

# بررسی کارایی مدل‌های هوش مصنوعی و آماری دو متغیره در تعیین مناطق حساس به وقوع زمین لغزش در استان آذربایجان غربی

آزاد آرام<sup>۱</sup>، محمدرضا دلایان<sup>۲\*</sup>، سیامک ساعدی<sup>۳</sup>، امید رفیعیان<sup>۴</sup> و صمد دربندی<sup>۵</sup>

- (۱) دانشجوی دکتری، گروه علوم و مهندسی خاک، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.
  - (۲) استادیار، گروه علوم و مهندسی خاک، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.
  - (۳) استادیار، گروه علوم و مهندسی خاک، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.
  - (۴) استادیار، گروه محیط‌زیست، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.
  - (۵) استادیار، گروه علوم و مهندسی آب، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.
- \* ایمیل نویسنده مسئول: mdalalian@iaut.ac.ir



شاپا چاپی: ۲۴۸۰-۲۲۵۱  
شاپا الکترونیکی: ۷۴۰۰-۲۲۵۰

نشریه حفاظت منابع آب و خاک

آدرس تارنما:

<https://wsrj.srbiau.ac.ir>

پست الکترونیک:

[iauwsrcj@srbiau.ac.ir](mailto:iauwsrcj@srbiau.ac.ir)  
[iauwsrcj@gmail.com](mailto:iauwsrcj@gmail.com)

سال یازدهم

شماره چهار

تابستان ۱۴۰۱

تاریخ دریافت:

۱۴۰۰/۱۱/۲۱

تاریخ پذیرش:

۱۴۰۱/۰۲/۳۰

صفحات: ۶۳-۷۴

## چکیده:

**زمینه و هدف:** زمین‌لغزش، یکی از مخاطرات طبیعی است که منجر به خسارات جانی و مالی فراوان می‌شود. پژوهشگران در موضوع حساسیت به وقوع زمین‌لغزش، به بررسی احتمال وقوع زمین‌لغزش با توجه به شرایط توپوگرافی و ژئومحیطی می‌پردازند و اطلاعات به‌دست آمده، در مدیریت خطر زمین‌لغزش حیاتی است. تهیه نقاط حساس به وقوع زمین‌لغزش یک ابزار ضروری برای ارزیابی خطر زمین‌لغزش بوده و در برنامه‌ریزی و مدیریت بهتر این مناطق بسیار کاربردی است. در این پژوهش مدل‌های مبتنی بر هوش مصنوعی و دو متغیره آماری در تعیین نقاط حساس به زمین‌لغزش در استان آذربایجان غربی مورد بررسی و مقایسه قرار گرفته است.

**روش پژوهش:** برای تهیه نقاط حساس به وقوع زمین‌لغزش در استان آذربایجان غربی که در شمال غربی ایران واقع شده است، از روش‌های مبتنی بر هوش مصنوعی و دو متغیره آماری بهره گرفته شد. این مطالعه در چهار مرحله صورت گرفت. مرحله اول شامل مطالعه زمین‌لغزش‌های منطقه بر اساس بانک اطلاعات سازمان جنگل‌ها، مراتع و آبخیزداری ایران (FRWO) و شناسایی ۱۱۰ زمین‌لغزش با بررسی‌های میدانی، تفسیر عکس‌های هوایی و تصاویر ماهواره‌ای گوگل ارث، مرحله دوم جمع‌آوری داده‌ها و ایجاد پایگاه داده‌های مکانی فاکتورهای مؤثر، مرحله سوم به‌کارگیری روش نسبت فراوانی (FR)، آنتروپی شانون (SE)، بگینگ (BA)، جنگل تصادفی (RF) و مدل ترکیبی جنگل‌های تصادفی و بگینگ (RF-BA) و مرحله چهارم: اعتبارسنجی روش‌ها با استفاده از روش منحنی مشخصه عملکرد سیستم (ROC) بود. بر اساس بررسی‌های میدانی و مطالعات مشابه، ۱۲ عامل مؤثر بر وقوع زمین‌لغزش شامل ارتفاع، زاویه شیب، جهت شیب، فاصله از گسل، فاصله از رودخانه، فاصله از جاده، تراکم زهکشی، تراکم جاده، بارندگی، خاک، کاربری زمین و سنگ‌شناسی شناسایی شد. در بررسی‌های میدانی، ۱۱۰ زمین‌لغزش در استان آذربایجان غربی مشخص شد. ۷۰ درصد از داده‌ها به‌طور تصادفی انتخاب و برای مدل‌سازی مورد استفاده قرار گرفتند و ۳۰ درصد داده‌ها برای اعتبارسنجی استفاده گردید.

**یافته‌ها:** در میان جهت‌های جغرافیایی، جهت جنوبی با وزن ۱/۴۹ دارای بیش‌ترین تأثیر بر وقوع زمین‌لغزش‌های استان بود. کمترین وزن نیز مربوط به مناطق مسطح بود که در آن هیچ‌گونه لغزشی رخ نداده است. نتایج فاکتور شیب نشان داد که شیب‌های میانی دارای بیشترین تأثیر بر وقوع زمین‌لغزش است، به‌طوری که در شیب‌های کم به دلیل وجود جاذبه کم، زمین‌لغزش کم‌تر رخ می‌دهد و شیب‌های بسیار تند نیز مربوط به مناطق کوهستانی بوده که با سنگ پوشیده شده و خاک بسیار نازکی وجود دارد که برای لغزش مناسب نمی‌باشد. بررسی عامل کاربری اراضی نشان داد که ۴۸ درصد از لغزش‌ها در مناطق کشاورزی رخ می‌دهد. بر طبق بررسی‌های این پژوهش، بیشتر زمین‌لغزش‌ها در نزدیکی رودخانه‌ها و گسل‌ها رخ داده است. همچنین در بعضی مناطق، نزدیک‌ترین فواصل به جاده، بیشترین خطر را برای زمین‌لغزش دارد.

**نتایج:** نتایج این تحقیق نشان داد که مدل‌های هوش مصنوعی (جنگل تصادفی RF و مدل ترکیبی جنگل‌های تصادفی و بگینگ RF-BA) دارای کارایی بالاتری نسبت به مدل‌های آماری (نسبت فراوانی FR و آنتروپی شانون SE) است. دقت مدل‌های ترکیبی بیشتر از مدل‌های منفرد بود. نتایج منحنی ROC دقت ۰/۹۲، ۰/۹۱، ۰/۸۹ و ۰/۸۸ را با مدل‌های RF-BA، RF، FR و SE نشان داد.

**کلیدواژه‌ها:** آنتروپی شانون، بگینگ، جنگل تصادفی، نسبت فراوانی، هوش مصنوعی



## مقدمه

فراوانی، وزن‌دهی شواهد، الگوریتم مبتنی بر تئوری بیزین و درخت تصمیم، الگوریتم درخت تصمیم، ماشین پشتیبان بردار، الگوریتم لجستیک درختی، الگوریتم طبقه‌بندی و رگرسیون درختی، الگوریتم‌های ترکیبی مثل ترکیب الگوریتم‌های درختی با الگوریتم‌های توابعی و درخت تصمیم و غیره می‌باشد. امروزه تکنیک‌های یادگیری ماشین که بخشی از تکنیک‌های هوش مصنوعی محسوب می‌گردد، از میان روش‌های احتمالی بیشتر مورد توجه واقع شده است (Bui et al., 2015; Pham et al., 2015b; Pham et al., 2016b; Pham et al., 2017b; Shirzadi et al., 2017a; Chen et al., 2017a; Pourghasemi and Rahmati, 2018).

در پژوهشی نگو<sup>۱</sup> و همکاران (۲۰۲۱) از دو الگوریتم شبکه عصبی بازگشتی (RNN) و شبکه عصبی کانولوشنال (CNN) برای تعیین نقاط حساس به زمین لغزش در ایران استفاده نمودند. یافته‌های آن‌ها نشان داد که حدود ۳۱ درصد از شهرهای ایران در مناطقی با حساسیت زمین لغزش زیاد و بسیار زیاد قرار دارد. در مطالعه‌ای، چن<sup>۲</sup> و همکاران (۲۰۲۱) جهت مدل‌سازی حساسیت زمین لغزش نتیجه گرفتند که مدل هیبریدی سیستم استنتاج عصبی فازی تطبیقی - الگوریتم مرغ آلاچیق سازی ANFIS-SBO مقادیر خطای کمتری را برای مجموعه داده اعتبارسنجی نشان می‌دهد و الگوریتم SBO قابلیت بهینه‌سازی بهتری نسبت به مدل سیستم استنتاج عصبی فازی تطبیقی - الگوریتم بهینه‌سازی مبتنی بر یادگیری و آموزش ANFIS-TLBO داشته است. در پژوهشی Youssef and Pourghasemi (2021) به تهیه نقشه حساسیت زمین لغزش با استفاده از الگوریتم‌های یادگیری ماشین و مقایسه عملکرد آن‌ها پرداختند. یافته‌ها نشان داد که مدل‌های جنگل تصادفی (RF) و آنالیز تشخیص خطی (LDA) بهترین عملکرد را در مقایسه با سایر مدل‌ها دارد. اهداف اصلی این پژوهش شامل تعیین عوامل مؤثر در وقوع زمین لغزش، تعیین نقاط حساس به زمین لغزش با استفاده از مدل‌های هوش مصنوعی و آماری دو متغیره و مقایسه کارایی این مدل‌ها است.

زمین لغزش به عنوان یکی از شایع‌ترین مخاطرات طبیعی است که در مناطق کوهستانی رخ می‌دهد و منجر به تلفات جانی و مالی می‌شود (Fang et al., 2020). شناسایی مناطق مستعد زمین لغزش یک گام اساسی در ارزیابی خطر زمین لغزش و کاهش خسارات ناشی از آن است (Ngo et al., 2021). شناسایی نقاط حساس به زمین لغزش در مدیریت ریسک و برنامه‌ریزی شهری بسیار مهم است (Gaidzik and Ramírez-Herrera, 2021). بر اساس زمین لغزش‌هایی که در گذشته و حال تحت شرایط توپوگرافی، زمین‌شناسی، هیدروژئولوژی و اقلیمی در یک منطقه اتفاق افتاده است، می‌توان اطلاعات مفیدی را برای پیش‌بینی پتانسیل وقوع زمین لغزش‌های آینده در آن منطقه تهیه نمود (Raghuvanshi et al., 2014).

مدل‌ها و روش‌های بسیاری برای بررسی حساسیت و خطر وقوع زمین لغزش وجود دارد. روش‌های کیفی، کمی، ابتکاری، آماری و قطعی از جمله‌ی این روش‌ها است - روش‌های آماری امکان همبستگی میان عوامل مؤثر که منجر به ناپایداری دامنه می‌شود با زمین لغزش‌هایی که در گذشته و حال حاضر اتفاق افتاده است را فراهم می‌آورد. این روش‌ها معمولاً نیازمند تعداد داده‌های زیادی بوده و معمولاً برای تعیین نقاط حساس و نیز خطر وقوع زمین لغزش در مقیاس منطقه‌ای مناسب است که به دو روش آماری دو متغیره و روش آماری چند متغیره تقسیم می‌شود (Sassa and Canuti, 2008). روش‌های قطعی (روش‌های در یک مقیاس خاص یا مقیاس دامنه) بر اساس مدل‌های فیزیکی و ریاضی در موضوع ناپایداری دامنه‌ها استوار است. چالش اصلی این روش نیازمند بودن به داده‌هایی همچون خصوصیت‌های مکانیکی خاک، سطح ایستابی، فشار آب منفذی، زاویه اصطکاک داخلی خاک، وزن واحد آب و خاک و غیره می‌باشد که تهیه این داده‌ها در یک مقیاس بزرگ‌تر از دامنه مثل یک حوزه آبخیز دشوار و هزینه‌بر می‌باشد (Terlien et al., 1995). در مدل‌سازی زمین لغزش، بیشتر روش‌ها و الگوریتم‌های توسعه داده شده متعلق به روش‌های رگرسیون لجستیک، روش آماری دو متغیره و چند متغیره، نسبت

<sup>1</sup> Ngo  
<sup>2</sup> Chen

## مواد و روش‌ها

### معرفی منطقه مورد مطالعه

سول<sup>۱</sup> و تا حدودی انتی‌سول<sup>۲</sup> و بخش غربی استان نیز عمدتاً دارای خاک از نوع انتی‌سول همراه بیرون‌زدگی سنگی می‌باشد. شکل ۱ موقعیت استان آذربایجان غربی در نقشه ایران را نشان می‌دهد.



شکل ۱. موقعیت جغرافیایی استان آذربایجان غربی در نقشه ایران

استان آذربایجان غربی واقع در شمال غربی ایران، بین ۳۵ درجه و ۵۸ دقیقه تا ۳۹ درجه و ۴۶ دقیقه عرض شمالی و ۴۴ درجه و ۳ دقیقه تا ۴۷ درجه و ۲۳ دقیقه طول شرقی واقع شده است و دارای آب و هوای معتدل، نیمه‌خشک و نیمه مرطوب کوهستانی تحت تأثیر جریان‌های مرطوب اقیانوس اطلس و مدیترانه می‌باشد. این استان بدون احتساب مساحت دریاچه ارومیه، ۳۷۴۱۱ کیلومتر مربع وسعت دارد. از نظر توپوگرافی، حداقل و حداکثر ارتفاع از سطح دریا به ترتیب ۶۰۵ و ۳۶۰۰ متر می‌باشد. حداقل و حداکثر درجه شیب به ترتیب صفر تا ۶۲ درجه می‌باشد. میانگین بارندگی سالیانه ۴۵۰ میلی‌متر در سطح استان متغیر می‌باشد. حداکثر، حداقل و میانگین سالیانه دما به ترتیب، ۳۱/۷ درجه سانتی‌گراد، ۶/۹- درجه سانتی‌گراد و ۱۰/۹ درجه سانتی‌گراد می‌باشد. از نظر اقلیم بر طبق تقسیم‌بندی اقلیمی دومارتن با ضریب ۱۶/۵ اقلیم منطقه مورد مطالعه نیمه خشک و بر اساس تقسیم‌بندی اقلیمی آمبرژه با ضریب ۳۱/۲ اقلیم از نوع نیمه خشک سرد می‌باشد (حنفی و حاتمی، ۱۳۹۲).

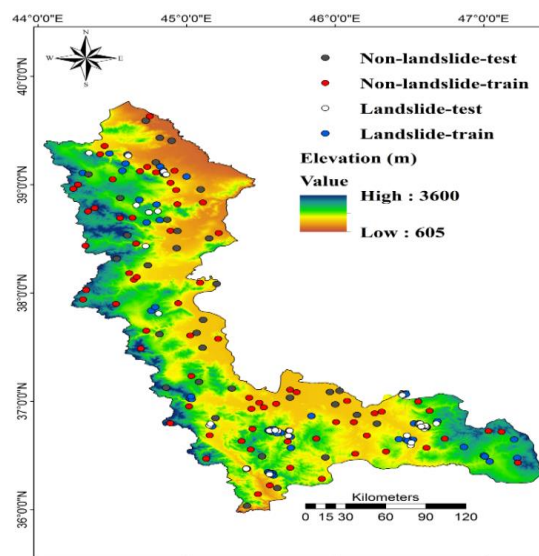
از نظر کاربری اراضی، بیشتر استان تحت پوشش کاربری باغ و زراعی با غالبیت کشت دیم از نوع گندم می‌باشد. از نظر ژئومورفولوژی، منطقه مورد مطالعه در بین دو گسل اصلی تبریز در شرق و ارومیه یا زرینه‌رود در غرب واقع است. در این ناحیه یکنواختی و سستی جنس سنگ‌ها، اشکال مشابهی را به وجود آورده که از ویژگی‌های آن، کوه‌های گنبدی با دامنه‌های ملایم، یکنواخت و دره‌های باز است. از ویژگی‌های ناهمواری‌های این بخش، وجود حصارهای کوهستانی از سنگ‌های دگرگونی و رسوبی است که دشت‌های مرتفع هموار یا تپه ماهوری را احاطه کرده است. از نظر خاک-شناسی، بخش غربی استان عمدتاً دارای خاک‌های اریدی-

### روش انجام تحقیق

این مطالعه در چهار مرحله برای تهیه نقاط حساس به وقوع زمین‌لغزش با استفاده از روش‌های مبتنی بر هوش مصنوعی و دو متغیره آماری انجام گرفت که عبارتند از مرحله اول: مطالعه زمین‌لغزش‌های منطقه بر اساس بانک اطلاعات سازمان جنگل‌ها، مراتع و آبخیزداری ایران (FRWO) و شناسایی ۱۱۰ زمین‌لغزش با بررسی‌های میدانی، تفسیر عکس‌های هوایی و تصاویر ماهواره‌ای گوگل ارث (شکل ۲)، مرحله دوم، جمع-آوری داده‌ها و ایجاد پایگاه داده‌های مکانی فاکتورهای مؤثر در نرم‌افزار ArcGIS 10.2، مرحله سوم: به‌کارگیری روش نسبت فراوانی (FR)، آنتروپی شانون، بگینگ (BA)، جنگل تصادفی (RF) و ترکیبی (RF-BA) و مرحله چهارم: اعتبارسنجی روش‌ها با استفاده از روش منحنی مشخصه عملکرد سیستم (ROC) بود.

<sup>1</sup>Inceptisol  
<sup>2</sup>Entisol

اگر داده‌های آموزشی یا تعلیمی (S) مثل درجه شیب، جهت شیب و غیره با تعداد m برچسب از نمونه‌ها (S = (x<sub>i</sub>, y<sub>i</sub>))، به طوری که  $x_i \in R$  و (وقوع و عدم وقوع زمین لغزش)  $y_i \in R$ ، القاء کننده گو تعداد نمونه‌های خود راه انداز N باشد. این الگوریتم با تکرار و جایگذاری منظم در داده‌های آموزشی دسته‌ای از داده‌های آموزشی جدید (S') را بر اساس تابع توزیع احتمال (رابطه ۲) ایجاد می‌کند که احتمال تکرار یا مشابهت برچسب یا رکوردها در داده‌های آموزشی برای این نوع الگوریتم وجود دارد. در مرحله بعد، برای هر کدام از دسته داده‌های آموزشی جدید تکرار شده با جایگذاری (S')، یک الگوریتم ضعیف پایه یا الگوریتم دسته‌بند (L<sub>i</sub>) اعمال می‌گردد. در نهایت الگوریتم‌های دسته‌بند مختلف به دست آمده از هر داده آموزشی تکرار شده با هم ترکیب تا یک الگوریتم دسته‌بند (L\*) قوی‌تر از الگوریتم پایه ضعیف با استفاده از شمارش آراء در رابطه ۵، به دست آید (Piao et al., 2015). فرایند یادگیری الگوریتم بگینگ به صورت رابطه‌های ۱ تا ۵ (دستورهای اجرا شده در نرم‌افزار) نشان داده شده است.



شکل ۲. موقعیت جغرافیایی زمین لغزش‌های سطحی منطقه مورد مطالعه و تهیه آن در نرم‌افزار ArcGIS

### فاکتورهای مؤثر بر وقوع زمین لغزش‌های سطحی

عوامل اساسی مؤثر بر وقوع زمین لغزش شامل عوامل توپوگرافیکی (از قبیل زاویه و جهت شیب، ارتفاع از سطح دریا)، عوامل هیدرولوژیکی و اقلیمی (بارندگی، فاصله از آبراهه‌ها یا رودخانه‌ها، تراکم آبراهه یا تراکم زهکشی)، عوامل زمین‌شناسی (لیتولوژی یا جنس سنگ بستر، فاصله از گسل و تراکم گسل)، عامل پوشش زمین (کاربری اراضی)، عوامل انسانی (فاصله از جاده و تراکم جاده) می‌باشد.

### الگوریتم‌های ترکیبی

#### مدل بگینگ

مدل بگینگ<sup>۱</sup> یک طراحی ساده و در عین حال مؤثر

است که چندین یادگیرنده را ترکیب می‌کند تا یک

مجموعه برای پیش‌بینی تشکیل دهد (Liang et al.,

2011). این روش که توسط بریمن در سال ۱۹۹۶ ارائه

شده است، یک روش مجموعه موازی کارآمد است که

عملکرد پیش‌بینی کننده‌های ناپایدار را بهبود می‌بخشد

(Liu et al., 2019).

**Input:** training dataset, Inducer, Integer (number of bootstrap samples)  
**Output:** Aggregated classifier L'  
**Start:**  
 (1) For i=1 to N {  
 (2) S' bootstrap sample from (sample with replacement)  
 (3) L<sub>i</sub> = ζ(S')  
 (4) }  
 (5) L'(x) argmax<sub>y∈Y</sub> = ∑<sub>i:C(x)=y</sub> 1 (the most often predicted label y)  
**Output:** classifier L\*

### مدل جنگل‌های تصادفی (RF)

این الگوریتم توسط رودریگز<sup>۲</sup> و همکاران (۲۰۰۶) ارائه شد. این روش، مجموعه‌ای از درختان تصمیم باینری آموزش داده شده جداگانه است. چندین درخت تصمیم از طریق زیرمجموعه‌های داده‌های مختلف ایجاد می‌شود و

<sup>2</sup> Rodriguez

<sup>1</sup> BAGGING

که در آن FR میزان تأثیر هر کلاس از هر پارامتر،  $F_i$  درصد نقاط قرارگرفته در کلاس  $i$  و  $P_i$  درصد پیکسل‌های کلاس  $i$  در کل حوضه آبخیز است و از روابط (۷) و (۸) بدست می‌آید:

$$F_i = \frac{n}{N} \times 100 \quad (7)$$

$$P_i = \frac{a}{A} \times 100 \quad (8)$$

که در آن  $n$ : تعداد نقاط لغزش در کلاس  $i$ ،  $N$ : تعداد کل نقاط لغزش،  $a$ : تعداد پیکسل‌های کلاس  $i$  و  $A$ : تعداد کل پیکسل‌ها می‌باشد. در نهایت نرخ‌های به دست آمده برای هر کلاس در نرم‌افزار ArcGIS 10.2 در لایه‌های مربوطه اعمال گردیده و با استفاده از تابع روی هم‌گذاری، نقاط حساس به لغزش استان آذربایجان غربی به دست خواهد آمد.

#### روش آنتروپی شانون

در یک سیستم، آنتروپی اندازه‌گیری عدم قطعیت، نوسانات، رفتار نامتعادل، توزیع انرژی و بی‌ثباتی است (Khosravi et al., 2016). این نظریه ابتدا توسط بولتزمن ارائه شد و سپس توسط شانون (۱۹۴۸) گسترش یافت. مدل آنتروپی به‌طور گسترده در مطالعات شناسایی خطر و مدیریت ریسک برای محاسبه نسبت اهرم خطرات طبیعی و نشان دادن اولویت برخی از متغیرها در تأثیرگذاری بر یک خطر خاص استفاده شده است (Al-Hinai and Abdalla, 2020). شاخص آنتروپی، مقیاسی از مقدار برابری است که در آن گروه‌ها به‌طور مساوی و یکنواخت در بین واحدهای سازمانی توزیع شده است (Malekinezhad et al., 2021). این شاخص به‌عنوان مقیاسی از تفاضل متوسط سهم گروه‌های واحد از سیستم کل تعریف شده است (Arora et al., 2021).

$$E_{ij} = \frac{FR}{\sum_{j=1}^{M_j} FR} \quad (9)$$

$$H_j = -\sum_{i=1}^{M_j} E_{ij} \log_2 E_{ij}, j = 1, \dots, n \quad (10)$$

نتایج قضاوت از آن‌ها برای به دست آوردن نتایج خروجی RF بررسی می‌شود (Sun et al., 2021). جنگل‌های تصمیم تصادفی یک روش یادگیری ترکیبی برای دسته‌بندی و رگرسیون می‌باشد که بر اساس ساختاری متشکل از شمار بسیاری درخت تصمیم، بر روی زمان آموزش و خروجی کلاس‌ها (کلاس‌بندی) یا برای پیش‌بینی‌های هر درخت به شکل مجزا، کار می‌کند. جنگل‌های تصادفی برای درختان تصمیم که در مجموعه آموزشی دچار بیش‌برازش می‌شود، مناسب است. عملکرد جنگل تصادفی معمولاً بهتر از درخت تصمیم است، اما این بهبود عملکرد تا حدی به نوع داده هم بستگی دارد (Piryonesi and El-Diraby, 2020).

#### روش‌های آماری دو متغیره

##### روش نسبت فراوانی (FR)

روش نسبت فرکانس یک مدل احتمالی است که قابل درک بوده و استفاده از آن آسان است و به این علت به‌طور گسترده استفاده می‌شود (Arca et al., 2020). مدل نسبت فراوانی را می‌توان برای کمی کردن رابطه پیش‌بینی بین متغیرهای وابسته و مستقل استفاده کرد. روش فوق یک روش تحلیل آماری پایه است که تأثیر هر دسته از عوامل شرطی‌سازی را بر سیل پیش‌بینی‌شده ارزیابی می‌کند (Sarkar et al., 2021). مجموعه نقاط لغزش گروه واسنجی به‌عنوان متغیر وابسته و پارامترهای انتخاب‌شده (طبقات ارتفاعی، شیب و...) به‌عنوان متغیرهای مستقل به تکنیک نسبت فراوانی معرفی می‌شود. با استفاده از تکنیک نسبت فراوانی، احتمال رخداد لغزش قرار گرفته در هر کلاس برای تمام پارامترها محاسبه می‌گردد. تکنیک نسبت فراوانی نشان‌دهنده احتمال حضور یک پدیده با مشخصاتی ویژه است. به‌منظور تعیین میزان تأثیر هر کلاس از هر متغیر مستقل از رابطه (۶) استفاده می‌گردد:

$$FR = \frac{F_i}{P_i} \quad (6)$$

آینده بوده و نشان‌دهنده توانایی پیش‌بینی مدل‌ها می‌باشد. ارتباط کمی - کیفی بین AUC و دقت پیش‌بینی برابر ۰/۶ - ۰/۵ (ضعیف)، ۰/۷ - ۰/۶۱ (متوسط)، ۰/۷۱ - ۰/۸ (خوب)، ۰/۹ - ۰/۸۱ (بسیار خوب) و ۰/۹۱ - ۱ (بسیار عالی) می‌باشد (Pourghasemi et al., 2013).

### بحث و نتایج

بررسی وضعیت مناطق زمین لغزشی بر اساس نسبت فراوانی و آنتروپی شانون

نتایج به دست آمده از روش FR نشان‌دهنده وزن‌های به دست آمده برای هر یک از کلاس‌های فاکتورهای مؤثر است. FR برای ۱۲ فاکتور مؤثر، بر طبق روابط آن‌ها با وقوع لغزش محاسبه گردید. نسبت مساحتی که لغزش در آنجا رخ داده به کل مساحت محاسبه گردید. هرچه نسبت فراوانی بیشتر باشد، نشان‌دهنده ارتباط قوی بین فاکتورهای مؤثر و رخداد لغزش خواهد بود و برعکس. به طوری که هرچه از ۱ بیشتر باشد، نشان‌دهنده ارتباط قوی و هرچه از ۱ کمتر باشد، نشان‌دهنده ارتباط ضعیف می‌باشد. نتایج نشان داد که در جهت‌های مختلف جغرافیایی، جهت جنوبی با وزن ۱/۴۹ دارای بیشترین تأثیر بر وقوع زمین لغزش‌های استان آذربایجان غربی دارد. کمترین وزن نیز مربوط به مناطق مسطح و با شیب بسیار اندک (کمتر از ۲ درجه) می‌باشد که در آن هیچ‌گونه لغزشی رخ نداده است. نتایج فاکتور شیب نشان داد که شیب‌های میانی (۱۱ تا ۱۷ درجه) دارای بیشترین تأثیر بر وقوع زمین لغزش بوده، به طوری که در شیب‌های کم، به دلیل وجود جاذبه کم زمین لغزش کمتر رخ می‌دهد و شیب‌های بسیار تند نیز مربوط به مناطق کوهستانی بوده که با سنگ پوشیده شده و خاک بسیار نازکی وجود دارد که برای لغزش مناسب نمی‌باشد. فاکتور مهم و مؤثر دیگر بر وقوع زمین لغزش، طبقات ارتفاعی می‌باشد. با افزایش طبقات ارتفاعی احتمال وقوع زمین لغزش به دلیل کاهش عمق خاک کاهش می‌یابد. نتایج فاکتور بارندگی نشان داد که به طور کلی هرچه مقدار بارندگی بیشتر باشد (مناطق

که در آن FR نسبت فراوانی و  $E_{ij}$  چگالی احتمال می‌باشد.

$$H_{j_{\max}} = \log_2 M_j, M_j = \text{number of classes} \quad (11)$$

$$I_j = (H_{j_{\max}} - H_j / H_{j_{\max}}), I = (0, 1), j = 1, \dots \quad (12)$$

$$V_j = I_j FR \quad (13)$$

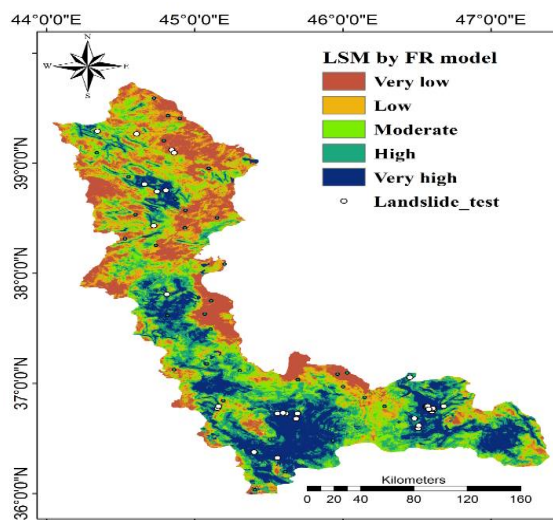
که  $H_j$  و  $H_{j_{\max}}$  مقادیر (ارزش‌های) آنتروپی،  $I_j$  ضریب اطلاعات و  $M_j$  تعداد کلاس‌ها می‌باشد. همچنین  $V_j$  نشان‌دهنده ارزش وزن برآیند از پارامتر کلی می‌باشد.

در تحقیق حاضر، اعتبارسنجی از طریق شاخص عملکرد منحنی مشخصه (ROC) صورت گرفت. منحنی فوق یک روش کمی برای بررسی و ارزیابی عملکرد کلی مدل‌ها می‌باشد. این معیار یک روش استاندارد و مناسب برای صحت‌سنجی کیفیت مدل زمین لغزش می‌باشد. در این روش، مساحت زیر منحنی ROC دارای مقادیر بین ۰/۵ تا ۱ بوده و از آن برای ارزیابی دقت مدل استفاده می‌شود (Khosravi et al., 2018). هرچه مساحت به عدد یک نزدیک‌تر باشد، مدل دارای دقت بالاتری می‌باشد. از ۷۰ درصد داده‌ها به عنوان داده‌های آموزشی و ۳۰ درصد باقی‌مانده برای اعتبارسنجی استفاده شد. شاخص درصد مساحت زیر این منحنی (AUCROC) بر اساس معادله زیر قابل محاسبه می‌باشد:

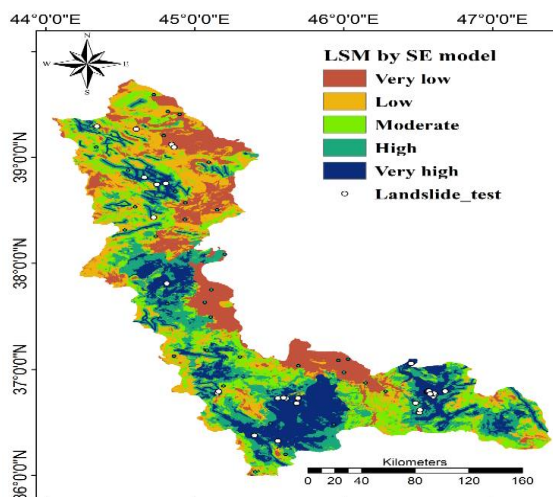
$$AUROC = \sum TP + \sum \frac{TN}{P} + N \quad (14)$$

که در آن، TP: تعداد پیکسل زمین لغزش‌هایی (کد ۱) که به درستی به عنوان زمین لغزش (کد ۱) طبقه‌بندی شده است، TN: تعداد پیکسل عدم حضور زمین لغزش‌هایی (کد صفر) که به درستی به عنوان عدم حضور زمین لغزش (کد صفر) طبقه‌بندی شده است، P و N به ترتیب تعداد کل پیکسل‌های زمین لغزش و عدم حضور زمین لغزش می‌باشد. اگر شاخص ROC بر اساس داده‌های آموزشی ترسیم گردد، این منحنی، منحنی نرخ موفقیت نامیده می‌شود. درحالی‌که بر اساس داده‌های صحت‌سنجی به آن منحنی نرخ پیش‌بینی می‌گویند. منحنی نرخ موفقیت بیان‌گر احتمال وقوع زمین لغزش بر اساس مدل ساخته شده با داده‌های آموزشی در زمان حال می‌باشد، در حالی که منحنی نرخ پیش‌بینی بیان‌گر احتمال وقوع زمین لغزش در

و تراکم آن‌ها از مهم‌ترین فاکتورهای مؤثر بر وقوع زمین‌لغزش در استان آذربایجان غربی می‌باشد. نقاط حساس به وقوع زمین‌لغزش توسط مدل آنتروپی شانون در سطح استان در شکل (۳-ب) که در نرم‌افزار ArcGIS 10.2 ترسیم شده است، نشان داده شده است. با توجه به هر دو شکل، نتایج بسیار شبیه به هم بوده و مناطق میانی و جنوبی استان بسیار مستعد وقوع زمین‌لغزش می‌باشد.



(الف)



(ب)

شکل ۳. تعیین نقاط حساس به لغزش استان آذربایجان غربی به- وسیله الف: روش نسبت فراوانی (RF) ب: روش آنتروپی شانون مقایسه کارایی الگوریتم‌های ترکیبی و منفرد

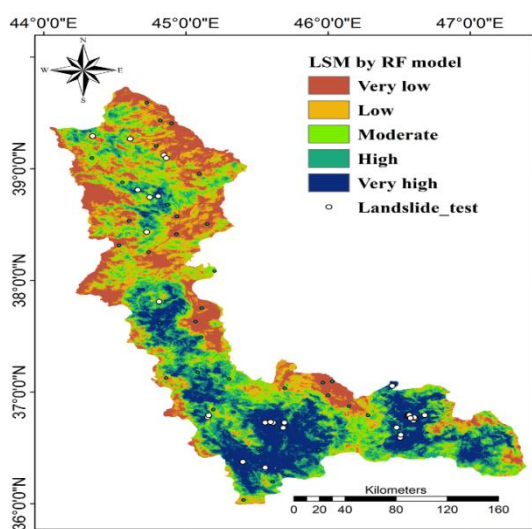
جنوبی استان)، احتمال رخداد زمین‌لغزش در استان نیز بیشتر می‌گردد. برای تأثیر گسل‌های منطقه بر وقوع زمین‌لغزش، از تأثیر فواصل متفاوت از گسل بر رخداد زمین‌لغزش‌ها استفاده گردید. نتایج جدول نسبت فراوانی (جدول ضمیمه) نشان داد که کلاس اول، یعنی فاصله ۰ تا ۵۰۰ متر با وزن ۵، دارای بیشترین تأثیر بر وقوع زمین‌لغزش‌های رخ داده در منطقه مورد مطالعه می‌باشد. برای تأثیر رودخانه‌های منطقه بر وقوع زمین‌لغزش، از تأثیر فاصله‌های متفاوت از رودخانه بر رخداد زمین‌لغزش‌ها استفاده شد. نتایج نشان داد که کلاس اول یعنی کمتر از ۵۰۰ متر، دارای بیشترین تأثیر بر رخداد زمین‌لغزش‌های استان می‌باشد، به طوری که رودخانه‌های دائمی منبع اصلی تأمین رطوبت مورد نیاز برای وقوع زمین‌لغزش‌ها می‌باشد. همچنین نتایج نشان داد که تراکم آبراهه و گسل دارای تأثیر بسیار زیادی بر وقوع زمین‌لغزش در استان آذربایجان غربی می‌باشد. به طور کلی هر چقدر تراکم و مقدار پوشش گیاهی بیشتر باشد، در نتیجه ریشه این گیاهان مانع از وقوع زمین‌لغزش می‌شود. نتایج مدل نسبت فراوانی نشان داد که مناطق جنگلی تنک دارای بیشترین تأثیر بر وقوع زمین‌لغزش در سطح استان می‌باشد. در این مناطق به دلیل خاک نفوذپذیر جنگلی، دامنه‌ها آب بسیار زیادی جذب کرده و به دلیل تنگی درختان، بسیار مستعد زمین‌لغزش می‌باشد. بررسی عامل کاربری اراضی نشان داد که ۴۸ درصد از لغزش‌ها در مناطق کشاورزی دیم رخ داده است که می‌توان دلیل آن را تغییر کاربری اراضی دانست، اما به دلیل مساحت زیاد کاربری کشاورزی دیم، وزن آن کمتر از مناطق جنگلی تنک در رخداد زمین‌لغزش می‌باشد. ترکیب خاک انتی سول/اریدیسول در بخش‌های میانی و جنوبی استان نیز دارای بیشترین تأثیر بر رخداد زمین‌لغزش در سطح استان می‌باشد. شکل (۳-الف) نقاط حساس به زمین‌لغزش را برای مدل FR نشان می‌دهد. فاکتورهای خاک، کاربری اراضی، فاصله از گسل، فاصله از رودخانه، فاصله از جاده

پژوهش استفاده گردید. نتایج مدل‌سازی مدل‌های داده کاوی RF و مدل ترکیبی Bagging-RF در جدول ۱ نشان داده شده است. نتایج حاکی از آن است که مدل ترکیبی دارای دقت و صحت بیشتری از مدل منفرد می‌باشد. به طوری که مقادیر حساسیت، شفافیت، صحت و ROC برای مدل ترکیبی از مقادیر مدل منفرد بیشتر بوده که حاکی از کارایی بهتر مدل ترکیبی در مقایسه با مدل منفرد در هر دو مرحله مدل‌سازی (آموزش) و مرحله ارزیابی (یا تست) می‌باشد. نقشه‌های تهیه شده به وسیله مدل‌های داده کاوی در شکل‌های (۴-الف) و (۴-ب) نشان داده شده است.

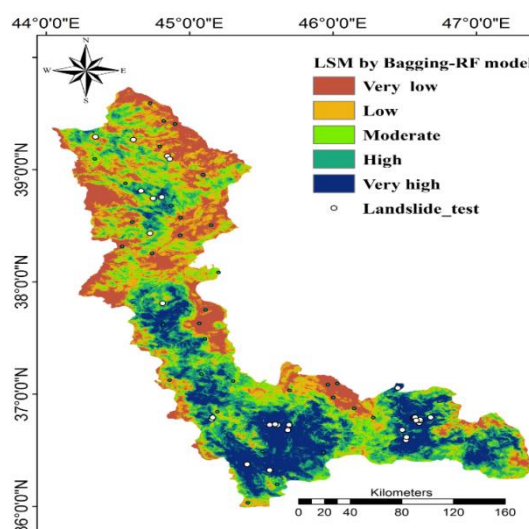
کارایی عملکرد الگوریتم‌های ترکیبی داده کاوی بستگی به انتخاب بهینه پارامترهای مورد استفاده آن الگوریتم از جمله تعداد تکرار برای اجرای مجدد مدل‌سازی با تعدادهای مختلف تکرار بر اساس داده آموزشی و تعداد seed (جهت تقسیم داده‌های آموزشی) دارد. در این پژوهش، از آنجا که این دو پارامتر هم راستا عمل می‌کند، ابتدا برای پارامتر تکرار، عدد ثابتی در نظر گرفته شد و سپس تعداد seed را تغییر داده و مقدار درصد مساحت زیر منحنی ROC (AUROC) ثبت گردید. از تکرار ۱۰ تا ۲۰ برای پارامتر تکرار و از seed برابر با ۱ تا ۱۰ در این

جدول ۱. مدل‌سازی الگوریتم‌های جنگل تصادفی و بگینگ-جنگل تصادفی

معیار	RF		Bagging	
	تست	آموزش	تست	آموزش
TP	۲۸	۷۳	۳۰	۷۵
TN	۲۹	۷۴	۳۱	۷۶
FP	۶	۴	۳	۲
FN	۵	۳	۲	۱
حساسیت	۰/۷۷۸	۰/۹۴۸	۰/۷۹۷	۰/۹۶۲
شفافیت	۰/۹۱۷	۰/۹۰۶	۰/۹۳۳	۱/۰۰۰
صحت	۰/۸۳۳	۰/۹۲۶	۰/۸۹۷	۰/۹۸۱
AUROC	۰/۸۹۸	۰/۹۸۴	۰/۹۲۳	۰/۹۹۳



(ب)



(الف)

شکل ۴. تعیین نقاط حساس به وقوع لغزش به وسیله روش الف: جنگل تصادفی (RF) ب: مدل ترکیبی Bagging-RF



### اعتبارسنجی نتایج

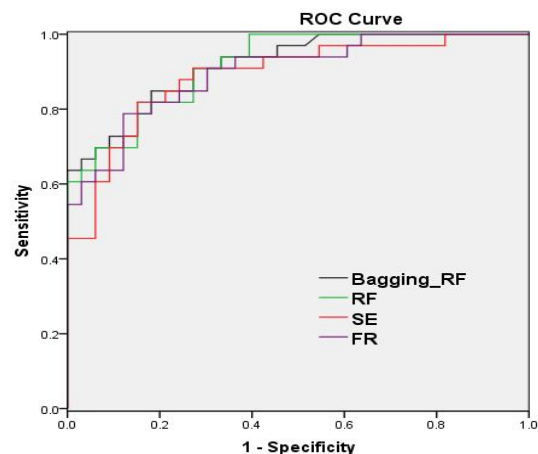
(شکل ۵). نتایج نشان داد که مدل ترکیبی هوش مصنوعی دارای کارایی بیشتری از بقیه مدل‌ها، حتی مدل منفرد هوش مصنوعی می‌باشد. همچنین نتایج نشان داد که مدل‌های هوش مصنوعی به‌طور کلی دارای دقت و صحت بالاتری نسبت به مدل‌های ساده آماری دو متغیره می‌باشد. مدل‌های آماری به دلیل ساختار ساده قادر به پیش‌بینی دقیق زمین‌لغزش‌های رخ داده نمی‌باشد. نتایج تحقیق حاضر با نتایج (Pourghasemi and Khosravi et al. (2019) and Rahmati (2018) که بیان نمودند روش‌های منطبق بر هوش مصنوعی و یادگیری ماشینی دارای صحت بالاتری نسبت به مدل‌های آماری است، منطبق است.

### نتیجه‌گیری

با توجه به مشاهده زمین‌لغزش‌ها در استان آذربایجان غربی، نتایج حاصل از تجزیه تحلیل آماری مدل نسبت فراوانی و آنتروپی شانون، مدل‌های هوش مصنوعی (جنگل تصادفی RF و بگینگ BA) و مدل ترکیبی RF-Bagging، نشان داد با افزایش ارتفاع تا یک حدی، حساسیت به لغزش بیشتر شده و دوباره حساسیت به زمین‌لغزش کاهش پیدا کرده است، زیرا در ارتفاعات بالا ممکن است به دلیل کوهستانی بودن منطقه، عمق خاک کم بوده و حالت صخره‌ای داشته باشد. بررسی عامل کاربری اراضی نشان داد که ۴۸ درصد از لغزش‌ها در مناطق کشاورزی دیم رخ داده است که می‌توان دلیل آن را تغییر کاربری اراضی دانست و چون در شیب‌های کم واقع شده، از دسترس بشر دور نبوده و دچار فرسایش می‌شود. از طرف دیگر، خاک اراضی به دلیل نزدیکی به رودخانه‌ها با جذب بالای آب و در صورت مساعد بودن سایر شرایط، باعث تشدید خطر وقوع زمین‌لغزش در منطقه می‌گردد.

نتایج حاصل از مدل آنتروپی شانون نشان داد که عوامل خاک و کاربری اراضی از مهم‌ترین فاکتورهای مؤثر بر وقوع زمین‌لغزش در استان آذربایجان غربی می‌باشد. همچنین نتایج حاکی از آن است که مدل‌های

اعتبارسنجی نقاط حساس به زمین‌لغزش، گامی اساسی در توسعه و تعیین مناطق حساس و تعیین کیفیت آن‌ها می‌باشد و بدون آن، مدل‌ها و نتایج آن‌ها فاقد اهمیت علمی خواهد بود (Nampak et al. 2014). با استفاده از روش‌های نرخ موفقیت و نرخ پیش‌بینی نتایج نقاط حساس به وقوع لغزش به‌وسیله مقایسه آن‌ها با موقعیت لغزشی موجود ارزیابی گردید که در واقع صحت آنها ارزیابی می‌گردد. نتایج نرخ موفقیت بر پایه داده‌های آموزش تهیه گردید، اما نرخ پیش‌بینی توسط مجموعه داده‌های اعتبارسنجی تهیه گردید. نتایج نرخ موفقیت نشان‌دهنده چگونگی تناسب مدل با داده‌های آموزش می‌باشد و بخاطر اینکه در نرخ موفقیت نقاط آموزش مدل (نقاطی که مدل بر اساس آن ساخته شده) به کار می‌رود، بنابراین روش نرخ موفقیت برای ارزیابی توانایی پیش‌بینی مدل نمی‌تواند مناسب باشد. سطح زیر منحنی نرخ پیش‌بینی برای مدل‌های نسبت فراوانی، آنتروپی شانون، جنگل تصادفی و مدل ترکیبی به ترتیب برابر ۰/۸۹، ۰/۸۸، ۰/۹۱ و ۰/۹۲ می‌باشد



شکل ۵. نتایج منحنی نرخ پیش‌بینی برای مدل‌های اجرا شده (Ba-RF): مدل ترکیبی بگینگ-جنگل تصادفی، RF: جنگل تصادفی، SE: آنتروپی شانون و FR: نسبت فراوانی (محور افقی نشان دهنده نرخ مثبت اشتباه (False positive rate) و محور عمودی نشان دهنده نرخ مثبت صحیح (True positive rate) می‌باشد

هوش مصنوعی دارای کارایی بالاتری از مدل‌های آماری بوده و مدل‌های ترکیبی دارای دقت بالاتری نسبت به مدل‌های منفرد است. بیشتر زمین لغزش‌ها در نزدیکی رودخانه‌ها و گسل‌ها رخ داده است. همچنین در بعضی مناطق، نزدیک‌ترین فواصل به جاده، بیشترین خطر را برای زمین‌لغزش دارد که باید در مدیریت تعریض جاده و ساخت و ساز در این مناطق دقت لازم را اعمال نمود.

بنابراین پیشنهاد می‌شود که در مناطق باغی، کشاورزی و جنگلی به دلیل بالا بودن تعداد زمین‌لغزش‌ها و تخریب خاک و نیز در جاده‌ها به‌منظور جلوگیری از تلفات انسانی، در مدیریت طراحی و تعریض جاده، دقت و نظارت کافی به عمل آید و به شناسایی مناطق مستعد و خطرناک و تثبیت دامنه‌های ناپایدار لغزشی بیشتر پرداخته شود.

#### Reference:

- Al-Hinai, H.Y. & Abdalla, R.I.F.A.A.T. (2020). Spatial Prediction of Coastal Flood Susceptible Areas in Muscat Governorate Using an Entropy Weighted Method. WIT Transactions on Engineering Sciences; WIT Press: Southampton, UK, 129: 121-133.
- Arca, D., Hacısalihoğlu, M. & Kutoğlu, Ş. H. (2020). Producing forest fire susceptibility map via multi-criteria decision analysis and frequency ratio methods. *Natural Hazards*, 104(1): 73-89.
- Aroora, A., Arabameri, A., Pandey, M., Siddiqui, M. A., Shukla, U.K., Bui, D.T., & Bhardwaj, A. (2021). Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. *Science of The Total Environment*, 750: 141565.
- Bui, D. T., Pradhan, B., Revhaug, I., Nguyen, D. B., Pham, H. V., & Bui, Q. N. (2015). A novel hybrid evidential belief function-based fuzzy logic model in spatial prediction of rainfall-induced shallow landslides in the Lang Son city area (Vietnam). *Geomatics, Natural Hazards and Risk*, 6(3), 243-271.
- Chen, W., Shirzadi, A., Shahabi, H., Ahmad, B. B., Zhang, S., Hong, H., & Zhang, N. (2017a). A novel hybrid artificial intelligence approach based on the rotation forest ensemble and naïve Bayes tree classifiers for a landslide susceptibility assessment in Langao County, China. *Geomatics, Natural Hazards and Risk*, 8(2), 1955-1977.
- Chen, W., Chen, X., Peng, J., Panahi, M. & Lee, S. (2021). Landslide susceptibility modeling based on ANFIS with teaching-learning-based optimization and Satin bowerbird optimizer. *Geoscience Frontiers*, 12(1): 93-107.
- Fang, Z., Wang, Y., Peng, L. & Hong, H. (2020). Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Computers & Geosciences*, 139: 104470.
- Gaidzik, K. & Ramírez-Herrera, M. T. (2021). The importance of input data on landslide susceptibility mapping. *Scientific reports*, 11(1): 1-14.
- Khosravi, K., Pham, B.T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., & Bui, D.T. (2018). A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Science of the Total Environment*, 627: 744-755.
- Khosravi, K., Pourghasemi, H. R., Chapi, K. & Bahri, M. (2016). Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models. *Environmental monitoring and assessment*, 188(12): 1-21.
- Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H. B., Grof, G., Ho, H. L., Hong, H., Chapi, K. & Prakash, I., (2019). A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods, *Journal of Hydrology*, 573: 311-323.
- Liang, G., Zhu, X. & Zhang, C. (2011). An empirical study of bagging predictors for different learning algorithms. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*.
- Liu, L., Chin, S.P. & Tran, T.D. (2019). Reducing sampling ratios and increasing number of estimates improve bagging in sparse regression. In *2019 53rd Annual Conference on Information Sciences and Systems (CISS)* (pp. 1-5). IEEE.
- Malekinezhad, H., Sepehri, M., Pham, Q.B., Hosseini, S.Z., Meshram, S.G., Vojtek, M. & Vojteková, J. (2021). Application of entropy weighting method for urban flood hazard mapping. *Acta Geophysica*: 1-14.
- Nampak, H., Pradhan, B., Manap, M.A. (2014). Application of GIS based data driven evidential belief function model to predict groundwater potential zonation, *Journal of Hydrology*, 513: 283-300.
- Ngo, P.T.T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A. & Lee, S. (2021). Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geoscience Frontiers*, 12(2): 505-519.

- Pham, B.T., Bui, D.T., Pourghasemi, H.R., Indra, P. and Dholakia, M.B. (2017b). Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theoretical and Applied Climatology*, 128(1-2), pp.255-273.
- Pham, B.T., Bui, D.T., Prakash, I. & Dholakia, M.B. (2016b). Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. *Natural Hazards*, 83(1), pp.97-127.
- Pham, B.T., Tien Bui, D., Indra, P., & Dholakia, M.B. (2015b). Landslide susceptibility assessment at a part of Uttarakhand Himalaya, India using GIS-based statistical approach of frequency ratio method. *Int J Eng Res Technol*, 4, 338-344.
- Piao, Y., Piao, M., Jin, C.H., Shon, H.S., Chung, J.-M., Hwang, B., Ryu, K.H. (2015). A new ensemble method with feature space partitioning for high-dimensional data classification. *Mathematical Problems in Engineering*
- Piryonesi, S.M. & El-Diraby, T. E. (2020). Data analytics in asset management: Cost-effective prediction of the pavement condition index. *Journal of Infrastructure Systems*, 26(1): 04019036.
- Pourghasemi, H.R., Pradhan, B., Gokceoglu, C., Mohammadi, M. & Moradi, H.R. (2013). Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. *Arabian Journal of Geosciences*, 6(7): 2351-2365.
- Pourghasemi, H.R. & Rahmati, O. (2018). Prediction of the landslide susceptibility: Which algorithm, which precision? *CATENA*, 162: 177-192.
- Raghuvanshi, T.K., Ibrahim, J., Ayalew, D. (2014). Slope stability susceptibility evaluation parameter (SSEP) rating scheme—an approach for landslide hazard zonation. *J. Afr Earth. Sci.*, 99:595-612
- Rodriguez, J.J., Kuncheva, L.I., & Alonso, C.J. (2006). Rotation Forest: A new classifier ensemble method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10), 1619-1630.
- Sarkar, D., Saha, S. and Mondal, P. 2021. GIS-based frequency ratio and Shannon's entropy techniques for flood vulnerability assessment in Patna district, Central Bihar, India. *International Journal of Environmental Science and Technology*: 1-22.
- Sassa, K. & Canuti, P. eds., (2008). *Landslides-disaster risk reduction*. Springer Science & Business Media.
- Shirzadi, A., Bui, D.T., Pham, B.T., Solaimani, K., Chapi, K., Kavian, A., Shahabi, H. and Revhaug, I. (2017a). Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environmental Earth Sciences*, 76(2), 60.
- Sun, D., Shi, S., Wen, H., Xu, J., Zhou, X. & Wu, J. (2021). A hybrid optimization method of factor screening predicated on GeoDetector and Random Forest for Landslide Susceptibility Mapping. *Geomorphology*, 379: 107623.
- Terlien, M.T., van Westen, C.J., & van Asch, T.W. (1995). Deterministic modelling in GIS-based landslide hazard assessment. In A. Carrara & F. Guzzetti (Eds.), *Geographical Information Systems in Assessing Natural Hazards*, PP: 57-77. Netherlands: Springer.
- Youssef, A.M. & Pourghasemi, H.R. (2021). Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geoscience Frontiers*, 12(2): 639-655.



Print ISSN: 2251-7480  
Online ISSN: 2251-7400

Journal of  
**Water and Soil  
Resources Conservation  
(WSRCJ)**

**Web site:**

<https://wsrcj.srbiau.ac.ir>

**Email:**

[iauwsrcj@srbiau.ac.ir](mailto:iauwsrcj@srbiau.ac.ir)  
[iauwsrcj@gmail.com](mailto:iauwsrcj@gmail.com)

**Vol. 11  
No. 4  
Summer 2022**

**Received:**  
2022-02-17

**Accepted:**  
2022-05-20

**Pages: 63-74**



 10.30495/WSRCJ.2022.20004

## Evaluation of the efficiency of artificial intelligence and bivariate statistical models in determining landslide prone areas in West Azerbaijan

Azad Aram<sup>1</sup>, Mohammad Reza Dalalian<sup>2\*</sup>, Siamak Saedi<sup>3</sup>, Omid Raffeian<sup>4</sup> and Samad Darbandi<sup>5</sup>

- 1) PhD Student, Department of Soil Science and Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.
- 2) Assistant Professor, Department of Soil Science and Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.
- 3) Assistant Professor, Department of Soil Science and Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.
- 4) Assistant Professor, Department of Environment, Tabriz Branch, Islamic Azad University, Tabriz, Iran.
- 5) Assistant Professor, Department of Water Science and Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.

\*Corresponding author email: [mdalalian@iaut.ac.ir](mailto:mdalalian@iaut.ac.ir)

**Abstract:**

**Background and Aims:** Landslide is one of the natural hazards that lead to a lot of human and financial losses. Researchers on the subject of landslide susceptibility are investigating the possibility of landslides with respect to topographic and geo-environmental conditions, and the obtained information is critical in landslide risk management. Preparation of landslide sensitive points is an essential tool for assessing landslide risk and is very useful in better planning and management of these areas. In this research, models based on artificial intelligence and two statistical variables in determining landslide sensitive points in West Azerbaijan province have been studied and compared.

**Methods:** Methods based on artificial intelligence and two statistical variables were used to prepare landslide-sensitive points in the province of West Azerbaijan, which is located in northwestern Iran. This study was conducted in four stages. The first stage: the study of landslides in the studied region based on the database of the Forests, Rangelands and Watershed Organization of Iran (FRWO) and the identification of 110 landslides through field surveys, interpretation of aerial photographs and Google Earth satellite images, the second stage: data collecting and creating a spatial databases of effective factors, the third stage: applying the Frequency Ratio (FR), Shannon Entropy (SE), Bagging (BA), Random Forest (RF) and hybrid model (RF-BA) and stage four: methods validating using the system performance curve (ROC). Based on field surveys and similar studies, 12 factors affecting landslide occurrence including altitude, slope angle, slope direction, distance from fault, distance from river, distance from road, drainage density, road density, rainfall, soil, land use and lithology were identified. In the field survey, 110 landslides were identified in West Azerbaijan. 70 percent of the data were randomly selected and used for modeling and 30 percent of the data were used for validation.

**Results:** In terms of geographical directions, the southern direction with a weight of 1.49 had the greatest impact on the occurrence of landslides in the province. The least weight was related to flat areas where no landslide occurred. The results of slope factor showed that the middle slopes had the greatest effect on the occurrence of landslides, so that in low slopes due to low gravity, less landslides occur and too much slopes were related to mountainous areas that were covered with rocks and there was very thin soil that is not suitable for landslide. The study of land use factor showed that 48 percent of landslides occurred in agricultural areas. The results showed that most of the landslides occurred near rivers and faults. Also, in some areas, the closest distances to the road had the greatest risk of landslide.

**Conclusion:** The results of this study showed that the artificial intelligence models (RF and the combined model RF-BA) had the higher efficiency than the statistical models (FR and SE). The accuracy of the combined models was higher than the single models. The ROC curve results showed the accuracy of 0.92, 0.91, 0.89 and 0.88 with RF-BA, RF, FR and SE models, respectively.

**Keywords:** Artificial intelligence, Bagging, Frequency ratio, Random Forest, Shannon entropy