GES	Journal of Geography and Environmental Studies, 13 (52), Winter 2024 https://sanad.iau.ir/journal/ges ISSN: 2008-7845 Doi: 10.71740/ges.2024.1130183

Research Paper
Received: 26 August 2024
Revised: 9 October 2024
Acceped: 11 December 2024

Investigating the Effect of Data Augmentation on the Intelligentization of Environmental Hazard Studies - Case Study: Real-Time Calculation of Earthquake Magnitudes in Early Warning Systems

Rezvan Esmaely[1], Rouholah Kimiaeifar[2*], Alireza Hajian[1], Khosro Soleimani[3], Maryam Hodhodi[1]

1. Departement of Physics, Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
2. Departement of Physics, Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran. (Corresponding Author)
E-Mail: r-kimiaefar@iaun.ac.ir
3. Departement of Mathematics, Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran.

Abstract

Natural hazards, including earthquakes, are serious challenges for human societies. These phenomena not only threaten human lives, but also cause significant economic and social losses. Considering that some of these risks occur suddenly and unpredictably, it is necessary that the evaluation and optimal management of these risks should be placed on the agenda of policymakers and crisis management officials. Recent advances and achievements in the field of intelligent algorithms have been able to reduce the human ability to reduce the effects caused by the occurrence. Among the examples of these achievements are Earthquake Early Warning Systems (EEWS), which have been proposed as an effective tool in reducing damages and human casualties, and can issue an early warning to the population affected by the event by calculating the basic parameters of the earthquake and provide the authorities with the necessary information. Considering that one of the basic pillars of using learning algorithms is the existence of sufficient training data, in cases such as earthquake data where the number of available samples is insufficient, data augmentation techniques are used. In this research, the impact of data augmentation in the calculation of the earthquake has been investigated in the real-time and based on the data of the strong motion network. Based on the analysis of the data of more than three thousand earthquakes, it has been shown that the use of data augmentation has improved the generalization performance of the trained network over test data by 37%.

Keywords: Intelligentization, earthquake magnitude, earthquake early warning system, powerful earth movement, data augmentation.

Citation: Esmaely, R.; Kimiaeifar, R.; Hajian, A.; Soleimani, Kh.; Hodhodi, M. (2024), Investigating the Effect of Data Augmentation on the Intelligentization of Environmental Hazard Studies - Case Study: Real-Time Calculation of Earthquake Magnitudes in Early Warning Systems, Journal of Geography and Environmental Studies, 13 (52), 6-17. Doi: 10.71740/ges.2024.1130183

Copyrights:
Copyright for this article is retained by the author (s), with publication rights granded to Journal of Geography and Environmental Studies. This is an open – acsses article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. **Publisher: Islamic Azad University of Najafabad**

Doi: 10.71740/ges.2024.1130183

مقاله پژوهشی

بررسی تاثیر داده‌افزایی در هوشمند سازی مطالعات مخاطرات محیطی- مطالعه موردی: محاسبه آنی بزرگا در سیستمهای هشدار سریع زلزله

رضوان اسماعیلی[۱]، روح‌اله کیمیایی‌فر[۱*]، علیرضا حاجیان[۱]، خسرو سلیمانی[۲]، مریم هدهدی[۱]

۱. گروه فیزیک، دانشکده مهندسی کامپیوتر، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران

۲. گروه ریاضی، دانشکده مهندسی کامپیوتر، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران

چکیده

مخاطرات طبیعی از جمله زمین‌لرزها، از چالش‌های جدی برای جوامع بشری به شمار می‌روند. این پدیده‌ها علاوه بر جانی، خسارات اقتصادی و اجتماعی قابل توجهی را نیز به دنبال دارند. با توجه به اینکه بعضی از این مخاطرات، به طور ناگهانی و غیرقابل پیش‌بینی رخ می‌دهند، لازم است که ارزیابی و مدیریت بهینه این مخاطرات در دستور کار سیاست گذاران و مسئولان مدیریت بحران قرار گیرد. پیشرفت‌ها و دستاوردهای اخیر در زمینه الگوریتم‌های هوشمند توانسته است توان بشر در کاهش آثار ناشی از وقوع را کاهش دهد. از نمونه این دستاوردها، سیستم‌های هشدار سریع زمین لرزه هستند که به عنوان ابزاری مؤثر در کاهش آسیب‌ها و تلفات انسانی مطرح شده و می‌توانند با محاسبه پارامترهای مبنایی زمین لرزه هشدار زودهنگامی را به جمعیت متاثر از رویداد صادر کرده اطلاعات مورد نیاز را در اختیار مسولان قرار دهند. با توجه به اینکه یکی از ارکان اساسی استفاده از الگوریتم‌های یادگیرنده، وجود داده‌های آموزشی کافی می‌باشد که بتواند شرایط متنوع قابل انتظار را پوشش دهد. این در حالی است که در مواردی مانند داده‌های مربوط به مطالعات زمین لرزه که تعداد نمونه‌های موجود در دسترس ناکافی است، تکنیک‌های داده‌افزایی استفاده می‌شوند. در این تحقیق، تاثیر داده‌افزایی در محاسبه بزرگای زمین لرزه به صورت آنی و بر اساس داده‌های جنبش نیرومند زمین مورد بررسی قرار گرفته است. بر اساس تحلیل‌های صورت گرفته بر روی داده‌های بیش از سه هزار زمین لرزه بوقوع پیوسته نشان داده شده است که استفاده از داده‌افزایی موجب بهبود عملکرد تعمیم‌دهی شبکه یادگیرنده به مقدار ۳۷ درصد شده است.

کلمات کلیدی: هوشممند سازی، بزرگای زمین لرزه، سیستم هشدار سریع زمین لرزه، جنبش نیرومند زمین، داده افزایی.

تاریخ ارسال: ۱۴۰۳/۰۶/۰۵

تاریخ بازنگری: ۱۴۰۳/۰۷/۱۸

تاریخ پذیرش: ۱۴۰۳/۰۹/۲۱

نویسنده مسئول: روح‌اله کیمیایی‌فر، گروه فیزیک، دانشکده مهندسی کامپیوتر، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران.r-kimiaefar@iaun.ac.ir

نحوه ارجاع به مقاله:

اسماعیلی، رضوان؛ کیمیایی‌فر، روح‌اله؛ حاجیان، علیرضا؛ سلیمانی، خسرو؛ هدهدی، مریم (۱۴۰۳)، بررسی تاثیر داده‌افزایی در هوشمند سازی مطالعات مخاطرات محیطی- مطالعه موردی: محاسبه آنی بزرگا در سیستمهای هشدار سریع زلزله، فصلنامه جغرافیا و مطالعات محیطی، ۱۳ (۵۲)، ۱۷-۶. Doi: 10.71740/ges.2024.1130183

Copyrights:

Copyright for this article is retained by the author (s), with publication rights granded to Journal of Geography and Environmental Studies. This is an open – acsses article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. **Publisher: Islamic Azad University of Najafabad**

مقدمه

مخاطرات طبیعی از جمله زلزله‌ها، همواره یکی از چالش‌های جدی برای جوامع بشری به شمار می‌روند. این پدیده‌ها نه تنها جان انسان‌ها را تهدید می کنند، بلکه خسارات اقتصادی و اجتماعی قابل توجهی را نیز به دنبال دارند. به طور خاص، با توجه به اینکه زلزله‌ها به عنوان به طور ناگهانی و غیرقابل پیش‌بینی رخ می‌دهند، لازم است که ارزیابی و مدیریت بهینه رویارویی با این پدیده طبیعی در دستور کار قرار گیرد. در این راستا، سیستم‌های هشدار سریع زلزله به عنوان ابزاری مؤثر در کاهش آسیب‌ها و تلفات انسانی مطرح می‌شوند (حاجیان و همکاران، ۲۰۲۳). این سیستم‌ها با استفاده از داده‌های زلزله‌شناسی و بکارگیری الگوریتم‌های محاسباتی و هوشمند، می‌توانند به طور زودهنگام وقوع زلزله را شناسایی کرده و اطلاعات لازم را به ساکنان مناطق در معرض خطر منتقل کنند. با این حال، کارایی این سیستم‌ها به شدت وابسته به کیفیت و کمیت داده‌های استفاده شده در مدل‌سازی و پیش‌بینی آن‌ها است (طباطبائی و همکاران، ۲۰۲۲). بنابراین، ضرورت داده‌افزایی در مطالعات مربوط به سیستم‌های هشدار سریع زلزله، به منظور ارتقاء دقت و قابل اعتماد بودن این سیستم‌ها، به یک نیاز اساسی تبدیل شده است. در این مقاله، به بررسی ابعاد مختلف مخاطرات طبیعی، اهمیت سیستم‌های هشدار سریع زلزله، و نقش کلیدی داده‌افزایی در افزایش اثربخشی این سیستم‌ها پرداخته خواهد شد.

پایش مخاطرات طبیعی یک فرایند کلیدی در مدیریت خطرات زیست‌محیطی است و فرآیندهای این حوزه، با استفاده از تحلیل‌های مبتنی بر شاخه‌های علوم مختلف و نیز تکنیک‌های متنوع در زمینه محاسبات به نتیجه می‌رسند. از طرفی با توجه به اینکه الگوریتم‌های هوشمند، به ویژه مدل‌های یادگیری ماشینی و هوش مصنوعی، توانایی تحلیل داده‌های بزرگ و شناسایی الگوهای پیچیده را داشته و به طور مؤثری در بهبود دقت پیش‌بینی‌ها و شناسایی نقاط بحران قابلیت‌های خود را نشان داده‌اند (یین و همکاران، ۲۰۲۰; ژن و همکاران، ۲۰۲۱; صمدی و همکاران، ۲۰۲۲). به صورت طبیعی، استفاده از روشهای مذکور در این حوزه حیاتی نیز کاملا توجیه شده است. یکی از تکنیک‌های نوظهور و توانمند حوزه یادگیری ماشین، تکنیک داده‌افزایی[1] است که به عنوان یکی از مؤلفه‌های اساسی در بعضی از تحلیل‌های مبتنی بر داده‌ها و در جهت تأمین داده‌های کافی و با کیفیت برای آموزش و بهینه‌سازی الگوریتم‌های هوشمند کمک می کند (چنگ و همکاران، ۲۰۱۹; اسماعیلی و همکاران، ۲۰۲۴) استفاده از این ابزار، منجر به افزایش دقت و قابلیت اطمینان مدل‌ها شده و می‌تواند به بهبود خروجی تحلیل‌های بیانجامد (لی و همکاران، ۲۰۲۱؛ گنزالس و همکاران، ۲۰۲۲). در این تحقیقات، سه مفهوم سیستم هشدار سریع زمین لرزه، شبکه یادگیری عمیق و داده‌افزایی به عنوان سه حلقه اساسی زنجیره یادگیری و استنتاجِ مورد نیاز در مسئله هشدار سریع زمین لرزه به صورت بلادرنگ در نظر گرفته شده و تاثیر داده‌افزایی در بهبود عملکرد شبکه یادگیرنده مورد بررسی واقع شده است. در ادامه مقدمه، نمونه‌هایی از استفاده از تکنیک داده‌افزایی در مطالعات حوزه مخاطرات طبیعی مورد بررسی واقع شده و پس از معرفی روش و داده‌های استفاده شده در تحقیق، آزمایشات صورت گرفته به صورت مبسوط شرح داده شده و نتایج و دستاوردهای تحقیق جمع بندی شده‌اند.

به صورت کلی، داده‌افزایی به مجموعه تکنیک‌هایی اطلاق می‌شود که با استفاده از داده‌ها یا نمونه‌های موجود، داده‌های جدید و متنوعی تولید می‌شود. این تکنیک‌ها عموماً در حوزه یادگیری ماشین و هوش مصنوعی کاربرد داشته و می‌توانند به بهبود تعمیم‌پذیری مدل‌ها کمک کنند. در اولین موارد استفاده از تکنیک داده‌افزایی و در حوزه مطالعات مخاطرات محیطی تکنیک‌های مورد استفاده بیشتر در زمینه ایجاد تغییرات در نمونه‌های تصاویر بوده است که شامل، کشیدگی، دوران، تغییر هیستوگرام و موارد مشابه بوده است اما به مرور، ضریب نفوذ تکنیکهای جدیدتر در سایر حوزه نیز افزایش یافته است. در ادامه، مواردی از تحقیقات

1. Data Augmentation

ارزشمند انتشار یافته در زمینه پایش مخاطرات طبیعی که به صورت خاص از تکنیکهای داده افزایی استفاده کرده‌اند به صورت اجمالی مرور شده‌اند.

در حوزه پایش آتش‌سوزی‌های جنگلی، فرناندز[1] و همکاران (۲۰۱۹) نشان دادند که بررسی الگوهای آتش‌سوزی با استفاده از داده‌های افزوده وضعیت خطر آتش‌سوزی را دقیق پیش بینی می‌کند. ژنگ و همکاران (۲۰۲۰) نیز در زمینه پیش‌بینی سیلاب‌ها با استفاده از تکنیک‌های یادگیری عمیق و داده‌افزایی تحقیق جامعی انجام داده و نشان دادند که استفاده از داده‌افزایی می‌تواند منجر به بهبود قابل توجه دقت مدل پیش‌بینی سیلاب‌ها شود. نتایج تحقیق ایشان نشان داد که با استفاده از داده‌های افزوده شده، دقت پیش‌بینی به میزان ۱۵ درصد افزایش یافته است. در یک تحقیق دیگر، ویگانو و همکاران (۲۰۲۱) به بررسی تأثیر تکنیک‌های داده‌افزایی در ارزیابی خطرات زلزله پرداخته و نشان دادند که با استفاده از شبیه‌سازی‌های مختلف و تکنیک‌های داده‌افزایی، می‌توان مدلی ایجاد کرد که قادر به پیش‌بینی دقیق‌تر و سریع‌تر تاثیرات زلزله بر زیرساخت‌ها است.

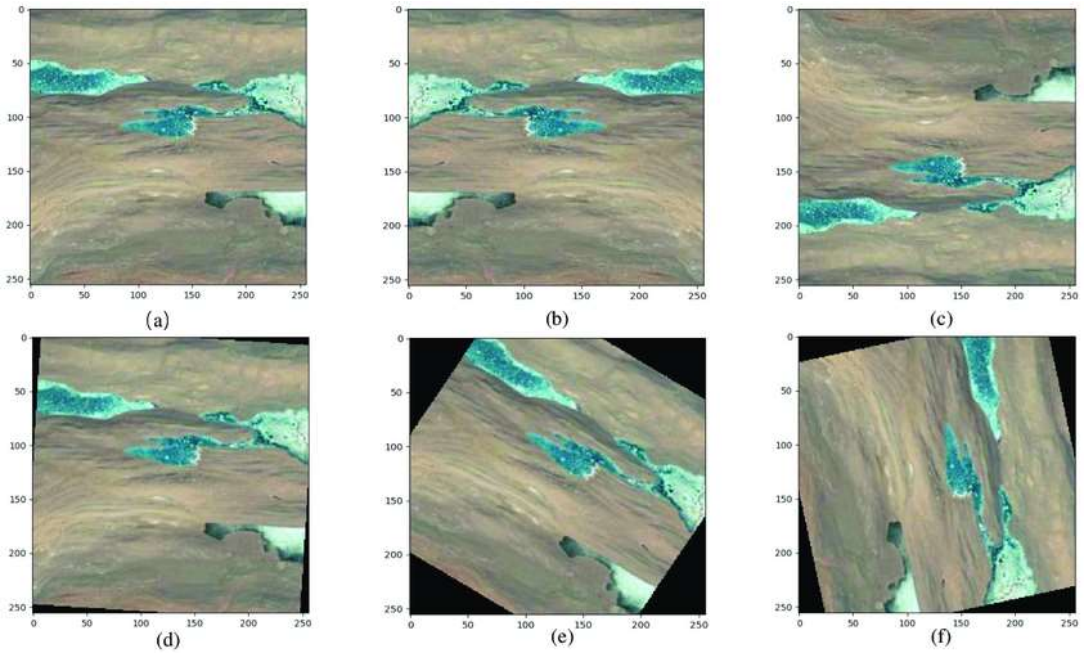
اگرچه در تحقیقات داخلی نیز مفهوم داده افزایی به وفور مورد استفاده واقع شده است، در حوزه پایش مخاطرات طبیعی، فراوانی تحقیقات محدود می‌باشد و در تحقیقات بررسی شده نیز، عمده کاربردها بر روی استفاده از تصاویر ماهواره‌ای متمرکز بوده‌اند. امین طوسی (۱۴۰۱) در مطالعه‌ای که به منظور تشخیص شعله‌های آتش، با روشی خلاقانه، تصاویر گردآوری شده که در خلال روز جمع آوری شده‌اند را به تصاویر از همان محل، در شرایط شبیه سازی شده شبانه، تبدیل کرده تا بتوانند نمونه‌های آموزشی بیشتری برای آموزش شبکه میها کنند. فرآیند مذکور با استفاده از یک شبکه مولد تخاصمی صورت گرفته است. در این تحقیق نتیجه گیری شده است که این فرآیند به صورت قابل توجه عملکرد شناسایی شعله آتش را بهبود می‌بخشد. همچنین، ایمانی (۱۴۰۳)، به منظور غلبه بر چالش کمبود نمونه‌های آموزشی مورد نیاز برای آموزش شبکه، از یک شبکه عصبی مولد متخاصمی برای داده افزایی نمونه‌های تصاویر ابر طیفی گردآوری شده توسط سامانه‌های سنجش از راه دور استفاده کرده چنین نتیجه گرفته است که روش ارائه شده نسبت به روشهای مرسوم کاراتر می‌باشد.

به صورت کلی و بر اساس داده‌ها و تحقیقات منتشر شده و موجود، می‌توان انتظار داشت که استفاده از داده‌افزایی منجر به بهبود عملکرد تعمیم دهی شبکه‌های یادگیرنده به مقدار ۱۰ تا ۲۵ شود. در شکل ۱، نمونه‌هایی از تکنیکهای متداول داده افزایی که بر روی یک تصویر ماهواره‌ای (تصویر اپتیکی به منظور شناسایی محدوده محتوای آب دریاچه صورت)، آورده شده است (ونگ و همکاران، ۲۰۲۰).

داده افزایی بر مبنای افراز دهکی داده‌ها

دو کاربرد اساسی داده‌افزایی در مباحث علوم زمین، غلبه بر مشکل عدم توازن داده‌ها و نیز ارتقا عملکرد الگوریتم یادگیری ماشین است که هر دو به نوعی در فضای داده‌های آموزشی دسته‌بندی می‌شوند. بنابراین، طبیعی است که الگوریتم داده‌افزایی بر مبنای داده‌های آموزشی طراحی شود. در تحقیق اباذری و همکاران (۲۰۲۴) نشان داده شد که استفاده از داده‌افزایی بر مبنای افراز داده‌های آموزشی به چهار دهک مستقل، مسئله عدم توازن داده‌ها در الگوریتم یادگیرنده را به نحو چشمگیری مرتفع می‌کند. همچنین اسماعیلی و همکاران (۲۰۲۴) نشان دادند که می‌توان داده‌های مدل یادگیرنده را با همجوشی[2] داده‌های آموزشی که به دو گروه خوشه‌بندی شده بودند، تبدیل به داده‌هایی نمود که آموزش و تعمیم شبکه بر روی آنها عملکرد بهتری خواهد داشت. در اینجا نیز داده‌افزایی با دیدگاهی مشابه توسعه داده شده است.

1. Fernandes
2. Fusion



شکل (۱): نمونه‌هایی از تصاویر داده‌افزایی شده از یک تصویر ورودی در قالب اعمال فیلترهای مختلف بر یک تصویر a) تصویر اصلی، b) تصویر برعکس شده افقی، c) تصویر برعکس شده عمودی، d تا e) تصویر چرخانده شده در زوایای مختلف (ونگ و همکاران، ۲۰۲۰).

تنوع داده‌های ورودی در مباحث مرتبط با علوم زمین و نیز پیچیگیده‌های ذاتی مدل‌های توصیف کننده باعث می‌شود که ذاتا، داده‌های آموزشی ورودی به شبکه یادگیرنده متنوع باشند و به دنبال آن، شبکه برای بهینه سازی وزن‌ها مجبور خواهد بود که با تمرکز بر یادگیری با عملکرد بهتر، تمرکز خود بر روی بعضی از نمونه‌ها را از دست بدهد. نمونه بارز این چالش، داده‌های نامتعادل هستند. شبکه‌های عصبی در مواجه با داده‌های کم شمار، همواره آن‌ها را به عنوان رویداد نادر کنار گذاشته و سعی در بهبود عملکرد با تمرکز با گروه داده‌های پرشمار می‌کنند. الگوریتم ارائه شده در این تحقیق به شرح زیر سعی در غلبه بر چالش مذکور دارد:

قدم اول: کلیه داده به عنوان یک کل در n گروه خوشه‌بندی می‌شوند. وجه تمایز این خوشه‌بندی با خوشه بندی‌های مرسوم این است که تعداد عناصر خوشه‌های مذکور با یکدیگر برابر می‌باشد. به این گونه خوشه بندی‌های مقید، خوشه‌بندی هم‌اندازه (Equal-Sizes Clustering) اطلاق می‌شود. در این تحقیق، فرآیند ذکر شده بر مبنای روش ارائه شده توسط لین و همکاران (۲۰۲۲) صورت می‌پذیرد.

قدم دوم، افراز داده‌های خوشه‌بندی شده به صورت یکسان از تمام خوشه‌ها به سه دسته آموزش، اعتبار سنجی و آزمون صورت می‌پذیرد. به عنوان نمونه، در اینجا ۷۰، ۲۰ و ۱۰ درصد از تعداد نمونه‌های هر خوشه به صورت تصادفی به عنوان داده‌های آموزشی، اعتبار سنجی و آزمون در نظر گرفته شدند.

قدم سوم، از داده‌های آموزشی متناظر با هر خوشه، افراز دهکی[1] در راستای هر ویژگی صورت گرفته و برای هر دهک، نمونه‌های جدید با اِعمال تغییرات متناسب با انحراف از معیار دهک مذکور در راستای آن ویژگی تولید می‌شود. مثلا، نمونه‌های جدید با پنج درصد از انحراف معیار کلی نمونه‌های متعلق به هر دهک، در راستای هر ویژگی تولید شده در حالی که مقدار هدف[2] بدون تغییر باقی می‌ماند.

قدم چهارم، با توجه به آمده شدن داده‌های آموزشی، شبکه با معماری مد نظر آموزش یافته سپس، با داده‌های آزمون اعتبارسنجی می‌شود. در صورت رسیدن به حد مورد قبول از عملکرد، فرآیند به پایان رسیده و در غیر اینصورت، مجدد پارامترهای اصلی مانند

1. Decile
2. Target

میان انحراف معیار تنظیم شده و شبکه مجدد آموزش می‌ابد. این فرآیند تا رسیدن به حد مطلوب از عملکرد ادامه پیدا خواهد کرد. بدین ترتیب، نمونه‌های جدید با دو قید عضویت در خوشه یکسان که بر مبنای تحلیل تمام ویژگی‌ها به صورت توامان بوده و سپس بر مبنای تحلیل آماری دهک‌بندی در راستای هر ویژگی تولید می‌شوند. فلوچارت روش ارائه شده در شکل ۲ به تصویر کشیده شده است.

داده‌ها و روشها

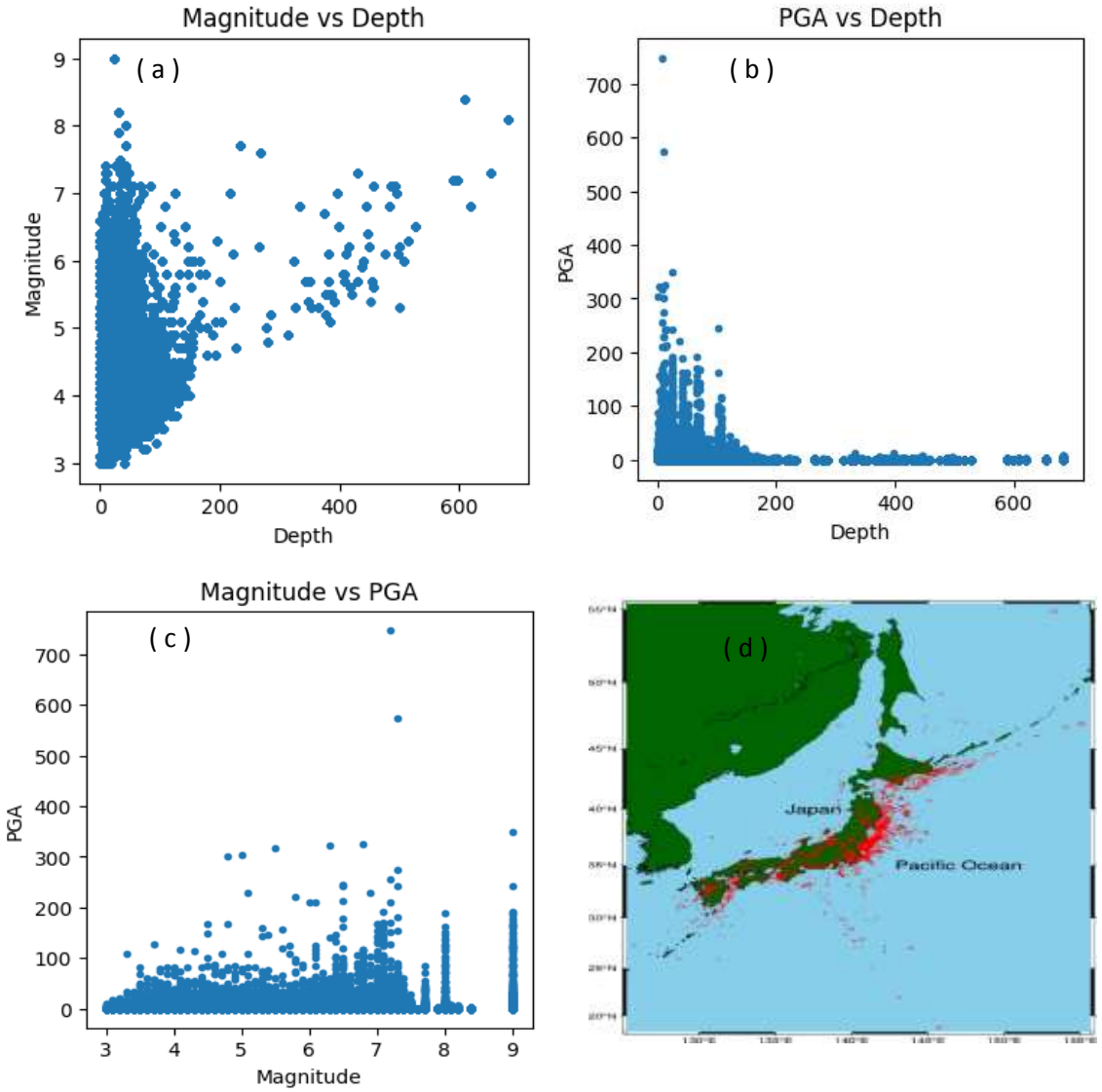
مجموعه داده‌های اصلی این پژوهش مربوط به ۳۰۳۶ زلزله ثبت شده توسط یک شبکه لرزه نگاری جنبش نیرومند زمین[1] در ژاپن است (شکل ۳-d). تعداد رکوردهای متناظر با رویدادهای مذکور ۱۰۳۵۸۹ عدد بوده که در بازه بزرگای ۳ تا ۹ واحد و عمق کانونی تا ۶۸۲ کیلومتری از سطح زمین توزیع شده‌اند. اطلاعات آماری مجموعه داده اصلی در جدول ۱ آورده شده است. همچنین، توزیع متقاطع داده‌ها بر مبنای ویژگی‌های اصلی، در شکل ۳ نیز آورده شده است. لازم به توضیح است که این اقدام در جهت کنار گذاشته شدن داده‌هایی است که با توجه به توزیع کنونی داده‌ها، در فواصل بسیار دورتری از میانگین آنها قرار دارند. رسم متقاطع داده‌ها بر مبنای پارامترهایی توصیف گر آنها کمک میکند که اتخاذ تصمیم در خصوص تعیین حدود، به صورت بصری آسانتر و سریعتر صورت پذیرد. بررسی شکل ۳ نشان می‌دهد که تنها تعداد اندکی از رویدادها در عمق بیشتر از ۱۸۰ کیلومتر قرار دارند، لذا به منظور اجتناب از تاثیر منفی احتمالی این رویدادها در آموزش شبکه عصبی به عنوان رویداد نادر (کینگ و ژنگ، ۲۰۰۱)، به عنوان اولین قدم در پالایش داده‌ها، رویدادها با عمق بیشتر از ۱۸۰ کیلومتر کنار گذاشته شدند. در قدم بعدی، با توجه به نتایج آزمایش‌های اولیه که پیش از تعیین قطعی ابرپارامترهای مربوط به مدل یادگیرنده و سایر متغیرهای آزمایشات، رویدادهایی که بزرگای بیشتر از ۵ واحد داشته‌اند، کنارگذاشته شدند. لازم به توضیح است که برای این رویدادها در آزمایشات اولیه، میزان خطای محاسبه شده برای گزارشات میانگینی بالاتر از یک واحد داشته و ترجیح داده شد که رفع این نقیصه برای مطالعات بعدی در نظر گرفته شود.

شکل (۲): فلوچارت روش ارائه شده جهت داده افزایی مقید بر مبنای عملکرد

1. https://www.kyoshin.bosai.go.jp

جدول (۱): توصیف آماری مجموعه داده‌های اصلی

	عرض جغرافیایی رومرکز	طول جغرافیایی رومرکز	میانگین سرعت موج برشی در ۳۰ متر	عمق کانونی	بیشینه شتاب زمین	بزرگای زمین لرزه
تعداد داده	103589	103589	103589	103589	103589	103589
میانگین	38.00	62.84	476.13	49.80	52.73	4.86
انحراف از معیار	2.84	22.23	228.03	68.49	3344.95	0.97
کمینه	31.09	25.00	111.11	0.00	0.00	3.00
بیشینه	45.49	100.00	2100.00	682.00	1633075.00	9.00



شکل (۳): نمایش متقاطع سه ویژگی عمق، بزرگا و PGA برای داده‌های اصلی، a) بزرگا بر حسب عمق کانونی b) بیشینه شتاب بر حسب عمق کانونی c) بزرگا بر حسب بیشینه شتاب و d) موقعیت رومرکز رویدادهای استفاده شده در تحقیق.

در نهایت، برای مقادیر بیشینه شتاب زمین[1]، که یک پارامتر مبنایی در طراحی سازه‌های منطقه بوده و در تدوین استانداردهای ساخت و ساز نیز مورد استفاده واقع می‌شود، محدودیت خاصی در نظر گرفته نشده و تمامی گزارشات برای مقادیر مذکور در تحقیق، مشارکت داده شد.

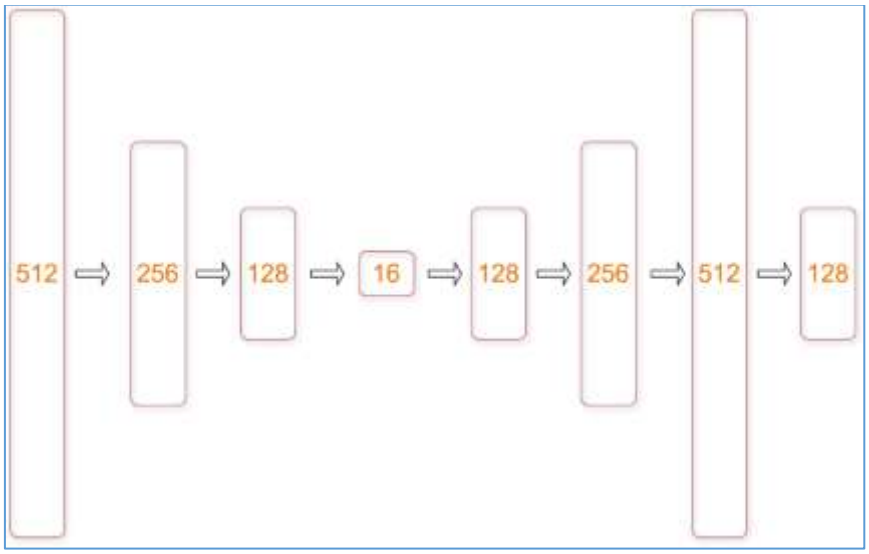
آزمایشات

یک شبکه یادگیری عمیق با معماری ترتیبی[2] با ۸ لایه تماما متصل[3] برای مسئله رگرسیونِ محاسبه بزرگای زمین لرزه در نظر گرفته شد. تابع فعالسازی Relu در تمامی لایه‌ها شده و از پارامتر میزان خسران[4] برای محاسبه عملکرد مجموعه داده‌های اعتبار سنجی استفاده شد. تعداد اپکهای[5] استفاده شده برابر ۵۰ اپک در نظر گرفته شده و متعاقبا، بهترین عملکرد در تعمیم وزن‌ها بر روی داده‌های اعتبارسنجی به عنوان بهینه‌ترین مدل وزن گرفته شده مورد استفاده واقع شد.

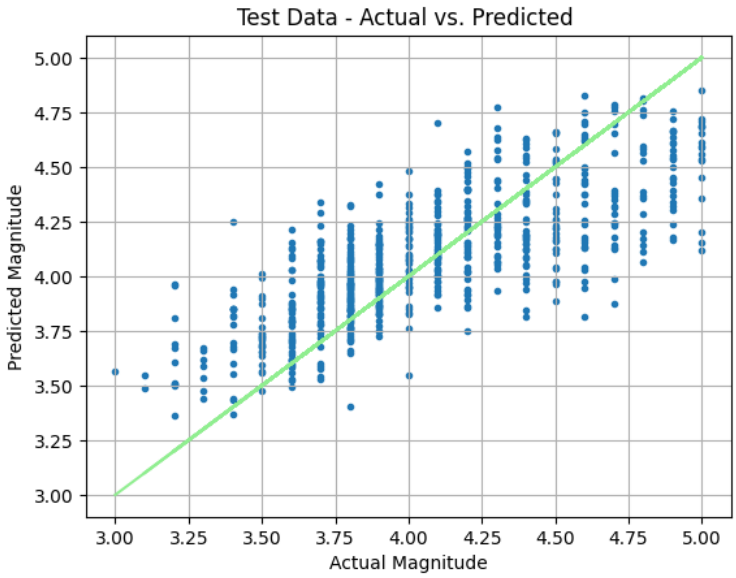
ورودیهای شبکه یادگیرنده عبارت بودند از مختصات جغرافیایی ایستگاه، سرعت موج برشی در ۳۰ متر اول سطح در موقعیت ایستگاه، و در نهایت مقادیر PGA برای دو مولفه افقی متعامد شمالی-جنوبی و شرقی-غربی. بجز آخرین مورد، بقیه ورودیهای برای هر ایستگاه در خلال استفاده از یک تنظیمات و دستگاه ثابت خواهد بود. در واقع متغیر مسئله در این جا فقط مقادیر PGA برای هر رویداد در هر ایستگاه از شبکه می‌باشد. چنانچه پیشتر نیز اشاره شد، ورودی‌های مذکور به منظور مدلسازی مقادیر بزرگای زلزله استفاده شدند. در شکل ۴ شماتیک معماری شبکه یادگیری عمیق استفاده شده آورده شده است. در خصوص معماری در نظر گرفته شده لازم به توضیح است که با توجه به میزان حساسیت PGA به تغییرات سنگ شناسی لایه‌های نزدیک به سطح زمین، معماری مدل عمیق مذکور با تاکید بر معماری شبکه‌های خودرمزنگار[6] در نظر گرفته شد تا از تاثیر نوسانات موجود در سطح که در اینجا به عنوان نویز در داده‌ها تلقی شده‌اند کاسته شود (حاجیان و همکاران، ۲۰۲۳).

با توجه به پیش پردازش‌های توضیح داده شده در قسمت قبل، در ادامه، کلیه داده‌ها به پنج خوشه هم اندازه تجزیه شده و سپس، برای تعیین داده‌های آموزشی، اعتبار سنجی و آزمون، به ترتیب مقادیر ۷۰، ۲۰ و ۱۰ درصد از هر خوشه در نظر گرفته شد. سپس، با توجه به ضرورت مقایسه عملکرد روش ارائه شده، داده‌های مذکور بدون هر گونه عملیات دیگر مانند نرمالایزسازی و داده‌افزایی، برای آموزش مدل ذکر شده استفاده شده و پس اتمام فرآیند آموزش، نتیجه تعمیم شبکه وزن گیری شده بر روی داده آزمون بدست آورده شد که نتیجه آن در شکل ۵ نمایش داده شده است. چنانچه مشاهده می‌شود، اختلاف بین مقادیر واقعی و مقادیر مدل شده تا یک واحد بزرگا خطا داشته است.

1. Peak Ground Acceleration, PGA
2. Sequential
3. Fully Connected
4. Loss
5. Epoch
6. Auto Encoder

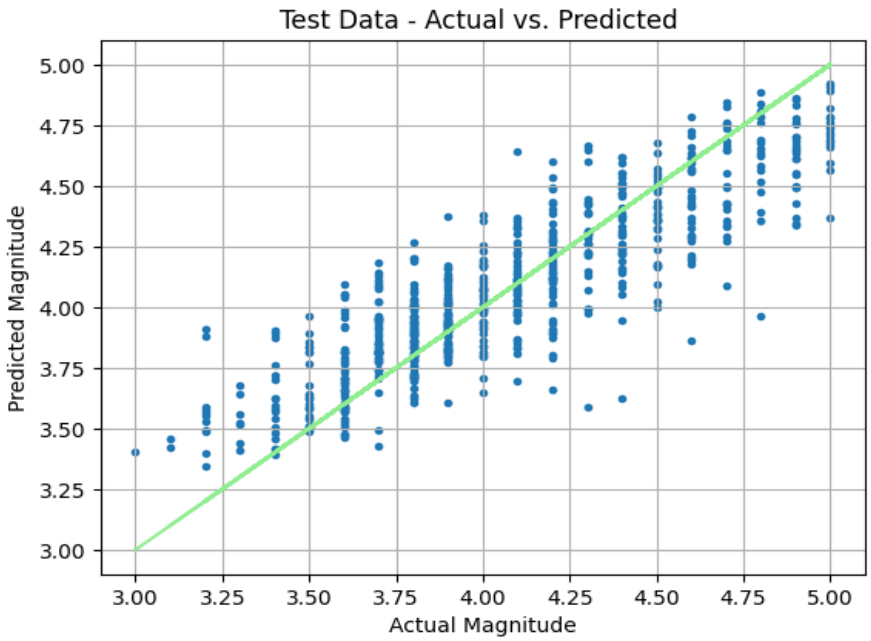


شکل (۴): معماری شبکه عمیق استفاده شده در تحقیق که از معماری مدلهای خودرمزنگار برگرفته شده است.



شکل (۵): رگرسیون مقادیر واقعی و خروجی تعمیم شبکه آموزش یافته برای محاسبه بزرگای زمین لرزه بدون اعمال الگوریتم داده افزایی مقید.

در این مرحله، داده‌های آموزشی در راستای هر ویژگی به دهک‌های جداگانه تقسیم بندی شده و فرآیند داده‌افزایی، مطابق توضیحات داده شده در قسمت روش، انجام گردید. بدین منظور، برای هر نمونه از هر دهک، میزان انحراف از معیار دو درصدی برای تولید نمونه‌های جدید در نظر گرفته شد. تعداد نمونه‌های تولید شده در هر دهک ۵۰۰ نمونه و همانگونه که اشاره شد، برای مقادیر هدف، مقادیر بزرگای گزارش شده بدون تغییر در نظر گرفته شد. پس از آموزش شبکه عصبی با معماری پیشین و تنها با یکبار تکرار فرآیند آموزش، خروجی تعمیم آن بر روی داده‌های آزمون محاسبه شد که در شکل ۶، خروجی رگرسیون آن، آورده شده است. بررسی کمی خروجی‌های دو آزمون تشریح شده نشان از بهبود ۳۷ درصدی در مدلسازی داده‌ها می‌دهد، چنانچه ماکزیمم میزان خطا ۰.۶۳ واحد بزرگا بوده است. لازم به توضیح است که در نظر گرفتن یک تکرار در فرآیند آموزش به خاطر قابلیت انجام مقایسه و بررسی عملکرد روش ارائه شده در دو حالت بوده است.



شکل (۶): رگرسیون مقادیر واقعی و خروجی تعمیم شبکه آموزش یافته برای محاسبه بزرگای زمین لرزه بعد از داده افزایی مقیدِ پیش از آموزش شبکه، با یکبار تکرار فرآیند آموزش

نتیجه گیری

پایش موثر مخاطرات طبیعی مستلزم بکارگیری سیستم‌های سخت‌افزاری و نرم‌افزاری و نیز الگوریتم‌های کارآمدی است که بتوانند فرآیندهای مرتبط مانند سیستمهای خودکارساز، پیش‌بینی کننده را میسر سازند. در این تحقیق سیستم‌های هشدار زودهنگام زمین لرزه، که کارایی خود را در کاهش تلفات ناشی از وقوع این مخاطره نشان داده‌اند، به عنوان یک مورد مطالعاتی جهت بررسی کارآمدی تاثیر داده افزایی در هوشمند سازی پایش خطر زمین لرزه درنظر گرفته شد. یکی از لوازمات پیاده سازی سیستمهای هشدار سریع زمین لرزه، محاسبات بلادرنگ یا نزدیک به بلادرنگ پارامترهای مبنایی زمین لرزه، به خصوص، بزرگای آن می‌باشد. این مهم از آن جهت اهمیت دارد که میزان بزرگا به صورت مستقیم می‌تواند تخمینی از قدرت جنبش نیرومند زمین و به دنبال آن، تلفات جانی و مالی در مناطق پیرامونی زمین لرزه ارائه کند. درفرآیندهای مرسوم محاسبه بزرگا، از دانش و تخصص کارشناس خبِره در شبکه لرزه نگاری استفاده می‌شود و به صورت طبیعی، استفاده از این دانش علاوه بر زمان‌بر بودن، وابسته به میزان تجربه و دانش و نیز سلیقه نیروی متخصص می‌باشد. با تاکید بر ارائه راه‌حلی مستقل، تحقیقات بسیاری در زمینه خودکارسازی فرآیندهای مرتبط با این مخاطره طبیعی صورت گرفته است. در این تحقیق نیز، با در نظر گرفتن ملاحظات مذکور، به منظور افزایش دقت خروجی تحلیل، روشی ارائه شده است که بتواند با داشتن اطلاعات ایستگاه‌های اطراف کانون زمین لرزه و نیز بیشینه شتاب زمین رکورد شده در ایستگاههای پیرامونی، بزرگای رویداد را ارائه کند.

با توجه به آزمایشات صورت گرفته، موارد زیر به عنوان نتایج این تحقیق قابل استنتاج می‌باشد:

- با توجه به پیچیدگی ذاتی رویداد زمین لرزه و نیز وابستگی شدید پارامترهای جنبش نیرومند زمین در سایت اندازه‌گیری، ضروری است که پایگاه داده استفاده شده در مدل یادگیرنده تقویت شود. با توجه به عدم در دسترس بودن داده‌های واقعی کافی، استفاده از داده‌افزایی به عنوان یک راه‌حل، قابل بررسی می‌باشد. این نتیجه با توجه به مقایسه خروجی شبکه یادگیری عمیق در دو حالت با و بدون داده‌های داده افزایی شده بدست آمده است.

- نتیجه بند قبلی، در یک شرایط برابر از نظر داده‌های ورودی و نیز مدل عمیق یادگیرنده، بدست آمده و خروجی بهینه شده با توجه به عملکرد بهتر در تعمیم شبکه آموزش یافته بر روی داده‌های آزمون بوده است، در نتیجه، می‌توان انتظار داشت که پیش بینی‌های آتی برای رویدادهای بعد از آموزش شبکه دقیق‌تر از حالتی است که شبکه بدون داده‌های شبیه سازی شده آموزش یابند.
- به عنوان یکی از پارامترهای مبنایی زلزله و در مورد خاص بزرگای زمین لرزه، بیشترین میزان خطای مورد انتظار با توجه به پایگاه داده استفاده شده در حد ۰.۶۳ واحد بزرگا و برای بیش از ده هزار داده استفاده شده در مجموعه داده‌های آزمون بوده است.

ادامه این تحقیق می‌تواند شامل موارد ذیل باشد:

- تحقیق در خصوص استفاده از سایر معماری‌های متدوال عمیق مانند شبکه‌های عصبی بازگشتی مانند شبکه حافظه بلند-کوتاه مدت[1]، مدلهای انتقالی می‌تواند مد نظر قرار گیرد تا عملکرد یادگیری-تعمیم شبکه ارتقا یابد.
- بررسی راهکارهای مقید سازی تعداد نمونه‌های مورد نیاز جهت اجتناب از تحمیل بار محاسباتی اضافی به شبکه عصبی در فرآیند آموزش می‌تواند در شرایطی که پیاده سازی الگوریتم در سیستمهای محاسباتی کوچکتر مد نظر است، مورد بررسی واقع شود.
- استفاده از الگوریتم ارائه شده در سایر حوزه‌ها و مسائل مرتبط با پایش مخاطرات که می‌تواند شامد محاسبه کلیه پارامترهای مرکز زلزله نیز باشد پیشنهاد می‌گردد.
- با توجه به پیشینه تحقیقات بررسی شده در زمینه تولید داده‌های شبیه سازی شده، استفاده از شبکه یادگیری مولد تخاصمی به منظور داده افزایی داده‌های زلزله پیشنهاد می‌شود.

منابع

امین طوسی، محمود (۱۴۰۱). انتقال سبک برای افزایش داده‌های آموزشی شبکه‌های کانولوشنی در شناسایی شعلهٔ آتش. *هوش محاسباتی در مهندسی برق*. ۱۳ (۴)، ۱۱۴-۹۷. doi: 10.22108/isee.2021.124044.1490

ایمانی، مریم (۱۴۰۳). داده‌افزایی مبتنی بر شبکه‌ی مولد رقابتی برای آشکارسازی اهداف تصاویر ابر طیفی با استفاده از شبکه‌های عصبی کانولوشن. *مجله بینایی ماشین و پردازش تصویر. مجله بینایی ماشین و پردازش تصویر*. ۱۳-۲۴. https://jmvip.sinaweb.net/article_207014.html.

Hajian, A.; Nunnari, G. & Kimiaefar, R. (2023). Intelligent methods and motivations to use in volcanology and seismology. *Intelligent Methods with Applications in Volcanology and Seismology*. Cham: Springer International Publishing, pp. 1–17. doi: 10.1007/978-3-031-15432-4_1.

Cheng, Y.; Wang, X. & Zhao, H. (2019). Data augmentation techniques for improving disaster prediction models. *International Journal of Disaster Risk Reduction*. 34, 380-393.

Esmaeili, R. et al (2024). Performance enhancement of deep neural network using fusional data assimilation and divide-and-conquer approach; case study: Earthquake magnitude calculation. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-024-10002-x.

Fernandes, P. M.; Silva, J. S. & Amraoui, M. (2019). Wildfire risk assessment using machine learning and data augmentation techniques. *Forest Ecology and Management* 432, 43-55.

Gonzalez, R.; Castro, J. & Vargas, A. (2022). Integrating intelligent algorithms with data augmentation for real-time disaster risk assessment. *Computers & Geosciences* 148, 104694.

King, G., & Zeng, L. (2001). Logistic Regression in Rare Events Data. *Political Analysis*. 9 (2), 137-163. DOI:10.1093/oxfordjournals.pan.a004868.

Li, J.; Liu, Y. & Wang, L. (2021). Enhancing disaster response with artificial intelligence: Opportunities and challenges. *Disaster Science and Management*. 30 (2), 160-176.

Lin, Y. et al (2022). Generating clusters of similar sizes by constrained balanced clustering. *Applied Intelligence*. 52, 5273–5289. https://doi.org/10.1007/s10489-021-02682-y.

1. Long-Short Term Memory, LSTM

Samadi, H.; Kimiaefar, R. & Hajian, A. (2022). Fast earthquake relocation using ANFIS neuro-fuzzy network trained based on the double difference method. *Scientific Quarterly Journal of Geosciences*. 32 (3), 93-102. doi: 10.22071/gsj.2022.296260.1922.

Tabatabae, S.M. et al (2022). Prediction of the peak ground acceleration for Zagros earthquakes using ANFIS and data partitioning approach. *Journal of Geography and Environmental Studies*. 42 (11), 92-104.

Viganò, D.; Gabriele, S. & Dell'Arciprete, P. (2021). Earthquake risk assessment using augmented data. *Earthquake Engineering and Structural Dynamics*. 50 (7), 1805-1820.

Yin, Z.; Zhang, Y. & Zhang, J. (2020). Machine learning in the assessment of natural disaster risks: A review. *Environmental Monitoring and Assessment*. 192 (12), 786.

Wang, Z. et al (2020). MSLWENet: A novel deep learning network for lake water body extraction of google remote sensing images. *Remote Sensing*. 12, 4140.

Zhan, Q.; Miao, L. & Xiao, J. (2021). Intelligent algorithms for natural disaster prediction and management: Current status and future prospects. *Journal of Natural Hazards*. 105 (1), 123-142.

Zhang, X.; Xu, Y. & Chen, Y. (2020). Deep learning for flood prediction: A review. *Natural Hazards*. 105 (2), 123-146.