکم ۱۶ (Dense)	(None, 64)	۳۲۰
حذف تصادفی ۱۴ (Dropout)	(None, 64)	۰
متراکم ۱۷ (Dense)	(None, 1)	۶۵

این معماری از یک لایه ورودی با شکل (۴, ۱۰۰) شروع می‌شود که داده‌های ترتیبی با طول ۱۰۰ و ویژگی‌های ۴ را دریافت می‌کند. سپس داده‌ها به لایه چندسر توجه^۲ با ۳۰۸ پارامتر وارد می‌شوند که با استفاده از مکانیسم توجه، روابط بین موقعیت‌های مختلف داده‌های ورودی را به‌صورت همزمان و پویا، مدل‌سازی می‌کند. این لایه قابلیت یادگیری ارتباطات پیچیده و وابستگی‌های بلندمدت در داده‌های سری زمانی را فراهم می‌آورد. پس از لایه توجه، داده‌ها از طریق

1 Transformer

2 MultiHeadAttention

لایه نرمال سازی^۱ عبور می کنند که با ۸ پارامتر به بهبود ثبات و سرعت همگرایی مدل کمک کرده و توزیع داده ها را در هر لایه بهبود می بخشد.

سپس یک لایه متوسط گیری سراسری یک بعدی^۲ اعمال می شود که خروجی را از شکل (None, 100, 4) به (None, 4) تبدیل می کند و با خلاصه سازی ویژگی های زمانی، ابعاد داده ها را کاهش می دهد و اطلاعات کلیدی را حفظ می کند. پس از آن، یک لایه کاملاً متصل^۳ با ۶۴ نورون و ۳۲۰ پارامتر برای یادگیری ترکیب های غیر خطی ویژگی های استخراج شده بکار می رود. این لایه نقش کلیدی در یادگیری الگوهای پیچیده و افزایش ظرفیت مدل برای انجام پیش بینی ها دارد. به منظور جلوگیری از بیش برآزش، یک لایه Dropout بدون پارامتر پس از آن قرار گرفته که برخی نورون ها را به صورت تصادفی در زمان آموزش غیرفعال می کند و به بهبود تعمیم پذیری مدل کمک می نماید. در نهایت یک لایه خروجی تمام متصل با یک نورون و ۶۵ پارامتر، مسئول تولید خروجی نهایی مدل است که معمولاً با تابع فعال سازی مناسب برای مسئله (مثلاً سیگموئید در مسائل طبقه بندی باینری) بکار می رود. کل معماری مذکور با کمترین تعداد پارامتر نسبت به مدل های بزرگتر، ساختاری کارآمد و بهینه جهت استخراج و مدل سازی ویژگی های پیچیده سری های زمانی فراهم می آورد و به کمک مکانیسم توجه، قابلیت یادگیری روابط بلندمدت و مهم بین نقاط داده را دارد. استفاده از لایه های نرمال سازی و دراپ اوت نیز باعث افزایش پایداری و تعمیم پذیری مدل شده است. این ترکیب منجر به مدلی می شود که می تواند به خوبی در مسائل پیش بینی و طبقه بندی داده های ترتیبی، عملکرد مناسبی ارائه دهد.

۳-۳-۴) معماری مدل ترکیبی

به منظور بهره گیری از مزایای مکمل ساختارهای گوناگون یادگیری عمیق در تحلیل داده های پیچیده و غیریستای لرزه ای، مدلی ترکیبی طراحی شده است که بطور همزمان از قابلیت های شبکه های عصبی پیچشی برای استخراج ویژگی های مکانی، شبکه های بازگشتی حافظه دار برای مدل سازی وابستگی های زمانی و سازوکار توجه چندسری^۴ برای درک تعاملی وابستگی ها میان ویژگی ها استفاده می کند. هدف اصلی از این ترکیب، ارتقای توان مدل در شناسایی الگوهای پنهان و پرنوسان موجود در داده های زلزله به ویژه در مناطق با رفتار لرزه ای پیچیده مانند زاگرس است. معماری این مدل در جدول ۶ به خوبی آمده است.

جدول ۶. معماری مدل ترکیبی

نوع لایه	شکل خروجی	پارامترها	متصل به
لایه ورودی	(None, 100, 4)	۰	-
کانولوشن ۱ بعدی ۲ (Conv1D)	(None, 100, 64)	۸۳۲	لایه ورودی ۱
پولینگ بیشینه ۱ بعدی ۱ (MaxPooling1D)	(None, 50, 64)	۰	کانولوشن ۱ بعدی ۲
کانولوشن ۱ بعدی ۳ (Conv1D)	(None, 50, 128)	۲۴,۷۰۴	پولینگ بیشینه ۱ بعدی ۱

1 LayerNormalization
2 GlobalAveragePooling1D
3 Dense
4 Multi-Head Attention

لایه ورودی_۱	۱۷,۶۶۴	(None, 100, 64)	lstm_2
کانولوشن_۱_بعدی_۳	۰	(None, 128)	پولینگ_بیشینه_سراسری_۱_بعدی_۱ (GlobalMaxPooling1D)
lstm_2	۱۲,۴۱۶	(None, 32)	lstm_3
پولینگ_بیشینه_سراسری_۱_بعدی_۱، lstm_3	۰	(None, 160)	به هم پیوستن_۱ (Concatenate)
به هم پیوستن_۱	۰	(None, 1, 160)	تغییر_شکل_۱ (Reshape)
تغییر_شکل_۱، تغییر_شکل_۱، تغییر_شکل_۱	۸۲,۴۶۴	(None, 1, 160)	توجه_چند_سر (MultiHeadAttention)
توجه_چند_سر	۰	(None, 160)	مسطح_سازی_۱ (Flatten)
مسطح_سازی_۱	۲۰,۶۰۸	(None, 128)	متراکم_۲ (Dense)
متراکم_۲	۰	(None, 128)	حذف_تصادفی_۳ (Dropout)
حذف_تصادفی_۳	۱۲۹	(None, 1)	متراکم_۳ (Dense)

در معماری ترکیبی پیشنهادی، از دو مسیر موازی برای استخراج اطلاعات مکانی و زمانی استفاده شده است تا تعامل میان این دو نوع ویژگی به خوبی شناسایی شود. در مسیر اول که شاخه کانولوشن نام دارد، ابتدا توالی‌های لرزه‌ای با ابعاد (۴, ۱۰۰) وارد لایه Conv1D با ۶۴ فیلتر می‌شوند تا ویژگی‌های محلی و الگوهای ابتدایی از داده‌های چهاربعدی استخراج شود. بهره‌گیری از اندازه کرنل مناسب، این امکان را می‌دهد که تغییرات کوچک مکانی-زمانی در رخدادها بطور موثری تشخیص داده شود. خروجی این لایه با عبور از لایه ادغام حداکثری، ابعاد زمانی خود را به نصف کاهش می‌دهد بطوریکه نویزهای غیرضروری حذف و تمرکز بر الگوهای غالب حفظ می‌شود. پس از آن، یک لایه Conv1D دوم با ۱۲۸ فیلتر برای استخراج عمیق‌تر ویژگی‌ها بکار می‌رود، به گونه‌ای که این لایه قادر باشد ساختارهای پیچیده‌تر مکانی در دنباله را مدل‌سازی کند. در نهایت با عبور از لایه متوسط‌گیری سراسری یک‌بعدی، تمامی اطلاعات مکانی-زمانی جمع‌بندی شده و به یک بردار ۱۲۸ بعدی تبدیل می‌شود که نماینده قوی از مهمترین ویژگی‌های مکانی کل توالی است.

در مسیر دوم که شاخه بازگشتی نامیده می‌شود، داده‌های ورودی به همان شکل (۴, ۱۰۰) مستقیماً وارد یک لایه LSTM با ۱۲۸ واحد و تنظیم return_sequences=True می‌گردند تا وابستگی‌های زمانی در تمامی گام‌های توالی حفظ شود. این رویکرد به مدل امکان می‌دهد نسبت به تغییرات تدریجی در الگوی لرزه‌ای حساسیت بیشتری داشته باشد. خروجی لایه اول LSTM با شکل (۱۲۸, ۱۰۰) به لایه LSTM دوم متصل می‌شود تا وابستگی‌های بلندمدت و پیچیده‌تر در دنباله‌های لرزه‌ای به یک بردار ۳۲ بعدی فشرده شود. این فشرده‌سازی موجب می‌شود اطلاعات زمانی کل توالی به صورت یک ویژگی ساخته شده هماهنگ به دست آید.

پس از طی شدن مراحل فوق، بردارهای ۱۲۸ بعدی مکانی و ۳۲ بعدی زمانی به وسیله لایه الحاق کننده^۱ با یکدیگر ادغام می شوند و یک بردار ۱۶۰ بعدی تشکیل می دهند. به منظور آماده سازی این بردار ترکیبی برای پردازش در لایه توجه چندسری، ابتدا ابعاد آن توسط لایه Reshape به شکل (۱، ۱۶۰) تغییر می یابد تا بتواند به خوبی به عنوان Q(Query)، K(Key) و V(Value) در سازوکار Attention استفاده شود. در لایه چندسری توجه^۲، با تخصیص چندین «سر توجه» مختلف، امکان تمرکز همزمان بر بخش های گوناگون بردار ترکیبی فراهم می آید. این فرآیند باعث می شود مدل بتواند روابط متقابل بین ویژگی های مکانی و زمانی را به صورت پویا و تطبیقی یاد بگیرد و بخش هایی از داده که در پیش بینی وقوع زلزله اهمیت بالاتری دارند را برجسته کند. خروجی این لایه با شکل (۱، ۱۶۰) سپس با لایه Flatten تخت می شود تا به فرم یک بردار ۱۶۰ بعدی یک بعدی درآید. در ادامه، یک بلوک شامل لایه Dense با ۱۲۸ نورون و تابع فعال ساز ReLU قرار دارد که وظیفه آن یادگیری ترکیب های غیرخطی میان ویژگی های به دست آمده از Attention است. شکل ۳ به خوبی فرآیند کار این مدل رو به نمایش گذاشته است. لازم به ذکر است برای جلوگیری از پدیده بیش برآزش در مدل ترکیبی که با بیش از ۱۵۹ هزار پارامتر طراحی شده است، از چندین تکنیک پیشرفته استفاده شد. لایه های Dropout با نرخ ۰.۲ به صورت استراتژیک پس از لایه های اصلی مانند LSTM و Dense قرار داده شدند تا از وابستگی بیش از حد مدل به ویژگی های خاص جلوگیری کرده و قابلیت تعمیم آن را افزایش دهند. علاوه بر این، از Early Stopping استفاده شد تا به صورت خودکار، آموزش مدل را در صورت عدم بهبود عملکرد بر روی مجموعه داده اعتبارسنجی متوقف کند و بدین ترتیب از حفظ کردن داده های آموزشی توسط مدل، جلوگیری به عمل آورد. این رویکردها اطمینان می دهند که مدل بجای حفظ جزئیات بی اهمیت، الگوهای اساسی و قابل تعمیم لرزه خیزی را یاد می گیرد.

1 Concatenate

2 Multi-Head Attention



در مجموع، این معماری شامل ۱۵۹'۸۱۷ پارامتر قابل آموزش است که با تقسیم دقیق میان استخراج ویژگی‌های پیچیده مکانی، تحلیل وابستگی‌های زمانی و تمرکز تطبیقی از طریق مکانیسم توجه، تعادلی مطلوب میان قدرت مدل‌سازی و پیچیدگی محاسباتی برقرار کرده‌است. نتیجه چنین ساختاری، آن است که مدل ترکیبی می‌تواند الگوهای نهفته و پیچیده در دنباله‌های لرزه‌ای را بهتر شناسایی کند و در مقایسه با مدل‌های پایه مجزای CNN یا LSTM یا ترنسفورمر، دقت و پایداری بالاتری در پیش‌بینی وقوع زلزله‌های بزرگ ارائه دهد.

۴) یافته‌ها

در این بخش به نمایش نتایج خود و مقایسه با کارهای پیشین می‌پردازیم.

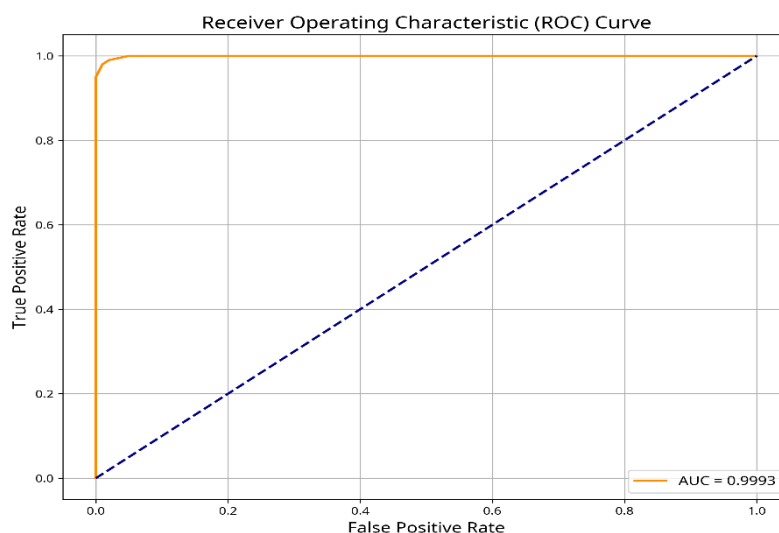
۴-۱) نتایج من

در ابتدا، سه مدل پایه شامل شبکه عصبی پیچشی، شبکه حافظه‌دار بلندمدت و ترنسفورمر پیاده‌سازی شدند. سپس مدل ترکیبی پیشنهادی شامل CNN + LSTM + Multi-Head Attention طراحی و ارزیابی گردید. نتایج ارزیابی این مدل‌ها در جدول ۷ ارائه شده‌اند.

جدول ۷. نتایج ارزیابی مدل‌های ساده و ترکیبی

مدل	Accuracy (%)	Loss
CNN	۹۶.۴۹	۷
LSTM	۹۴.۸۳	۱۹.۷
Transformer	۹۴.۸۶	۱۹
Hybrid	۹۹.۵۱	۴

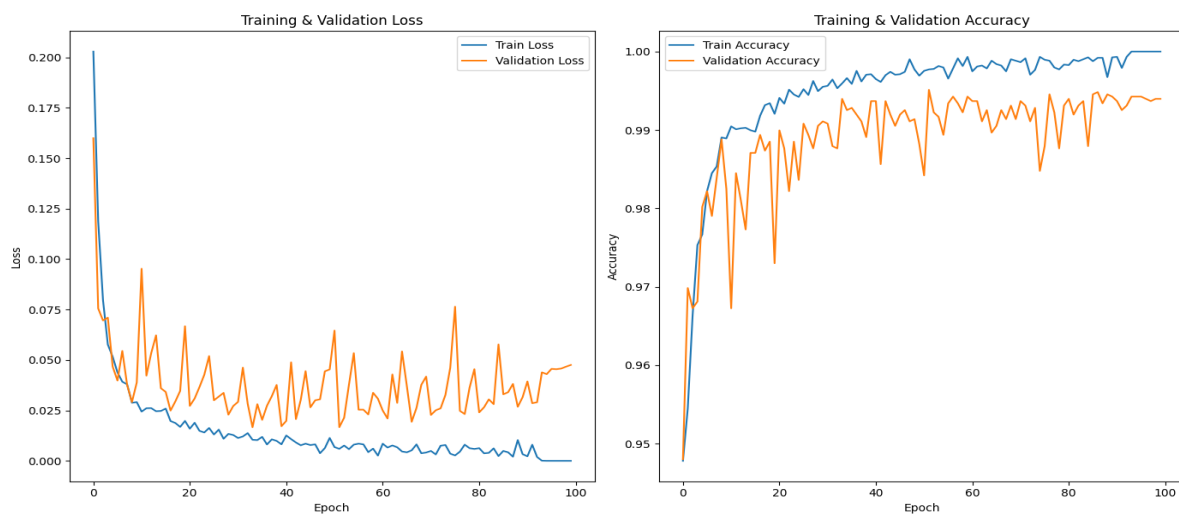
همچنین شکل ۴، منحنی ROC مدل ترکیبی پیشنهادی را نشان می‌دهد.



شکل ۴. نمودار ROC مدل ترکیبی

در این نمودار، منحنی نزدیک به گوشه بالا - چپ قرار دارد که نشان از عملکرد بسیار خوب مدل دارد. همچنین مقدار $AUC = 0.9993$ (مساحت زیرمنحنی)، بسیار نزدیک به عدد ایده‌آل ۱ است و نشان می‌دهد مدل توانایی بسیار بالایی در تشخیص تفکیک کلاس‌ها (وقوع یا عدم‌وقوع زلزله) دارد. این مقدار AUC بیانگر آن است که مدل در ۹۹.۹۳٪

درصد مواقع قادر است یک نمونه مثبت را از یک نمونه منفی به درستی تمیز دهد که معیاری قوی برای عملکرد کلی مدل محسوب می‌شود. چنین عملکردی در کاربردهای عملی به‌ویژه در حوزه‌های حساس مانند پیش‌بینی زلزله، بسیار ارزشمند است زیرا نشان می‌دهد مدل، توانایی بالایی در کاهش نرخ هشدارهای غلط و افزایش دقت هشدارهای واقعی دارد. در ادامه، شکل ۵ نشان‌دهنده روند کاهش خطای آموزش و اعتبارسنجی در طول ۱۰۰ دوره آموزش است که کاهش یکنواخت و پایداری را در خطای آموزش و نوسانات محدودی را در خطای اعتبارسنجی نشان می‌دهد. نمودار سمت راست نیز بیانگر روند افزایشی دقت در هر دو مجموعه آموزش و اعتبارسنجی است که به دقت نهایی بیش از ۹۹ درصد در پایان آموزش منتهی شده‌است. این نتایج بیانگر یادگیری موثر مدل و تعمیم‌پذیری قابل قبول آن در داده‌های جدید است.



شکل ۵. نمودار تغییرات تابع هزینه و دقت در مراحل آموزش و اعتبارسنجی مدل ترکیبی

۴-۲) مقایسه نتایج خود با کارهای پیشین

به‌منظور سنجش برتری نسبی مدل پیشنهادی، نتایج آن با نتایج گزارش‌شده در پژوهش‌های مرتبط هاشمی و همکاران (۲۰۲۴) در زمینه پیش‌بینی زلزله مقایسه شده‌است. اگرچه تحقیقات متعددی در این زمینه صورت گرفته‌است اما اغلب این مطالعات بر مناطق دیگر تمرکز داشته یا از روش‌های آماری و کلاسیک استفاده کرده‌اند. در میان این پژوهش‌ها، کار هاشمی و همکاران (۲۰۲۴) به‌عنوان یکی از معتبرترین و نزدیک‌ترین مطالعات به پژوهش حاضر، به‌دلیل تشابه در دو عامل کلیدی انتخاب شد. این مقاله با استفاده از داده‌های کاتالوگ زلزله مرکز لرزه‌نگاری ایران و با هدف پیش‌بینی زلزله‌های بزرگ (بزرگی ۵.۵ ریشتر و بالاتر) در منطقه زاگرس، به بررسی و ارزیابی عملکرد سه مدل یادگیری ماشین پرداخته‌است. این مدل‌ها شامل شبکه‌های عصبی مصنوعی، جنگل تصادفی و ماشین بردار پشتیبان هستند. درنهایت با ارزیابی عملکرد این مدل‌ها براساس معیارهای آماری مختلف، مشخص شده‌است که مدل ANN عملکرد بهتری در پیش‌بینی زلزله داشته‌است. بنابراین این مقاله می‌تواند به‌عنوان یک مرجع کلیدی برای مقایسه و تحلیل رویکرد شما در زمینه پیش‌بینی زلزله در این منطقه مورد استفاده قرار گیرد.

این مقایسه در جدول شماره ۸ ارائه شده‌است.

جدول ۸. مقایسه نتایج کار خود و کار پیشین (مجموعه داده زاگرس)

منبع	الگوریتم	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
روش پیشنهادی	ترکیبی	۹۹/۵۱	۹۶/۶	۹۳/۹۲	۹۵/۲۴
هاشمی و همکاران	SVM	۹۱/۶	۶۲/۶	۹۳	۷۳
	RF	۹۳	۷۳	۹۳	۸۱
	ANN	۹۶	۸۵	۹۷	۹۰

همانطور که مشاهده می شود، مدل ترکیبی پیشنهادی در اغلب شاخص ها از جمله دقت، یادآوری و امتیاز F1، نسبت به سایر روش های ذکر شده در ادبیات پژوهش، عملکرد بهتری داشته است. علت این برتری را می توان در استفاده از ساختار ترکیبی و چندمرحله ای مدل دانست که از نقاط قوت هر مولفه ویژگی های مکانی CNN، وابستگی زمانی LSTM و توجه پویا در مدل چندمرحله ای بهره می برد. برتری مدل ما از ترکیب هوشمندانه لایه های CNN (برای ویژگی های مکانی)، LSTM (برای وابستگی های زمانی) و Transformer (برای روابط پیچیده) حاصل شده است که تحلیل جامع تری را ممکن می سازد. با وجود عدم استفاده از داده های ژئوفیزیکی، قابلیت تعمیم مدل قابل قبول است و این محدودیت می تواند در پژوهش های آتی رفع شود. همچنین پایداری مدل در مواجهه با داده های نویزی و توانایی آن در یادگیری الگوهای نهان زمانی- مکانی، آن را به گزینه ای مناسب برای توسعه سیستم های هشدار زلزله در مناطق لرزه خیز تبدیل می کند. در این پژوهش، یکی از چالش های اصلی، اثبات کارایی و قابلیت تعمیم مدل پیشنهادی در مقایسه با کارهای پیشین بود. رویکرد ما بر پایه ارزیابی توانایی مدل های یادگیری عمیق در استخراج الگوهای زمانی- مکانی از داده های لرزه ای تمرکز داشت. درحالی که این رویکرد ممکن است در ابتدا به نظر محدود کننده برسد اما به ما اجازه داد تا عملکرد خالص مدل را بدون تاثیر ویژگی های جانبی بسنجیم. برای پاسخ به این چالش، ما عملکرد مدل ترکیبی را نه تنها با مدل های پایه بلکه با نتایج کارهای پیشین در مجموعه داده های دیگر نیز مقایسه کردیم. جدول ۹، عملکرد مدل ما را با برخی از کارهای پیشین در زمینه پیش بینی زلزله نشان می دهد.

جدول ۹. ارزیابی نتایج مدل ترکیبی با مجموعه داده های دیگر

داده	منبع	الگوریتم	نتایج		
			Accuracy	precision	Recall
آمریکا	روش پیشنهادی	ترکیبی	۹۸/۱۱	۹۶/۸۰	۹۸/۲۰
	Asim و همکاران	SVR	۸۸/۲	۸۹/۹	۵۵/۸
	(۲۰۱۸)	HNN	۸۷/۷	۸۶/۸	۵۵
	Alvarez و همکاران	AdaBoost	۸۷/۴	۸۴/۲	۶۸/۷
	(۲۰۱۸)				
	Panakkat و همکاران	RNN	۷۵/۲	۷۱	۷۱

(۲۰۰۷)					
۹۸/۵۲	۹۸/۶۳	۹۸/۴۰	۹۷/۶۸	ترکیبی	روش پیشنهادی
۳۷	۴۳/۷	۷۲/۳	۷۹/۹	SVR	Asim و همکاران
۴۷	۵۴/۷	۷۳/۸	۸۱/۸	HNN	(۲۰۱۸)
۵۵	۶۱/۲	۸۰/۲	۸۴/۵	AdaBoost	Alvarez و همکاران (۲۰۱۸)
***	۸۹/۲	۶۱/۱	۷۹/۷	ANN	Reyes و همکاران (۲۰۱۲)
۹۹/۳	۹۹/۳۲	۹۸/۶۷	۹۸/۱۰	ترکیبی	روش پیشنهادی
۳۹	۵۰	۶۹/۲	۷۶/۱	SVR	Asim و همکاران
۳۹	۵۴/۱	۶۴	۷۵/۲	HNN	(۲۰۱۸)
۴۷	۵۷/۷	۷۴/۳	۷۸/۷	AdaBoost	Alvarez و همکاران (۲۰۱۸)
***	۸۶	۶۹	۷۹	EPSO	Aslam و همکاران (۲۰۲۱)

۵) نتیجه‌گیری و پیشنهادات

در این پژوهش، با هدف پیش‌بینی وقوع زمین‌لرزه‌هایی با بزرگی بیش از ۵.۵ ریشتر در منطقه شمال زاگرس به‌ویژه ناحیه سرپل ذهاب، از مجموعه‌ای از مدل‌های یادگیری عمیق بهره گرفته شد تا با تلفیق اطلاعات مکانی و زمانی، الگوهای پنهان در داده‌های لرزه‌ای استخراج شود. مجموعه داده شامل بیش از ۲۳ هزار رخداد لرزه‌ای طی سال‌های ۱۹۹۰ تا ۲۰۲۴ بود که پس از پاک‌سازی و پردازش، در قالب دنباله‌های ۱۰۰ تایی برای آموزش مدل‌ها آماده شد. برچسب‌گذاری داده‌ها نیز براساس وقوع زمین‌لرزه‌های شدید در بازه ۳۰ روزه پس از هر دنباله صورت گرفت که امکان مدل‌سازی مسئله به‌صورت طبقه‌بندی دودویی را فراهم کرد. سه معماری یادگیری عمیق شامل CNN، LSTM، Transformer و ترکیب آنها طراحی و ارزیابی شدند. مدل ترکیبی پیشنهادی که از مسیرهای موازی CNN و LSTM برای استخراج ویژگی‌های مکانی و زمانی

و همچنین لایه توجه برای تمرکز تطبیقی استفاده می‌کرد، توانست عملکرد بهتری نسبت به سایر مدل‌ها ارائه دهد. دقت این مدل در مجموعه آزمون ۹۹٫۵۱ درصد و مقدار AUC آن ۰٫۹۹۹۳ گزارش شد درحالی‌که مقادیر دقت و موفقیت نیز نشان‌دهنده توان بالای مدل در کاهش هشدارهای غلط و شناسایی موثر رخدادها و واقعی بود. نوآوری اصلی پژوهش حاضر، طراحی یک مدل ترکیبی نوین برای پیش‌بینی زلزله است که از سه معماری قدرتمند یادگیری عمیق یعنی CNN، LSTM و Transformer بهره می‌برد. این ترکیب به مدل امکان می‌دهد تا بطور همزمان، الگوهای مکانی، وابستگی‌های زمانی و روابط پیچیده بین رویدادها را به صورت جامع تحلیل کند که این امر در مقایسه با روش‌های پیشین که عموماً از مدل‌های ساده یا روش‌های آماری استفاده می‌کردند، یک گام به جلو محسوب می‌شود. همچنین برتری روش پیشنهادی در مقایسه با روش‌های موجود، دقت بالاتر، قابلیت تعمیم‌پذیری و مقاومت بیشتر در برابر داده‌های پرت است. مدل ما توانسته است با موفقیت، تعادلی میان پیچیدگی مدل برای یادگیری الگوهای عمیق و قابلیت تعمیم آن به مجموعه داده‌های مختلف برقرار کند که این امر آن را به یک ابزار قدرتمند و قابل اعتماد برای پیش‌بینی زلزله تبدیل می‌کند. ازجمله نقاط قوت این پژوهش، می‌توان به استفاده از داده‌های به‌روز و گسترده، انتخاب پنجره‌های زمانی مناسب، بکارگیری ساختارهای پیشرفته یادگیری عمیق و نیز بهره‌مندی از روش‌های بهینه‌سازی مدل مانند Dropout و نرمال‌سازی اشاره کرد. این عوامل موجب شدند مدل نهایی به تعادلی میان دقت بالا و پایداری در برابر نویز برسد. در این پژوهش، مدل ترکیبی پیشنهادی ما به گونه‌ای طراحی شده که قادر به شناسایی و مدل‌سازی تغییرات بلندمدت در الگوهای لرزه‌ای است. لایه‌های LSTM با بهره‌گیری از قابلیت خود در حفظ اطلاعات بلندمدت و لایه Transformer با استفاده از مکانیزم توجه قادر به درک روابط بین رویدادهای لرزه‌ای با فاصله زمانی زیاد هستند. این معماری به مدل اجازه می‌دهد تا پدیده‌هایی مانند تغییرات در نرخ لرزه‌خیزی را به صورت پویا در نظر بگیرد و از این الگوها برای پیش‌بینی استفاده کند. با این حال، لازم به ذکر است که مدل حاضر بطور مستقیم از داده‌های ژئوفیزیکی و تکتونیکی مانند تغییر شکل گسل‌ها یا تنش‌های پوسته زمین استفاده نکرده است. این رویکرد به منظور ارزیابی کارایی مدل، صرفاً براساس داده‌های لرزه‌ای اتخاذ شده است. برای مقابله موثرتر با تغییرات بلندمدت ساختاری در پوسته زمین و افزایش دقت پیش‌بینی، پیشنهاد می‌شود در پژوهش‌های آینده، داده‌های ژئوفیزیکی مرتبط (نظیر داده‌های GPS، InSAR یا مدل‌های تنش منطقه‌ای)، به مدل اضافه شوند. این کار می‌تواند درک مدل از فرآیندهای فیزیکی مرتبط با زلزله را به شکل چشمگیری بهبود بخشد و به قابلیت تعمیم آن کمک کند. در ادامه مسیر، توصیه می‌شود ترکیب داده‌های لرزه‌ای با ویژگی‌های ژئوفیزیکی و ساخت مدل‌های تطبیقی برای مناطق مختلف ایران و جهان مورد توجه قرار گیرد. همچنین طراحی معماری‌های سبک‌تر و سریع‌تر که بتوانند عملکرد قابل قبول با منابع محدود ارائه دهند، می‌تواند گامی مهم در جهت پیاده‌سازی عملی سامانه‌های هشدار زلزله باشد.

منابع

- A. B. Asma and P. J. B. Aranas, "Forecasting Earthquake Using Machine Learning," *Journal of Artificial Intelligence, Machine Learning and Neural Network*, vol. 4, no. 3, pp. 31–40, Apr.–May 2024, DOI:10.55529/jaimlenn.43.31.40
- A. Bhatia, S. Pasari, and A. Mehta, "Earthquake forecasting using artificial neural networks," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, no. 5, pp. 823–827, 2018.
<https://doi.org/10.5194/isprs-archives-XLII-5-823-2018>

- A. Gupta and R. Kumar, "A REAL TIME PERFORMANCE BASED MODEL FOR PREDICTING THE EARTHQUAKE EARLY WARNING ALARMS USING A MACHINE LEARNING TECHNIQUE," IEEE Internet of Things Journal, vol. 8, no. 6, pp. 4567-4578, June 2021. DOI: <http://doi.org/10.26480/mjg.02.2024.115.120>
- B. Aslam, A. Zafar, U. Khalil, and U. Azam, "Seismic activity prediction of the northern part of Pakistan from novel machine learning technique," J. Seismol., vol. 25, pp. 639-652, 2021. DOI: [10.1007/s10950-021-09982-3](https://doi.org/10.1007/s10950-021-09982-3)
- H. Ebrahimian, F. Jalayer, B. Maleki Asayesh, S. Hainzl, and H. Zafarani, "Improvements to seismicity forecasting based on a Bayesian spatio-temporal ETAS model," Scientific Reports, vol. 12, Article no. 20970, 2022, <https://doi.org/10.1038/s41598-022-24080-1>
- I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA: MIT Press, 2016.
- J. P. Huang, X. A. Wang, Y. Zhao, C. Xin, and H. Xiang, "Large earthquake magnitude prediction in Taiwan based on deep learning neural network," Neural Network World, vol. 28, no. 2, pp. 149-160, 2018, DOI: [10.14311/NNW.2018.28.009](https://doi.org/10.14311/NNW.2018.28.009)
- J. Reyes, A. Morales-Esteban, and F. Martínez-Álvarez, "Neural networks to predict earthquakes in Chile," Appl. Soft Comput., vol. 13, no. 2, pp. 1314-1328, 2013. DOI: [10.1016/j.asoc.2012.10.014](https://doi.org/10.1016/j.asoc.2012.10.014)
- Keogh, E., Chu, S., Hart, D., & Pazzani, M. (2001). An online algorithm for segmenting time series. *Proceedings IEEE International Conference on Data Mining*, pp. 289-296. <https://doi.org/10.1109/ICDM.2001.989531>
- K. M. Asim, A. Idris, T. Iqbal, and F. Martínez-Álvarez, "Seismic indicators based earthquake predictor system using Genetic Programming and AdaBoost classification," Soil Dyn. Earthq. Eng., vol. 111, pp. 1-7, 2018. DOI: [10.1016/j.soildyn.2018.04.020](https://doi.org/10.1016/j.soildyn.2018.04.020)
- Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. In Proceedings of the 32nd International Conference on Machine Learning (ICML).
- M. Apriani, S. K. Wijaya, and Daryono, "Earthquake magnitude estimation based on machine learning: Application to earthquake early warning system," *Journal of Physics: Conference Series*, vol. 1951, no. 1, p. 012057, 2021. DOI: [10.1088/1742-6596/1951/1/012057](https://doi.org/10.1088/1742-6596/1951/1/012057)
- M. Asim, J. Taboada, A. García-Jerez, and E. Suárez-Rey, "Earthquake prediction model using support vector Programming and hybrid neural networks," PLOS ONE, vol. 13, no. 12, p. e0199004, 2018, doi:10.1371/journal.pone.0199004.
- M. Bhatia, T. A. Ahanger, and A. Manocha, "Artificial intelligence based real-time earthquake prediction," *Engineering Applications of Artificial Intelligence*, vol. 120, p. 105856, 2023. DOI: [10.1016/j.engappai.2023.105856](https://doi.org/10.1016/j.engappai.2023.105856)
- M. H. Al Banna, K. A. Taher, M. S. Kaiser, M. Mahmud, M. S. Rahman, A. S. Hosen, and G. H. Cho, "Application of artificial intelligence in predicting earthquakes: State-of-the-art and future challenges," IEEE Access, vol. 8, pp. 192880-192923, 2020. DOI: [10.1109/ACCESS.2020.3029859](https://doi.org/10.1109/ACCESS.2020.3029859)
- M. Merdasse, M. Hamdache, J. A. Peláez, J. Henares, and T. Medkour, "Earthquake Magnitude and Frequency Forecasting in Northeastern Algeria using Time Series Analysis," *Applied Sciences*, vol. 13, no. 3, p. 1566, 2023, DOI: [10.3390/app13031566](https://doi.org/10.3390/app13031566)
- M. T. Green and S. K. Patel, "Machine learning approaches for earthquake risk assessment: A comprehensive survey," IEEE Reviews in Geosciences, vol. 29, pp. 102-118, 2022. doi.org/10.1016/j.pdisas.2024.100398
- N. B. Jarah, A. H. H. Alasadi, and K. M. Hashim, "Earthquake prediction technique: a comparative study," IAES International Journal of Artificial Intelligence (IJ-AI), vol. 12, no. 3, pp. 1026-1032, 2023, DOI: [10.11591/ijai.v12.i3.pp1026-1032](https://doi.org/10.11591/ijai.v12.i3.pp1026-1032)
- Panakkat and H. Adeli, "Neural network models for earthquake magnitude prediction using multiple seismicity indicators," Int. J. Neural Syst., vol. 17, no. 1, pp. 13-33, 2007.
- S. Ommi and M. Hashemi, "Machine learning technique in the north Zagros earthquake prediction," *Applied Computing and Geosciences*, vol. 22, p. 100163, 2024. DOI: [10.1016/j.acags.2024.100163](https://doi.org/10.1016/j.acags.2024.100163)
- T. Perol, M. Gharbi, and M. A. Denolle, "Convolutional neural network for earthquake detection and location," *Science Advances*, vol. 4, no. 2, p. e1700578, Feb. 2018, DOI: [10.1126/sciadv.1700578](https://doi.org/10.1126/sciadv.1700578)
- T. Wang, J. D. Griffin, M. Brenna, D. Fletcher, J. Zeng, M. Stirling, P. W. Dillingham, and J. Kang, "Earthquake forecasting from paleoseismic records," *Nature Communications*, vol. 15, p. 1944, Mar. 2024, <https://doi.org/10.1038/s41467-024-46258-z>
- Z. Bao, J. Zhao, P. Huang, S. Yong, and X. Wang, "A deep learning-based electromagnetic signal for earthquake magnitude prediction," *Sensors*, vol. 21, no. 13, p. 4434, Jun. 2021. [doi: 10.3390/s21134434](https://doi.org/10.3390/s21134434)