



مقایسه الگوریتم‌های ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک در پهنه‌بندی خطر زمین‌لغزش در جاده مهاباد سردشت

تاریخ پذیرش: ۱۴۰۱/۱۰/۱۱

تاریخ دریافت: ۱۴۰۱/۰۹/۱۲

جواد سدیدي*^۱، رضا مالکی^۲

۱- دکترای تخصصی، دانشیار، گروه آموزشی سنجش از دور و سیستم اطلاعاتی جغرافیایی، دانشکده علوم جغرافیایی، دانشگاه خوارزمی
۲- دانشجوی کارشناسی ارشد، دانشکده علوم جغرافیایی، دانشگاه خوارزمی

چکیده

در تحقیق حاضر به ارزیابی و پهنه‌بندی خطر زمین‌لغزش در محدوده مهاباد تا سردشت با کاربرد مدل‌های تحلیلی مختلف از قبیل ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک پرداخته شده است. روش انجام تحقیق حاضر مبتنی بر روش توصیفی - تحلیلی و کاربرد مقایسه‌ای صحت الگوریتم‌های ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک بوده است. بر این اساس از داده‌ها و معیارهای مختلف محیطی در فرآیند تجزیه و تحلیل استفاده گردیده است. ابتدا بر اساس تعیین نقاط نمونه، سه مدل مذکور به منظور تهیه نقشه پهنه‌بندی زمین‌لغزش به اجرا درآمده است و سپس بر اساس نتایج، اقدام به ارزیابی و اعتبارسنجی نتایج مدل‌های مورد استفاده شده است. نتایج پهنه‌بندی زمین‌لغزش محدوده مورد مطالعه حاکی از این بوده است که به‌طور کلی نیمه جنوبی منطقه به دلیل تأثیر عواملی از قبیل ساختارهای متراکم گسلی، شیب بالاتر و تراکم بیشتر آبراهه از پتانسیل بالاتری نسبت به نیمه شمالی آن برخوردار است و بر اساس ماشین بردار پشتیبان ۷۱/۰۴ درصد، بر اساس جنگل تصادفی ۵۳/۴۴ درصد و بر اساس رگرسیون لجستیک ۷۷/۳۹ درصد از مجموع وسعت منطقه دارای حساسیت لغزشی متوسط به بالا بوده است. ارزیابی دقت حاصله برای الگوریتم‌ها بر اساس منحنی ROC چنین مشخص نموده است که ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک به ترتیب مقدار صحت ۰/۷۶، ۰/۸۷ و ۰/۸۴ را به خود اختصاص داده‌اند و از این نظر الگوریتم جنگل تصادفی بهترین دقت را ارائه کرده است. همچنین شاخص Precision - Recall نیز به ترتیب برابر با ۰/۸۰۹، ۰/۸۷۳ و ۰/۸۴۴ به دست آمده است که بیانگر دقت بالاتر الگوریتم جنگل تصادفی نسبت به دو الگوریتم دیگر در زمینه پهنه‌بندی پتانسیل خطر وقوع زمین‌لغزش در مسیر مهاباد - سردشت می‌باشد.

کلمات کلیدی: زمین‌لغزش، یادگیری ماشین، جنگل تصادفی، ماشین بردار پشتیبان، مهاباد - سردشت.

۱- مقدمه

زمین‌لغزش‌ها با توجه به عدم قطعیت زمانی و فضایی خود اغلب به سختی قابل پیش‌بینی هستند و هر فناوری یا تکنیکی که بتواند امکان پیش‌بینی دقیق‌تر از این مخاطره را فراهم نماید از اهمیت زیادی برخوردار است (۲). تهیه نقشه حساسیت زمین‌لغزش در دهه‌های اولیه به دلیل تجربی بودن رویکردهای سنتی و اعتبار پایین نتایج آن‌ها فرآیندی نسبتاً دشوار به شمار می‌رفت اما امروزه توسعه فناوری‌های کامپیوتری، داده‌ها و ابزارهای سنجش از دور (RS) و سیستم اطلاعات جغرافیایی (GIS)، دستیابی به نقشه‌های دقیق و جامع از پتانسیل زمین‌لغزش را تسهیل نموده است. در راستای این توسعه تا به امروز رویکردها و تکنیک‌های متعددی در زمینه پهنه‌بندی زمین‌لغزش به کار گرفته شده است (۲۸). در سالیان اخیر تکنیک‌های یادگیری ماشین (Machine Learning)، هوش مصنوعی و یادگیری عمیق به‌منظور پردازش مجموعه داده‌های فضایی در زمینه پهنه‌بندی خطر زمین‌لغزش مورد توجه بسیاری قرار گرفته‌اند. تکنیک‌های مذکور در مقایسه با سایر رویکردهای موجود همچون روش‌های آماری یا فیزیکی دارای ظرفیت پیش‌بینی‌کنندگی بالاتری بوده (۲۰) و شواهد مطالعاتی نیز نشان می‌دهد که کاربرد هر یک از الگوریتم‌های پیشرفته در این راستا می‌تواند تأثیر زیادی در ارتقای کیفیت ارزیابی خطر زمین‌لغزش و تعیین نقاط قوت و ضعف هر یک داشته باشد (۲۶). امروزه با توجه به کارایی و قابلیت‌های تکنیک‌های یادگیری ماشین توجهات جهانی در عرصه ارزیابی زمین‌لغزش به آن‌ها معطوف گردیده است (۱۵). تکنیک‌های یادگیری ماشین به لحاظ به‌کارگیری الگوریتم‌های تخمین توزیع، ماهیت داده‌محور و تکرار بالای فرایند مدل‌سازی قابلیت بالایی در شناسایی رفتار وقوع پدیده زمین‌لغزش داشته و در مطالعات مختلفی مزیت خود را نسبت به مدل‌های آماری دومتغیره و چندمتغیره به اثبات رسانده‌اند (۱۱). تکنیک‌های یادگیری ماشین به صورت مجزا، گروهی یا تلفیقی همراه با سنجش از دور و سیستم اطلاعات جغرافیایی برای مطالعات پهنه‌بندی زمین‌لغزش استفاده می‌شوند. بر این اساس شناسایی و انتخاب بهترین الگوریتم از میان تکنیک‌های یادگیری ماشین برای پهنه‌بندی خطر زمین‌لغزش در یک ناحیه موجب می‌شود که اطمینان بالاتری به نتایج در راستای پیش‌گیری یا بهبود شرایط صورت گرفته و اقدامات آتی به درستی به انجام برسد (۱۳).

رویکرد یادگیری ماشین دربرگیرنده الگوریتم‌های مختلفی به منظور پهنه‌بندی خطر زمین‌لغزش است. در این ارتباط تا به امروز الگوریتم‌های داده‌محور همچون ماشین بردار پشتیبان (Support Vector Machine)، جنگل تصادفی (Random Forest)، و رگرسیون لجستیک (Logistic Regression) با توجه به مزایای آن‌ها نسبت به روش‌های سنتی به دفعات در راستای مطالعه زمین‌لغزش در مناطق مختلف به کار گرفته شده‌اند (۳). الگوریتم ماشین بردار پشتیبان که روشی مبتنی بر تئوری یادگیری آماری و از نوع غیرپارامتری نظارت‌شده می‌باشد (۱۲)، شامل یک مرحله آموزشی مرتبط با مقادیر ورودی و خروجی هدف است و مطابق عقیده کولکسن و

همکاران (۲۰۱۶)، استحکام و اثربخشی آن در حل مجموعه داده‌های پیچیده و ساختاریافته سبب شده است که در حال حاضر به‌طور فزاینده‌ای در زمینه ارزیابی زمین‌لغزش مورد استفاده واقع گردد (۱۶). الگوریتم جنگل تصادفی که ثابت شده است نتیجه خوبی برای پهنه‌بندی حساسیت زمین‌لغزش دارد، می‌تواند اضافه بر ارزش ناشی از درخت تصمیم را کاهش دهد و بنابراین قادر است که پیش‌بینی بسیار دقیقی نسبت به الگوریتم درخت تصمیم ارائه دهد (۱۹). از طرفی دیگر در الگوریتم رگرسیون لجستیک نیز امکان یافتن بهترین مدل بر ارزش برای توصیف رابطه میان وجود یا عدم وجود زمین‌لغزش و مجموعه متغیرهای مستقل که از آن تحت عنوان عوامل زمین‌لغزش یاد می‌شود، میسر می‌باشد (۱۷). هنگامی که صحبت از پیش‌بینی زمین‌لغزش‌ها می‌شود، رویکردهای سنتی و فیزیکی پهنه‌بندی زمین‌لغزش که بر اساس قوانین و نظریات زمین‌شناسی، ژئوتکنیک، هیدرولوژی یا هواشناسی توسعه یافته‌اند و از طرفی به دلیل پیچیدگی زیاد حرکات توده‌ای، این رویکردها ممکن است برای پیش‌بینی مکان و زمان حرکات توده‌ای و مطالعه تمامی عوامل شناخته‌شده و ناشناخته مؤثر بر زمین‌لغزش امکان پاسخگویی نداشته باشند (۲۱). بر همین اساس در سال‌های اخیر، با توسعه سریع الگوریتم‌های یادگیری ماشین به عنوان یک شاخه نوین علم داده و گسترش آن در بسیاری از زمینه‌های مطالعات محیطی، بسیاری از محققان با توجه به قابلیت این الگوریتم‌ها در مدیریت حجم گسترده‌ای از مجموعه داده‌ها و یافتن الگوهای پیچیده موجود در آن‌ها اقدام به به‌کارگیری الگوریتم‌های مختلف یادگیری ماشین در راستای پهنه‌بندی خطر زمین‌لغزش در نواحی مختلف نموده‌اند. کیانی و همکاران (۱۰)، با استفاده از دو مدل رگرسیون لجستیک و آنفیس (ANFIS) خطر زمین‌لغزش در حوضه آبریز هشتجین در استان اردبیل را مورد مطالعه قرار داده‌اند. آنان نقشه پهنه‌بندی خطر لغزشی را در پنج کلاس خطر ترسیم نموده‌اند. مطابق نتایج حاصله از این پژوهش، شاخص منحنی ROC (Receiver Operating Characteristic) چنین نشان داده است که مدل آنفیس دقتی برابر ۸۸/۲۳ درصد و مدل رگرسیون لجستیک دقتی برابر با ۸۶/۴۵ درصد ارائه نموده است. سو و همکاران (۱۸)، در تحقیقی سه روش غیرخطی برای پهنه‌بندی زمین‌لغزش را مورد ارزیابی و مقایسه قرار داده‌اند. این سه روش شامل ماشین بردار پشتیبان، رگرسیون لجستیک و شبکه عصبی مصنوعی (ANN) بوده و نقشه پهنه‌بندی خطر لغزشی نیز در چهار کلاس به نمایش درآمده است. نتایج حاصل از اجرای سه مدل مذکور به منظور تعیین بهترین روش، با استفاده از شاخص‌های ROC و AUC (Area Under Curve) مورد مقایسه قرار گرفته است که نشان می‌دهد مقادیر صحت پیش‌بینی در مرحله آموزشی بین ۷۳/۴۳٪ تا ۸۷/۴۵٪ و در مرحله اعتبارسنجی بین ۶۷/۱۶٪ تا ۷۳/۱۳٪ متغیر بوده و مقادیر AUC حاصله نیز از ۰/۸۰۷ تا ۰/۹۰۶ در مرحله آموزشی و ۰/۷۵۳ تا ۰/۹۴۴ در مرحله اعتبارسنجی حکایت داشته است. نتایج نهایی نشان‌دهنده تعادل نسبی بالاتر روش SVM نسبت به دو روش دیگر بوده و این مدل بهره پیش‌بینی دقیق‌تری را در مقایسه با دو مدل دیگر حاصل کرده است. آشور و پورقاسمی (۱)، به ارزیابی تطبیقی مقادیر دقت حاصله برای سه مدل یادگیری ماشین از

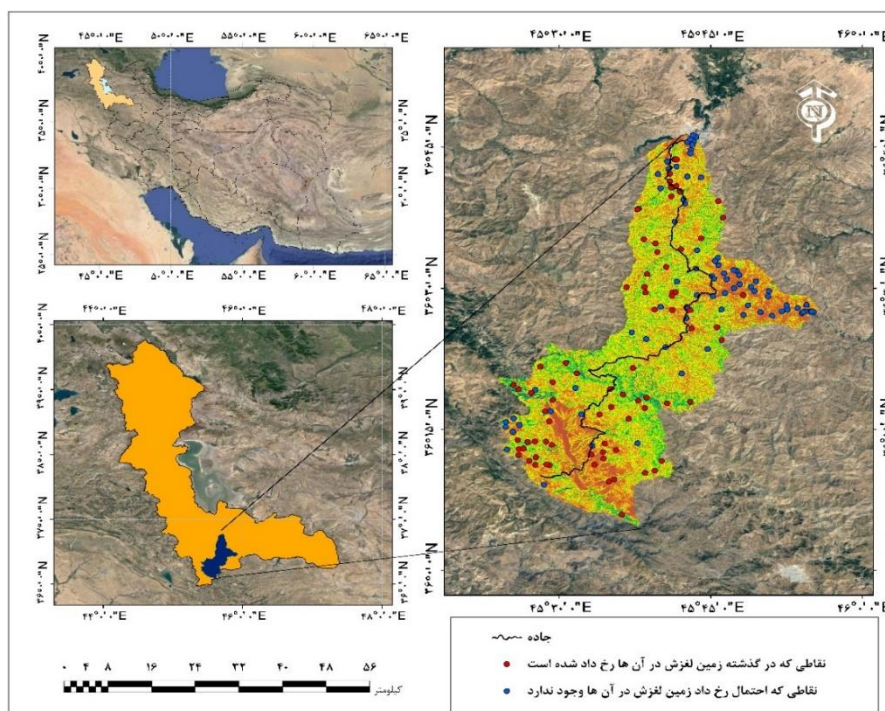
قبیل جنگل تصادفی، ماشین بردار پشتیبان و درخت رگرسیونی تقویت شده در زمینه حساسیت وقوع زمین لغزش در محور بزرگراه A1 در کشور الجزایر پرداخته‌اند. آنان برای ارزیابی دقت حاصله برای مدل‌های فوق‌الذکر، از سه شاخص منحنی راک، خطای استاندارد (Standard Error) و سطح اطمینان در حد ۹۵ درصد بهره گرفته‌اند. بر اساس نتایج ارزیابی‌های آنان، مدل جنگل تصادفی با مقدار مساحت زیر منحنی ۹۷/۲٪ بیشترین دقت پیش‌بینی را نسبت به دو مدل دیگر از خود نشان داده است. پی و همکاران (۲۷)، اقدام به مقایسه مدل‌های مختلف یادگیری ماشین در راستای پهنه‌بندی حساسیت زمین لغزش نموده‌اند. برای این منظور آنان سه الگوریتم رگرسیون لجستیک، ماشین بردار پشتیبان و بیز ساده (Naïve Bayes) را در سطح محدوده مطالعاتی خود که بخشی از استان سیچوان (Sichuan) در کشور چین بوده است پیاده نموده‌اند. بر اساس مقادیر ارزیابی دقت و مساحت زیر منحنی، نتایج این مطالعه چنین مشخص نموده است که مدل رگرسیون لجستیک در مقایسه با دو مدل دیگر بهترین عملکرد را داشته است و مدل بیز ساده در رتبه دوم و مدل ماشین بردار پشتیبان در رتبه سوم به لحاظ دقت ارزیابی جای گرفته است.

محدوده مهاباد - سردشت در شمال غربی ایران به دلیل شرایط خاص خود از قبیل توپوگرافی کوهستانی و عبور از گردنه زمزیران، شیب زیاد، شرایط اقلیمی خاص که موجب بارش‌های بهاری و ذوب برف‌ها در فصل بهار شده یا برف‌گیر بودن و وجود یخبندان و وجود ساختارهای سنگ‌شناسی سست و نامقاوم در برخی نواحی و نهایتاً لرزه‌خیزی، از پتانسیل بالایی به لحاظ وقوع زمین لغزش برخوردار است و از آنجایی که وقوع زمین لغزش می‌تواند موجب خسارت و آسیب به زیرساخت‌های ارتباطی، نواحی مسکونی و فرسایش خاک در این منطقه گردد، لذا لازم است که با استفاده از دقیق‌ترین و بهترین الگوریتم‌های یادگیری ماشین ارزیابی مناسبی از پتانسیل لغزشی آن که حداکثر انطباق را با واقعیات موجود داشته باشد به انجام برسد. با توجه به موارد بیان شده، تحقیق حاضر با هدف مقایسه عملکرد و کارایی الگوریتم‌های مختلف از قبیل ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک و تلفیق داده‌های جغرافیایی مختلف به منظور پهنه‌بندی خطر زمین لغزش در محدوده جاده مواصلاتی مهاباد - سردشت در استان آذربایجان غربی در شمال غربی ایران به انجام رسیده است تا با بتوان ارزیابی دقیقی از پتانسیل وقوع زمین لغزش در نواحی مختلف آن انجام داده و علاوه بر اینکه امکان شناسایی مناطق در معرض خطر میسر شود، از طرفی به مقایسه کارایی مدل‌های تحلیلی مورد استفاده در این زمینه پرداخته شود و بهترین مدل در این راستا معرفی گردد.

۲- روش تحقیق

۲-۱- محدوده مورد مطالعه

محدوده مورد مطالعه در مختصات جغرافیایی بین ۳۶ درجه و ۰۵ دقیقه تا ۳۶ درجه و ۴۷ دقیقه طول شرقی و ۴۵ درجه و ۲۵ دقیقه تا ۴۵ درجه و ۵۵ دقیقه عرض شمالی و در منتهی الیه جنوبی استان آذربایجان غربی واقع گردیده است که شامل بخشی از اراضی شهرستان‌های سردشت و مهاباد با مساحتی بالغ بر ۱۳۷۲/۶۸ کیلومتر مربع است. کمینه ارتفاع محدوده مورد مطالعه ۹۷۱ و بیشینه آن نیز ۲۳۸۱ متر از سطح دریاهای آزاد می‌باشد. اغلب محدوده مورد مطالعه کوهستانی بوده و بخش‌های کمی دارای منظر جلگه‌ای می‌باشد که این اراضی عمدتاً در قسمت شمالی محدوده مورد مطالعه قرار گرفته است. پوشش گیاهی غالب از نوع تیپ پوشش جنگل‌های زاگرس است که وسعت این جنگل‌ها ۱۶۶۰ کیلومتر مربع بوده و درختان منطقه بیشتر از نوع بلوط می‌باشند. وضعیت توپوگرافی منطقه مطالعاتی به‌گونه‌ای است که از لحاظ حرکات دامنه‌ای انتظار می‌رود که قسمت‌های جنوبی محدوده مورد مطالعه به لحاظ پتانسیل‌های سطحی نسبت به قسمت‌های شمالی دارای احتمال وقوع حرکات توده‌ای بیشتری باشند. در شکل (۱) نقشه موقعیت جغرافیایی محدوده مورد مطالعه نشان داده شده است.



شکل (۱): موقعیت جغرافیایی محدوده مطالعاتی
Fig (1): Location of study area

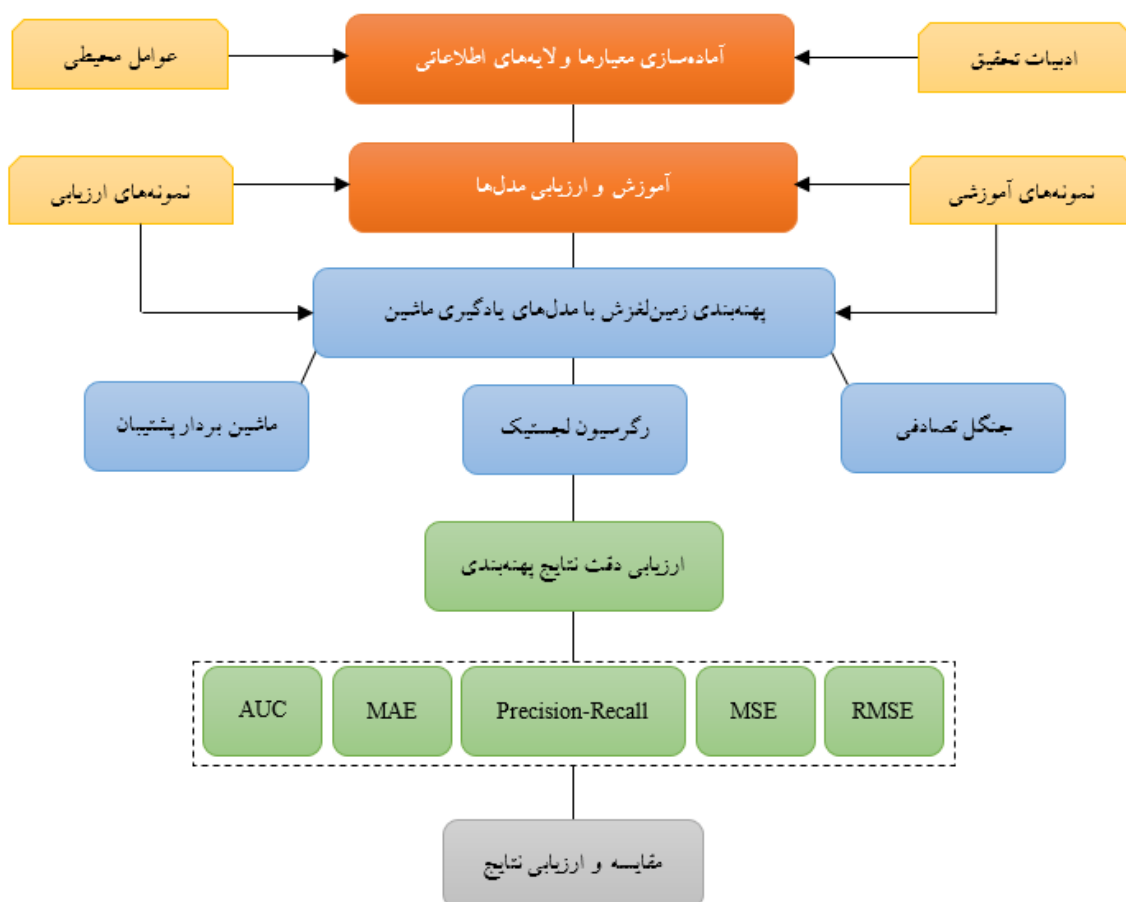
۲-۲- داده‌های مورد استفاده

در مطالعه حاضر ابتدا بر اساس پیشینه تحقیق و با توجه به خصوصیات محدوده مورد مطالعه، عوامل مختلف مؤثر در وقوع زمین‌لغزش شامل ارتفاع، جهت شیب، شیب، فاصله از گسل، تراکم آبراهه، فاصله از آبراهه، خاک‌شناسی، کاربری اراضی، شاخص پوشش گیاهی (NDVI)، سنگ‌شناسی، انحنای شیب، شاخص توان آبراهه (SPI) و شاخص رطوبت توپوگرافی (TWI) شناسایی گردیده و در مرحله بعد لایه‌های اطلاعاتی مربوط به هر یک از عوامل، در محیط ArcGIS تهیه گردیده است. داده‌ها و لایه‌های مکانی موردنیاز این تحقیق از منابع ذیل تهیه یا گردآوری شده است:

- لایه‌های سنگ‌شناسی، خاک‌شناسی و گسل (برگرفته از نقشه‌های زمین‌شناسی ۱:۱۰۰۰۰۰ سازمان زمین‌شناسی کشور)؛
- لایه موقعیت نقاط زمین‌لغزش (اخذ شده از سازمان جنگل‌ها و مراتع و نیز برداشت شده با استفاده از عملیات میدانی و GPS)؛
- لایه مقدار بارندگی در بازه زمانی ۲۰۲۰ - ۲۰۰۰ (اخذ شده از سازمان هواشناسی کشور)؛
- لایه کاربری اراضی (برگرفته از لایه کاربری اراضی جهانی سال ۲۰۲۱ در ۱۲ کلاس مؤسسه ESRI)؛
- لایه‌های شیب، جهت شیب، انحنای شیب، ارتفاع و آبراهه‌ها (برگرفته از مدل رقومی ارتفاعی (DEM) ۳۰ متری سنجد Aster موجود در پایگاه اطلاعاتی سازمان زمین‌شناسی ایالات متحده آمریکا (USGS))؛
- لایه‌های SPI و TWI (تولید شده در محیط نرم‌افزار SAGA GIS)؛
- لایه پوشش گیاهی (NDVI) (برگرفته از تصاویر سنجنده سنتینل-۲).

۲-۳- روش تجزیه و تحلیل

به منظور پهنه‌بندی حساسیت زمین‌لغزش در محدوده مطالعاتی از روش تطبیقی و مقایسه الگوریتم‌های ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF) و رگرسیون لجستیک (LR) بهره گرفته شده است. مطابق فرآیند کلی تحقیق، که در شکل (۲) نشان داده شده است، کاربرد الگوریتم‌های فوق‌الذکر نیازمند نمونه‌های آموزشی (Training Samples) با تعداد و توزیع کافی در سطح کلاس‌های مختلف می‌باشد و در همین راستا نمونه‌های مربوط به نقاط زمین‌لغزش و غیر زمین‌لغزش به تعداد ۸۰ مورد با پراکنش مناسب از سطح محدوده مورد مطالعه برداشت گردیده است که از این تعداد، ۵۶ نقطه (۷۰ درصد از کل نقاط) در مرحله طبقه‌بندی و ۲۴ نقطه (۳۰ درصد باقی از کل نقاط) در مرحله اعتبارسنجی مورد استفاده واقع شده است.



شکل (۲): فرآیند کلی انجام تحقیق

Fig (2): General process of the research

ماشین بردار پشتیبان به عنوان یک روش استاندارد در راستای پهنه‌بندی و پیش‌بینی حساسیت زمین‌لغزش در نظر گرفته شده است. این روش یادگیری ماشینی ابتدا داده‌ها را از فضای ورودی اصلی به یک فضای ویژگی با ابعاد بالا ترسیم می‌کند که در آن می‌توان از یک ابرصفحه برای انجام طبقه‌بندی داده‌ها استفاده کرد (۲۲). مدل SVM یک مرز طبقه‌بندی از مجموعه نقاط آموزشی ایجاد می‌کند تا حاشیه بین دو کلاس تا حد امکان گسترده باشد. در این راستا، خروجی کلاس ۱- یک رخداد غیر لغزشی را نشان می‌دهد و ۱+ نشان‌دهنده یک رخداد زمین‌لغزش است (۷). الگوریتم ماشین بردار پشتیبان به منظور تهیه نقشه خطر زمین‌لغزش در سطح محدوده مورد مطالعه با استفاده از تابع شعاعی (RBF) و با تعداد ۵۶ نقطه آموزشی به اجرا درآمده است. از طرفی سادگی و انعطاف مدل رگرسیون لجستیک گاهی اوقات می‌تواند منجر به عملکرد بهتر از دیگر مدل‌های غیرخطی

پیچیده مانند ماشین بردار پشتیبان شود (۶). با این حال، در صورتی که متغیر پاسخ از یک حجم نمونه کوچک ترسیم شود، مدل رگرسیون لجستیک مناسب نبوده و برای پاسخ‌های باینری (Binary) ضعیف عمل می‌کند (۴). مدل رگرسیون لجستیک یک روش ریاضی رایج برای ایجاد ارتباط میان عوامل محرکه مختلف و خطر زمین‌لغزش است (۲۵) و بر همین اساس با استفاده از معیارهای مربوطه و نقاط لغزشی مشاهده‌ای به اجرا درآمده است. مدل جنگل تصادفی نیز یک الگوریتم کلاسیک در یادگیری گروهی است که درخت‌های تصمیم‌گیری (Decision Trees) متعددی را برای ساختن یک طبقه‌بندی کننده مجموعه آموزش ارائه می‌دهد. بر اساس فرآیند الگوریتم جنگل تصادفی، چندین زیرمجموعه آموزشی مختلف با نمونه‌گیری با جایگزینی از مجموعه آموزشی به دست می‌آیند. هر زیرمجموعه آموزشی برای آموزش یک درخت تصمیم استفاده می‌شود و چندین درخت تصمیم آموزش دیده می‌توانند یک طبقه‌بندی کننده کلی را تشکیل دهند (۱۴). هنگامی که یک نمونه به طبقه‌بندی کننده کلی وارد می‌شود، هر درخت تصمیم یک نتیجه طبقه‌بندی را خروجی می‌دهد و نتیجه طبقه‌بندی نهایی از برآیند طبقه‌بندی‌ها حاصل می‌شود (۸). پس از برداشت و آماده‌سازی نمونه‌ها، پیاده‌سازی فاز عملیاتی و طبقه‌بندی مد نظر قرار گرفته که بدین منظور با بهره‌گیری از محیط کدنویسی Anaconda Spyder IDE الگوریتم‌های سه‌گانه به اجرا درآمده است.

۴-۲-۴-۲-۴ اعتبارسنجی

در مطالعه حاضر به منظور ارزیابی دقت و اعتبارسنجی نقشه‌های پهنه‌بندی حاصل از هر سه الگوریتم، ۲۴ نقطه از مجموع نقاط نمونه برداشت شده مورد استفاده قرار گرفته و با بهره‌گیری از پارامترهای ارزیابی از قبیل مساحت زیر منحنی (AUC)، صحت و یادآوری (Precision-Recall)، خطای میانگین مربعات (MSE)، خطای جذر میانگین مربعات (RMSE) و میانگین خطای مطلق (MSE)، نقشه‌ها ارزیابی گردیدند. منحنی ROC نرخ مثبت واقعی یا حساسیت لغزشی طبقه‌بندی شده درست را در برابر نرخ‌های منفی واقعی، غیرزمین‌لغزش‌ها یا ویژگی‌های طبقه‌بندی شده درست ترسیم می‌کند (۲۴). نتایج منحنی راک از طریق AUC ارزیابی می‌شود که در آن طبقه‌بندی به صورت $AUC = 0.5$ و موفقیت کامل به صورت $AUC = 1$ نشان داده می‌شود (رابطه ۱).

$$AUC = \frac{\sum TP + \sum TN}{P + N} \quad (1)$$

که در این رابطه، N و P به ترتیب تعداد نقاط غیر لغزش و زمین لغزش را نشان می‌دهند. پارامتر RMSE نیز خطای محاسبه شده بین داده‌های پیش‌بینی و مشاهده‌ای را نشان می‌دهد (۵) (رابطه ۲).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{predicted} - X_{actual})^2} \quad (2)$$

که در این رابطه، n مجموع نقاط آموزشی تعریف شده، $X_{\text{predicted}}$ و X_{actual} به ترتیب نمونه پیش‌بینی شده در داده‌های آموزشی و مقدار واقعی زمین‌لغزش هستند. شاخص MSE نیز میانگین مربعات خطاها را اندازه‌گیری می‌کند؛ یعنی میانگین مجذور اختلاف بین مقادیر پیش‌بینی شده و مقادیر واقعی (۹) (رابطه ۳).

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (3)$$

که در این رابطه نیز Y_i ، مقادیر مشاهده‌ای واقعی، \hat{Y}_i ، مقادیر مشاهده‌ای پیش‌بینی شده و n تعداد کل مشاهدات است. شاخص MAE میانگین بزرگی خطاها را در مجموعه‌ای از پیش‌بینی‌ها بدون در نظر گرفتن جهت آن‌ها اندازه‌گیری می‌کند (۲۳) (رابطه ۴).

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_{\text{predicted}} - X_{\text{actual}}| \quad (4)$$

که در آن n تعداد خطاها، $X_{\text{predicted}}$ مقدار پیش‌بینی و X_{actual} مقدار واقعی است. شاخص صحت (که ارزش پیش‌بینی مثبت نیز نامیده می‌شود) کسری از نمونه‌های مربوطه در بین نمونه‌های بازبایی شده است، در حالی که یادآوری (همچنین به عنوان حساسیت شناخته می‌شود) کسری از نمونه‌های مربوطه است که بازبایی شده‌اند. بنابراین، هم صحت و هم یادآوری بر اساس ارتباط مقادیر داده‌ها سنجیده می‌شوند. این شاخص از طریق روابط زیر محاسبه می‌شود:

$$Precision = \frac{tp}{tp+fp} \quad (5)$$

$$Recall = \frac{tp}{tp+fn} \quad (6)$$

که در آن tp ، مقادیر مثبت حقیقی، fp ، مقادیر مثبت کاذب و N کل مشاهدات می‌باشد.

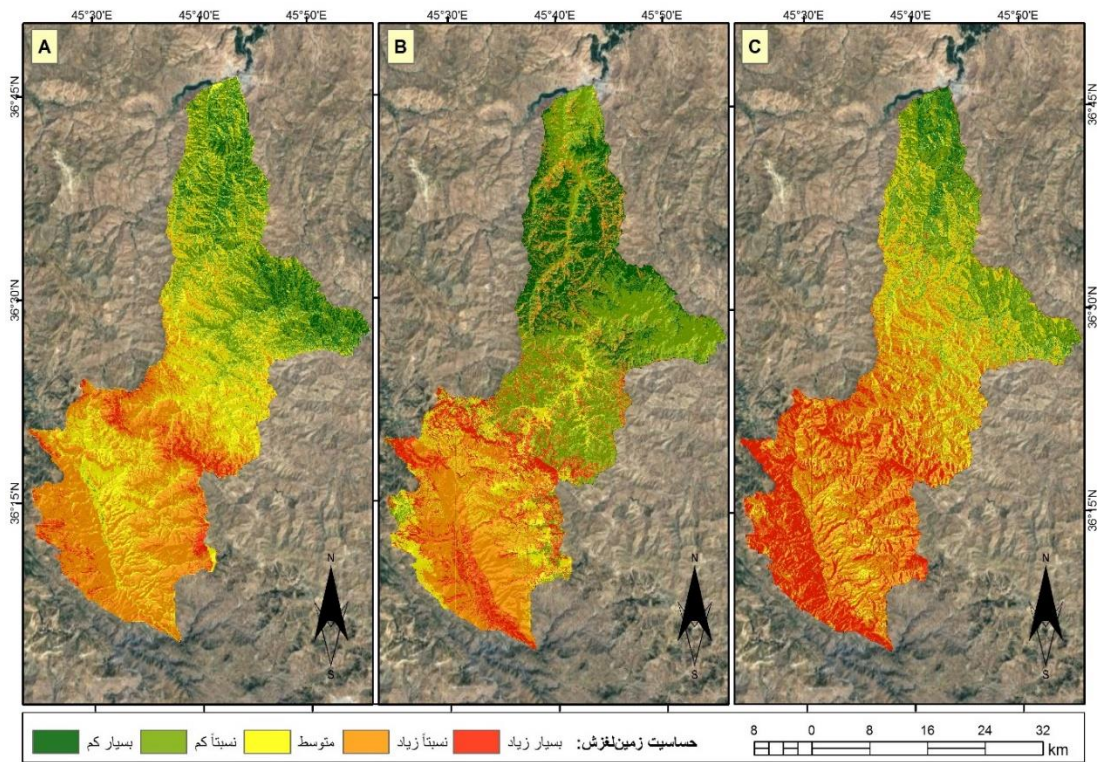
۳- نتایج و بحث

تهیه نقشه پهنه‌بندی زمین‌لغزش، اولین فاز ارزیابی در محدوده مورد مطالعه بوده است که مطابق آن شدت و ضعف حساسیت فضایی و نواحی پرخطر شناسایی گردیده است. در این راستا محدوده مورد مطالعه با تعیین روش طبقه‌بندی Natural Breaks به پنج کلاس حساسیت در بازه بسیار کم تا بسیار زیاد طبقه‌بندی شده است (شکل ۳). مطابق نقشه حاصل از الگوریتم ماشین بردار پشتیبان، محدوده مورد مطالعه دارای حساسیت متوسط به بالایی از نظر خطر زمین‌لغزش می‌باشد و از این نظر بخش جنوبی منطقه دارای حساسیت بسیار بالاتری نسبت به بخش شمالی آن بوده و به عبارتی دیگر احتمال وقوع زمین‌لغزش در نیمه جنوبی محدوده بسیار بالاتر از نیمه شمالی آن می‌باشد و از طرفی نقاط پراکنده‌ای در شمال و شمال غربی منطقه با خطر لغزشی بالا

شناسایی شده است. تحلیل خطرپذیری محیطی به لحاظ وقوع زمین لغزش چنین مشخص نموده است که بیشترین پتانسیل لغزشی در سطح منطقه در مجاورت گسل‌ها و نواحی تکتونیکی بوده و از طرفی نواحی مرتفع با شیب نسبی زیاد (بیشتر از ۳۵ درجه) نیز در این ارتباط نقش بسیار زیادی را داشته است (شکل ۳-الف). بر اساس نقشه حاصل از مدل جنگل تصادفی و همچون پهنه‌بندی صورت گرفته برای ماشین بردار پشتیبان، خطر زمین لغزش در محدوده مورد مطالعه در دو نیمه شمالی و جنوبی قابل بررسی است که نیمه جنوبی منطقه به دلیل شرایط توپوگرافی و تکتونیکی پتانسیل بسیار بالاتری به لحاظ وقوع زمین لغزش دارا می‌باشد اما نقشه مربوط به مدل جنگل تصادفی در قیاس با نقشه حاصله از مدل SVM نمایش محتاطانه‌تری از خطر زمین لغزش منطقه ارائه نموده است به گونه‌ای که کلاس حساسیت متوسط به بالا میزان کم‌تری را نمایش داده است اما از طرفی کلاس حساسیت بسیار زیاد میزان بالاتری را دربر گرفته است (شکل ۳-ب). همچنین بر اساس ارزیابی عوامل محیطی مؤثر بر پهنه‌بندی خطر زمین لغزش با مدل جنگل تصادفی، پارامترهای شیب، شاخص توان آبراهه و فاصله از گسل به عنوان مهم‌ترین عوامل در وقوع زمین لغزش در سطح منطقه شناسایی شده‌اند درحالی‌که پارامترهای خاک، زمین‌شناسی و کاربری اراضی کم‌ترین تأثیر را در این ارتباط دارا بوده‌اند. بر این اساس می‌توان گفت که شیب‌های تند کوهستانی در محدوده مسیر مهاباد - سردشت نقش اصلی را در بروز زمین لغزش ایفا می‌کند و تأثیر جریان آبراهه در شیب‌های بالا که در طول بریدگی‌های کوهستانی به سمت پایین دست روانه می‌شود و البته نقش لرزه‌خیزی که موجب ناپایداری شیب در نواحی مجاور گسلی می‌گردد، عوامل اصلی وقوع زمین لغزش در این منطقه قلمداد می‌شوند. نقشه پهنه‌بندی مربوط به مدل رگرسیون لجستیک نیز حاکی از این بوده است که گستره وسیعی از محدوده مطالعاتی دارای حساسیت متوسط به بالا به لحاظ وقوع زمین لغزش می‌باشد به نحوی که به نظر می‌رسد مساحت زیادی در گروه خطر نسبتاً زیاد و بسیار زیاد طبقه‌بندی شده است و علاوه بر این بسیاری از مناطق که در دو مدل قبلی به عنوان نواحی دارای حساسیت لغزشی شناسایی نشده بودند، در مدل رگرسیون لجستیک به منزله نواحی دارای حساسیت نمایش داده شده‌اند و برای مثال درصد قابل توجهی از اراضی دارای شیب پایین (زیر ۵ درجه) نیز حساسیت بالایی ارائه نموده‌اند که این امر در مورد نیمه شمالی منطقه که شیب زمین کمتر از نیمه جنوبی آن است بیشتر صدق می‌کند (شکل ۳-ج).

نتایج پهنه‌بندی حساسیت زمین لغزش در جدول ۱ نشان داده شده است. بر اساس نتایج الگوریتم ماشین بردار پشتیبان و از کل مساحت محدوده مطالعاتی، ۷/۶ درصد دارای حساسیت بسیار کم، ۲۱/۳۶ درصد دارای حساسیت نسبتاً کم، ۲۶/۶۸ درصد دارای حساسیت متوسط، ۲۷/۲۷ درصد دارای حساسیت نسبتاً زیاد و ۷/۰۹ درصد دارای حساسیت بسیار زیاد می‌باشد. بر اساس نتایج، این میزان برای الگوریتم جنگل تصادفی به ترتیب برابر با ۱۴/۴۲، ۳۲/۱۴، ۱۴/۵۷، ۲۵/۶۸ و ۱۳/۱۹ درصد و برای الگوریتم رگرسیون لجستیک به ترتیب برابر با

۴/۷۳، ۱۷/۸۸، ۲۲/۱۸، ۳۲/۱۵، ۲۳/۰۶ درصد بوده است. به‌طور کلی نتایج حاکی از این است که الگوریتم رگرسیون لجستیک ۵۵/۲۱ درصد، الگوریتم ماشین بردار پشتیبان ۴۴/۳۶ درصد و الگوریتم جنگل تصادفی ۳۸/۸۷ درصد از سطح محدوده مسیر مهاباد - سردشت را به عنوان نواحی پرمخاطره (کلاس‌های حساسیت زیاد و بسیار زیاد) شناسایی کرده‌اند.



شکل (۳): نقشه پهنه‌بندی حساسیت زمین‌لغزش (الف: ماشین بردار پشتیبان؛ ب: جنگل تصادفی؛ ج: رگرسیون لجستیک)
Fig (3): Landslide susceptibility map (A: Support vector machine; B: Random forest; C: Logistic regression)

جدول (۱): نتایج پهنه‌بندی حساسیت زمین لغزش

Table (1): Landslide susceptibility zoning results

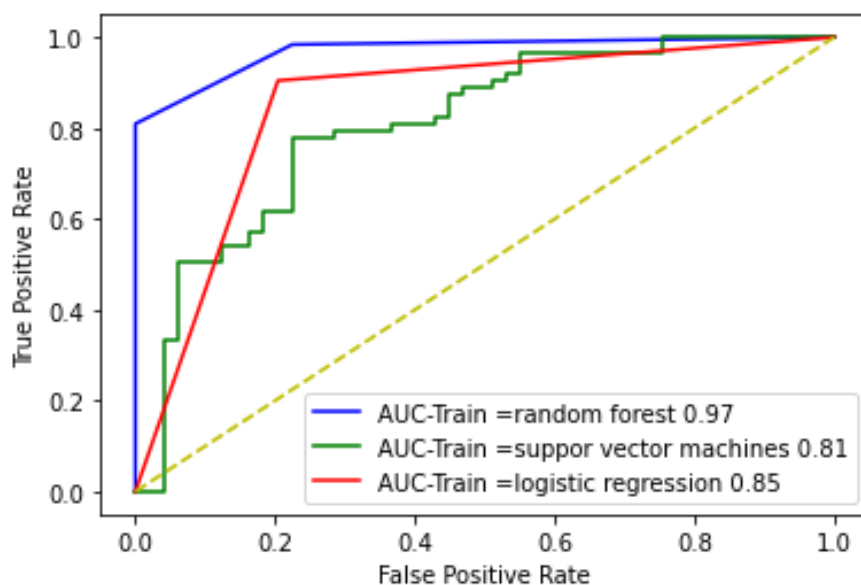
رگرسیون لجستیک (LR)		جنگل تصادفی (RF)		ماشین بردار پشتیبان (SVM)		کلاس‌های پهنه‌بندی
نسبت (درصد)	مساحت (کیلومتر مربع)	نسبت (درصد)	مساحت (کیلومتر مربع)	نسبت (درصد)	مساحت (کیلومتر مربع)	
۴/۷۳	۶۴/۷۶۹	۱۴/۴۲	۱۹۷/۳۹۷	۷/۶	۱۰۴/۰۹۵	حساسیت خیلی کم
۱۷/۸۸	۲۴۴/۹۰۷	۳۲/۱۴	۴۴۰/۱۲۳	۲۱/۳۶	۲۹۲/۴۹۷	حساسیت کم
۲۲/۱۸	۳۰۳/۷۳۸	۱۴/۵۷	۱۹۹/۴۴۳	۲۶/۶۸	۳۹۶/۴۳۵	حساسیت متوسط
۳۲/۱۵	۴۴۰/۲۶۵	۲۵/۶۸	۳۵۱/۷۱۸	۳۷/۲۷	۵۱۰/۴۳۹	حساسیت زیاد
۲۳/۰۶	۳۱۵/۷۳۱	۱۳/۱۹	۱۸۰/۷۲۹	۷/۰۹	۹۶/۹۴۵	حساسیت خیلی زیاد
۱۰۰	۱۳۶۹/۴۱۱	۱۰۰	۱۳۶۹/۴۱۱	۱۰۰	۱۳۶۹/۴۱۱	مجموع

تحلیل حساسیت و مقایسه نقشه پهنه‌بندی با نقشه توزیع فضایی زمین لغزش‌های موجود چنین نشان داده است که الگوریتم رگرسیون لجستیک بیشترین درصد انطباق را با لغزش‌های پیشین منطقه داشته است به نحوی که ۶۵ درصد زمین لغزش‌های قبلی در کلاس‌های دارای حساسیت نسبتاً زیاد و بسیار زیاد رخ داده است. این میزان برای الگوریتم جنگل تصادفی برابر با ۵۵ درصد و برای الگوریتم ماشین بردار پشتیبان نیز برابر با ۵۰ درصد بوده است. از طرفی کارایی هر سه الگوریتم در شناسایی بهینه نواحی دارای حساسیت لغزشی بسیار کم و نسبتاً کم نیز اثبات گردیده است (شکل ۴).



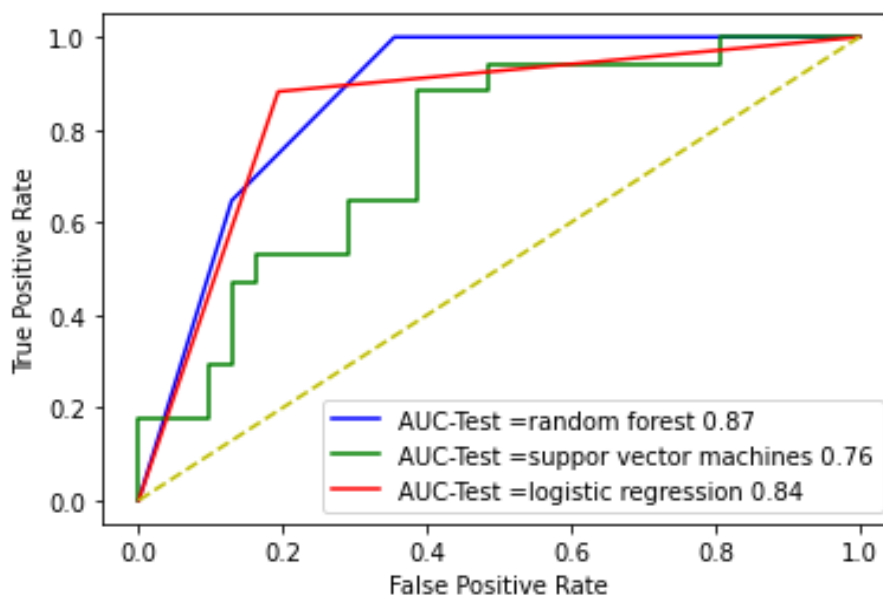
شکل (۴): مقایسه نتایج پهنه‌بندی با زمین لغزش‌های موجود
 Fig (4): Comparison of zoning results with existing landslides

همانطور که بیان گردید، در منحنی ROC هرچه قدر سطح زیر منحنی بیشتر باشد به معنی دقت بالای پیش‌بینی‌های صورت‌گرفته است که این منحنی در بازه ۰/۵ تا ۱ متغیر می‌باشد و تفسیر آن بدین‌گونه است که مقادیر بین ۱ - ۰/۹ عملکرد عالی، ۰/۸ - ۰/۹ عملکرد خیلی خوب، ۰/۷ - ۰/۸ عملکرد خوب، ۰/۶ - ۰/۷ عملکرد متوسط و ۰/۵ تا ۰/۶ عملکرد ضعیفی را برای مدل نشان می‌دهند. نتایج اعتبارسنجی بیانگر این بوده است که با توجه به مقادیر حاصله، هر سه الگوریتم مورد استفاده در مرحله آموزش مدل (Train) دارای عملکرد خیلی خوب تا عالی در راستای پیش‌بینی خطر وقوع زمین‌لغزش بوده‌اند و ارزیابی نمودار مربوطه برای مرحله آزمون (Validation) نیز بیانگر این بوده که در بین سه الگوریتم، جنگل تصادفی با نرخ ۰/۸۷ و رگرسیون لجستیک با نرخ ۰/۸۴ عملکرد خیلی خوب و ماشین بردار پشتیبان نیز با نرخ ۰/۷۶ عملکرد خوب از خود ارائه نموده‌اند (شکل ۵ و ۶) و بنابراین الگوریتم جنگل تصادفی بهترین عملکرد را نسبت به دو الگوریتم دیگر در راستای پهنه‌بندی خطر وقوع زمین‌لغزش از خود نشان داده است. همچنین منحنی صحت - یادآوری ارزیابی نیز مؤید این نکته است که الگوریتم جنگل تصادفی در مقایسه با دو الگوریتم دیگر عملکرد دقیق‌تری را در پهنه‌بندی خطر وقوع زمین‌لغزش در محدوده مورد مطالعه ارائه نموده است، به طوری که مقدار حاصله برای جنگل تصادفی برابر با ۰/۸۷۳ و برای ماشین بردار پشتیبان و رگرسیون لجستیک به ترتیب برابر با ۰/۸۰۹ و ۰/۸۴۴ به دست آمده است (شکل ۷).



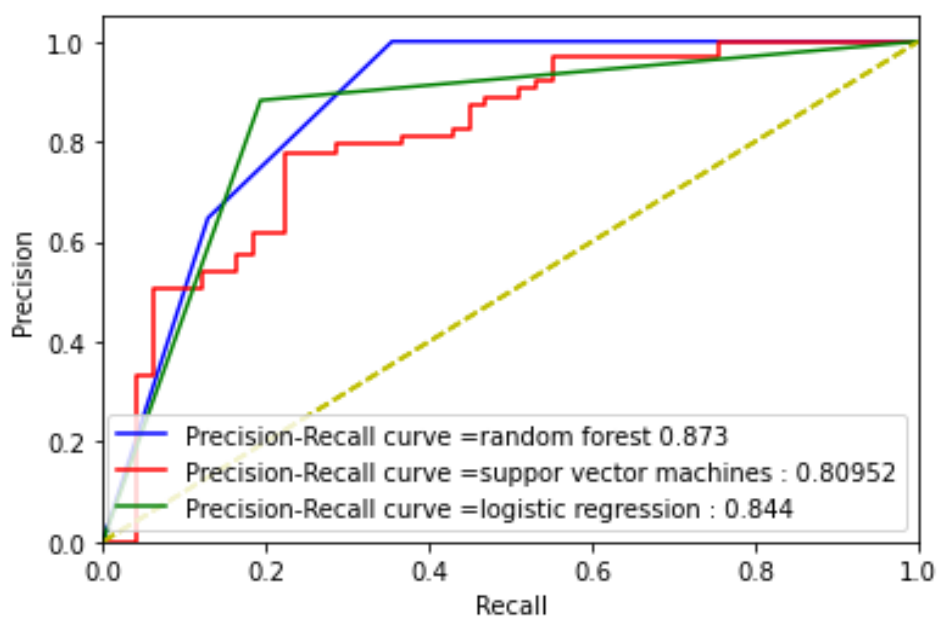
شکل (۵): مقایسه نتایج الگوریتم‌ها در مرحله آموزش

Fig (5): Comparison of algorithms in the training phase



شکل (۶): مقایسه نتایج الگوریتم‌ها در مرحله آزمون

Fig (6): Comparison of algorithms in the validation phase



شکل (۷): مقایسه نتایج منحنی صحت-یادآوری الگوریتم‌ها

Fig (7): Comparison of precision-recall curve of algorithms

از طرفی سایر شاخص‌های ارزیابی صحت نیز حکایت از عملکرد مطلوب مدل‌های مورد استفاده در پهنه‌بندی زمین‌لغزش در محدوده مطالعاتی داشته است. بر این اساس، مقدار میانگین مطلق خطا (MAE) برای مدل جنگل تصادفی برابر با ۰/۲۴۵۸ و مقدار مجذور خطای آن (MSE) نیز برابر با ۰/۱۱۷۸ به دست آمده است که نشان‌دهنده مقدار خطای پایین ارزیابی است. همچنین مقدار RMSE نیز برابر با ۰/۳۴۳۳ به دست آمده است که مقدار پایینی بوده و دقت بالای ارزیابی این مدل را بیان می‌کند. همچنین برای مدل ماشین بردار پشتیبان، مقدار میانگین مطلق خطا برابر با ۰/۳۴۴۳۹ و مقدار میانگین مجذور خطا برابر با ۰/۲۱۸۹۳ و مقدار RMSE برای هر دو نقاط آموزشی و صحت‌سنجی برابر با ۰/۴۶۷۹۰ به دست آمده است. نهایتاً تحلیل نتایج مدل رگرسیون لجستیک حاکی از این می‌باشد که مقدار میانگین مطلق خطا برابر با ۰/۲۹۱۲ و مقدار مجذور خطا برابر با ۰/۱۶۰۵ بوده که حاکی از مقادیر پایین خطای ارزیابی است. از طرفی نرخ خطای RMSE مدل رگرسیون لجستیک برابر با ۰/۳۲۳۰ بوده که مقدار مناسبی به لحاظ خطای ریشه‌ای ارائه نموده است.

۴- نتیجه‌گیری

با توجه به خسارات ناشی از وقوع خطر زمین‌لغزش در مناطق مختلف، ضروری است که تهیه نقشه پهنه‌بندی حساسیت زمین‌لغزش و شناسایی مناطق مستعد با روش‌ها و ابزارهای مناسب مد نظر قرار گیرد. بدین منظور در تحقیق حاضر به منظور پهنه‌بندی حساسیت زمین‌لغزش در محدوده مهاباد - سردشت در شمال غربی ایران، سه الگوریتم ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون لجستیک مورد استفاده واقع شده و کارایی آن‌ها در این زمینه با یکدیگر مقایسه و ارزیابی گردید. مطابق نتایج و بر اساس شاخص‌های مختلف ارزیابی صحت، مدل جنگل تصادفی بهترین و بالاترین دقت را نسبت به مدل رگرسیون لجستیک و ماشین بردار پشتیبان در زمینه پهنه‌بندی زمین‌لغزش ارائه نموده است. از طرفی معیارهای محیطی مختلف در قالب ساختارهای توپوگرافی، زمین‌شناسی، هیدرولوژی و انسانی وارد فرآیند ارزیابی گردیده و تأثیر نسبی هر یک از عوامل در زمینه وقوع زمین‌لغزش در محدوده مورد مطالعه مشخص گردید. بر اساس نقشه‌های حاصله مشخص گردید که نیمه شمالی منطقه به دلایلی از قبیل شیب نسبی کم‌تر، ارتفاع کم‌تر، عدم وجود ساختار گسلی متراکم و تراکم کم‌تر آبراهه دارای حساسیت لغزشی پایین‌تری نسبت به نیمه جنوبی آن می‌باشد و از طرفی نواحی دارای حداکثر حساسیت لغزشی در بخش‌های جنوب غربی و جنوب شرقی منطقه مستقر شده و درصد قابل توجهی از وقایع پیشین زمین‌لغزش نیز در آن‌جا به وقوع پیوسته است. علاوه بر این چنین مشخص گردیده است که سه ناحیه عمده در بخش جنوبی محدوده مهاباد - سردشت دارای حداکثر حساسیت به لحاظ وقوع خطر زمین‌لغزش می‌باشد که این سه ناحیه عبارتند از گردنه زمزیران از روستای کله گاوی تا روستای گویران در منتهی الیه جنوب غربی منطقه، مسیر کوهستانی گرد شیلان تا نستان در بخش جنوب شرقی منطقه و محدوده

زیوه تا سردشت در غرب رودخانه زاب کوچک. بر اساس نتایج حاصله و با توجه به توسعه الگوریتم‌ها و تکنیک‌های مختلف پردازشی، پیشنهاد می‌شود در مطالعات آتی به مقایسه کارایی و عملکرد مدل‌های تحلیلی با ماهیت مختلف به منظور پهنه‌بندی حساسیت زمین لغزش در منطقه مورد مطالعه پرداخته شود و با در نظر گرفتن این موضوع که برخی مطالعات از روش فیوژن الگوریتم‌های یادگیری ماشین بهره برده‌اند و دقت و اعتبار نتایج پهنه‌بندی زمین لغزش تا حدی ارتقا یافته است، لذا پیشنهاد می‌شود که این مسئله در تحقیقات آتی برای نواحی مورد مطالعه مد نظر قرار بگیرد.

۵- منابع

- 1- Achour, Y. Pourghasemi, H.R. (2020). How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? *Geoscience Frontiers*, 11(3); 871-883.
- 2- Ahmad, MN. Shao, Z. Aslam, RW. Ahmad, I. Liao, M. Li, X. Song, Y. (2022). Landslide hazard, susceptibility and risk assessment (HSRA) based on remote sensing and GIS data models: a case study of Muzaffarabad Pakistan. *Stochastic Environmental Research and Risk Assessment*. May 27; 1-6.
- 3- Akinci, H. Zeybek, M. (2021). Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey. *Natural Hazards*, 108(2); 1515-1543.
- 4- Anastasiy, Motrenko. Strijov, Vadim. Weber, Gerhard-Wilhelm. (2014). Sample size determination for logistic regression. *Journal of Computational and Applied Mathematics* 255; 743-752.
- 5- Bennett, N.D. Croke, B.F. Guariso, G. Guillaume, J.H. Hamilton, S.H. Jakeman, A.J. Marsili-Libelli, S. Newham, L.T. Norton, J.P. Perrin, C. (2013). Characterizing performance of environmental models. *Environ. Model*, 4; 1-20.
- 6- Chen, W. Peng, J. Hong, H. Shahabi, H. Pradhan, B. Liu, J. Zhu, A.X. Pei, X. Duan, Z. (2018). Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren county, Jiangxi province, China. *Sci. Total Environ*, 626; 1121-1135.
- 7- Hoang, N.D. Nguyen, Q.L. Bui, D.T. Image Processing-Based Classification of Asphalt Pavement Cracks Using Support Vector Machine Optimized by Artificial Bee Colony. (2018). *J. Comput. Civ. Eng*, 32; 04018037.
- 8- Hong, H. Shahabi, H. Shirzadi, A. Chen, W. Chapi, K. Ahmad, B.B. Roodposhti, M.S. Hesar, A.Y. Tian, Y. Bui, D.T. (2019). Landslide susceptibility assessment at the Wuning area, China: A comparison between multi-criteria decision making, bivariate statistical and machine learning methods. *Nat. Hazards*, 96; 173-212.
- 9- Hota, H. S. Handa, R Shrivastava, A. K. (2017). Time series data prediction using sliding window based RBF neural network. *International Journal of Computational Intelligence Research*, 13(5); 1145-1156.
- 10- Kiani, S. Karimkhani, A. Mazidi, A. (2021). Landslide Risk Zoning using Logistic Regression and ANFIS Models in Hashtjin Catchment Area in Ardabil Province. *Geography and Environmental Sustainability*. 11(2); 55-73. doi: 10.22126/ges.2021.6461.2391. (In Persian).
- 11- Kornejady, A. Pourghasemi, H. R. (2019). 'Landslide Susceptibility Assessment Using Data Mining Models, A case study: Chehel-Chai Basin', *Watershed Engineering and Management*, 11(1); 28-42. doi: 10.22092/ijwmse.2019.118436. (In Persian).
- 13- Mahdavi-fard, M, Jafari, S. Valizadeh Kamran, K, Karimzadeh, S. (2022). Rapid monitoring of mangrove cover changes using support vector machine algorithm in Google Earth Engine

- computing platform (Case study: Qeshm mangrove forests). *Remote Sensing and GIS Applications in Environmental Sciences*, 2(2); 36-23. (In Persian).
- 14- Mersha, T. Meten, M. (2020). GIS-based landslide susceptibility mapping and assessment using bivariate statistical methods in Simada area, northwestern Ethiopia. *Geoenvironmental disasters*, 7(1); 1-22.
- 15- Miao, T. Wang, M. Susceptibility analysis of earthquake-induced landslide using random forest method. *In proceedings of the international conference on computer information systems and industrial applications*, Bangkok, Thailand, 28–29 June 2015; 1–1006. 136.
- 16- Ngo, T. Q. Dam, N. D. Al-Ansari, N. Amiri, M. Phong, T. V. Prakash, I. Pham, B. T. (2021). Landslide susceptibility mapping using single machine learning models: a case study from Pithoragarh District, India. *Advances in civil engineering*.
- 17- Niazi, Y. Mendoza, ME. Talebi, A. Bidaki, H. (2021). GIS-based support vector machine model in shallow landslide hazards prediction: A case study on Ilam dam watershed, Iran. *Journal of Nature and Spatial Sciences (JONASS)*. 1(1); 59-84.
- 18- Pardeshi, SD. Autade, SE. Pardeshi, SS. (2013). Landslide hazard assessment: recent trends and techniques. *SpringerPlus*. 2(1); 1-1.
- 19- Su, Q. Zhang, J. Zhao, S. Wang, L. Liu, J. Guo, J. (2017). Comparative assessment of three nonlinear approaches for landslide susceptibility mapping in a coal mine area. *ISPRS International Journal of Geo-Information*. 6(7); 228.
- 20- Sukristiyanti, S. Wikantika, K. Sadisun, IA. Yayusman, LF. Soebowo, E. (2021). Preliminary Study of Landslide Susceptibility Modeling with Random Forest Algorithm Using R (Case Study: The Cisangkuy Sub-watershed). *In IOP Conference Series: Earth and Environmental Science*. 1(1); 012015. IOP Publishing.
- 21- Taheri, V. Karam, A. Safari, A. Shataii joibari, S. (2020). 'Landslides zoning of Golestan province using combination algorithm least squares support vector machines and artificial bee colony', *Geographical Planning of Space*, 10(37); 213-230. doi: 10.30488/gps.2021.177771.3022. (In Persian).
- 22- Tehrani, FS. Calvello, M. Liu, Z. Zhang, L. Lacasse, S. (2022). Machine learning and landslide studies: Recent advances and applications. *Natural Hazards*. 20; 1-49.
- 23- Tien Bui, D. Shahabi, H. Shirzadi, A. Chapi, K. Alizadeh, M. Chen, W. Mohammadi, A. Ahmad, B. Panahi, M. Hong, H. (2018). Landslide detection and susceptibility mapping by airsar data using support vector machine and index of entropy models in Cameron highlands, Malaysia. *Remote Sens*. 10; 1527.
- 24- Tsangaratos, P. Ilija, I. (2017). Applying machine learning algorithms in landslide susceptibility assessments. *In Handbook of Neural Computation*. 1: 433-457. Academic Press.

- 25- Umar, Z. Pradhan, B. Ahmad, A. Jebur, M.N. Tehrany, M.S. (2014). Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena*. 118; 124-135.
- 26- Wang, L.J. Guo, M. Sawada, K. Lin, J. Zhang, J. (2016). A comparative study of landslide susceptibility maps using logistic regression, frequency ratio, decision tree, weights of evidence and artificial neural network. *Geosciences Journal*. 20(1); 117-36.
- 27- Wubalem, A. (2021). Landslide susceptibility mapping using statistical methods in Uatzau catchment area, northwestern Ethiopia. *Geoenvironmental Disasters*. 8(1); 1-21.
- 28- Yi, Y. Zhang, Z. Zhang, W. Xu, C. (2019). Comparison of different machine learning models for landslide susceptibility mapping. In IGARSS 2019. *IEEE International Geoscience and Remote Sensing Symposium*. 9318-9321. IEEE.
- 29- Zhou, X. Wu, W. Lin, Z. Zhang, G. Chen, R Song, Y. Wang, Z. Lang, T. Qin, Y. Ou, P. Huangfu, W. (2020). Landslide risk zoning in Ruijin, Jiangxi, China. *Natural Hazards and Earth System Sciences Discussions*. 5; 1-21.