

تعیین پتانسیل آب زیرزمینی با استفاده از مدل‌های شبکه عصبی مصنوعی، جنگل تصادفی، ماشین بردار پشتیبان و رگرسیون خطی (مطالعه موردی: حوضه آبخیز دریاچه ارومیه)

علیرضا رابط^۱، علی دسترنج^{۲*}، سورنا اسدی^۳، امید اسدی نلیوان^۴

۱. بخش تحقیقات اقتصادی-اجتماعی، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی زنجان، سازمان تحقیقات، آموزش

و ترویج کشاورزی، زنجان

۲. بخش تحقیقات حفاظت خاک و آبخیزداری، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی خراسان رضوی، سازمان

تحقیقات، آموزش و ترویج کشاورزی، مشهد

۳. مدیرعامل شرکت سایه گستر دشت البرز، کرج

۴. دکترای آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

(تاریخ دریافت: ۱۳۹۹/۰۵/۱۰، تاریخ تصویب: ۱۳۹۹/۰۸/۱۶)

چکیده

هدف از تحقیق حاضر، تعیین مناطق دارای پتانسیل آب زیرزمینی با استفاده از مدل‌های شبکه عصبی مصنوعی (ANN)، جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM) و رگرسیون خطی (GLM) است. در تحقیق حاضر از ۱۴ پارامتر شامل ارتفاع، شیب، جهت شیب، انحنای فاصله از آبراهه و گسل، تراکم آبراهه و گسل، لیتولوژی، متوسط بارندگی، کاربری اراضی، شاخص موقعیت توپوگرافیک (TPI)، موقعیت شیب نسبی (RSP) و شاخص رطوبت توپوگرافیک (TWI) برای بررسی پتانسیل آب زیرزمینی استفاده شده است. از مجموع ۱۰۶۲۴ چشمه، به صورت تصادفی ۷۰ درصد به عنوان داده‌های آزمون و ۳۰ درصد به عنوان داده‌های اعتبارسنجی طبقه‌بندی شدند. همچنین، برای تعیین مهم‌ترین پارامترها از مدل RF استفاده شد. تست هم‌خطی بین پارامترها با استفاده از نرم‌افزار SPSS انجام شد. از منحنی تشخیص عملکرد نسبی برای قدرت پیش‌بینی مدل‌ها و شاخص سطح سلول هسته (SCAI) به منظور دقت تفکیک بین طبقات استفاده شد. نتایج نشان داد بین پارامترها هم‌خطی وجود ندارد. نتایج مدل RF نشان داد به ترتیب پارامترهای ارتفاع، کاربری اراضی، شیب، فاصله از گسل، TWI و لیتولوژی مهم‌ترین عوامل تأثیرگذار بر پتانسیل آب زیرزمینی هستند. همچنین، بر اساس منحنی ROC در هر دو بخش آموزش (AUC=۰/۹۱۵) و اعتبارسنجی (AUC=۰/۹۰۹)، مدل ANN دارای بیشترین دقت بودند و مدل‌های RF، SVM و GLM در رده‌های بعدی قرار گرفتند. همچنین، نتایج شاخص سطح سلول هسته نشان داد هر چهار مدل با دقت مناسبی به تفکیک طبقات پرداخته‌اند. بر اساس مدل ANN، ۳۱/۴ درصد حوضه پتانسیل آب زیرزمینی زیاد و خیلی زیاد دارد.

کلیدواژگان: پتانسیل آب زیرزمینی، جنگل تصادفی، دریاچه ارومیه، شبکه عصبی مصنوعی، یادگیری ماشینی.

مقدمه

در دهه‌های اخیر در بسیاری از کشورها از جمله ایران به دلیل افزایش جمعیت و توسعه صنعت، آب‌های زیرزمینی به عنوان یکی از منابع طبیعی مهم از جنبه‌های مختلف آب آشامیدنی، فعالیت‌های صنعتی و کشاورزی و همچنین حمایت از توسعه اقتصادی، سلامت انسان و تنوع اکولوژیکی مورد توجه قرار گرفته‌اند [۱ و ۲]. بنابراین، شناسایی مناطق دارای آب زیرزمینی از ضروریات مدیریت منابع آب محسوب می‌شود و می‌تواند به عنوان اقدام اولیه در زمینه مدیریت منابع آب‌های زیرزمینی در نظر گرفته شود [۳]. روش‌های متداول برای تعیین مناطق دارای آب زیرزمینی شامل آزمایش‌های هیدروژئولوژیکی، حفاری، روش‌های زمین‌شناسی، هیدروژئولوژیکی و ژئوفیزیکی می‌شود که بسیار هزینه‌بر و زمان‌بر هستند. سنجش از دور و تکنیک سیستم اطلاعات جغرافیایی در مقایسه با روش‌های سنتی در زمینه ارزیابی پتانسیل آب‌های زیرزمینی که بسیار وقت‌گیر، کم‌دقت و پرهزینه هستند، کارایی زیادی دارند و سبب افزایش دقت و سرعت پژوهش‌ها در زمینه ارزیابی آب‌های زیرزمینی شده‌اند [۴]. هدف از ارزیابی ذخایر آب زیرزمینی، فراهم کردن اطلاعات در رابطه با وضعیت کنونی ذخایر آب و فراهم کردن درکی از وضعیت آب‌های زیرزمینی در آینده است [۳]. در سال‌های اخیر، پژوهشگران مختلفی تلاش کرده‌اند تا پتانسیل منابع آب زیرزمینی را با استفاده از تکنیک‌های سنجش از دور، سیستم اطلاعات جغرافیایی و یادگیری ماشینی ارزیابی کنند [۵-۸]. تکنیک‌های یادگیری ماشینی توجه پژوهشگران را به خود جلب کرده است، زیرا این روش‌ها الگوهای پنهان در آب‌های زیرزمینی را شناسایی می‌کنند و رابطه غیرخطی بین متغیرها را به دست می‌آورند [۹]. روش‌های مختلف یادگیری ماشینی مثل Naive Bayes، AdaBoost، Bagging، GAM [۱۰]، وزن شواهد، رگرسیون لجستیک و درخت توابع [۵]، شبکه عصبی مصنوعی [۱۱]، جنگل تصادفی [۲]، حداکثر آنتروپی [۲]، درخت رگرسیون [۱۲]، درخت رگرسیون تقویت شده [۱۳]، درخت طبقه‌بندی شده تقویت شده [۱۴] و روش‌های آماری نسبت فراوانی [۱۴] و دمپسترشفر [۱۵] در زمینه تعیین مناطق احتمالی وجود آب‌های زیرزمینی مفید بوده‌اند و استفاده شده‌اند. عوامل

مختلفی مانند شیب، جهت شیب، شاخص رطوبت توپوگرافی (TWI)، شاخص موقعیت توپوگرافیک (TPI)، تراکم آبراهه، بارندگی، شاخص توان آبراهه (SPI)، موقعیت شیب نسبی (RSP)، کاربری اراضی، ارتفاع، انحنای سطح، شاخص ارتفاع از سطح نزدیک‌ترین زهکش (HAND)، لیتولوژی و تراکم گسل بر میزان پتانسیل آب زیرزمینی تأثیرگذارند [۱۴-۲۰]. در زمینه تعیین مناطق دارای پتانسیل آب زیرزمینی تحقیقات فراوانی انجام گرفته است. شفیعی و همکاران (۱۳۹۸) در تحقیقی با استفاده از ۹ شاخص مؤثر، پتانسیل آب زیرزمینی در دشت نورآباد ممسنی را با استفاده از شاخص آنتروپی شانون بررسی کردند. آنها بر اساس مقدار ROC مدل یادشده را در تعیین مناطق دارای پتانسیل آب زیرزمینی مناسب ارزیابی کردند [۲۱]. عرب عامری و همکاران (۱۳۹۷) در مطالعه‌ای، پتانسیل آب زیرزمینی در حوضه نجف‌آباد را با استفاده از مدل‌های شواهد قطعی، شاخص آنتروپی و مدل ترکیبی تابع شواهد قطعی-شاخص آنتروپی بررسی کردند. نتایج مطالعه آنها نشان داد مؤثرترین فاکتورها در پتانسیل آب زیرزمینی فاصله از گسل، کاربری اراضی و زمین‌شناسی بودند و مدل ترکیبی بر اساس مقدار AUC دقت بیشتری نسبت به سایر مدل‌ها داشت. همچنین، بر اساس SCAI مقادیر این شاخص برای هر سه مدل مناسب بود [۳]. ذبیحی و همکاران (۱۳۹۴) پتانسیل آب زیرزمینی در دشت بجنورد را با استفاده از مدل‌های آنتروپی شانون و جنگل تصادفی بررسی کردند. بر اساس نتایج و مقدار ROC مدل آنتروپی شانون نسبت به روش جنگل تصادفی بهتر تشخیص داده شد و لایه‌های کاربری اراضی، لیتولوژی، فاصله از آبراهه و طبقات ارتفاعی بیشترین تأثیر را بر پتانسیل آب زیرزمینی داشتند. رحمتی و همکاران (۲۰۱۶) با استفاده از دو مدل جنگل تصادفی و حداکثر آنتروپی، مناطق دارای پتانسیل آب زیرزمینی در دشت مهران را تعیین کردند. نتایج پژوهش آنها نشان داد بر اساس مقدار AUROC هر دو مدل توانایی زیادی دارند و مدل حداکثر آنتروپی توانایی بیشتری دارد و عوامل ارتفاع، تراکم آبراهه، لیتولوژی و کاربری اراضی به ترتیب مهم‌ترین عوامل مؤثر هستند [۲]. میرکی و همکاران (۲۰۱۹) بر اساس روش جدید (RS-RF) و مقایسه آن با روش‌های رگرسیون لجستیک، جنگل تصادفی و Naive Bayes به

آب استانی قرار بگیرد و به این‌وسیله از حفر چاه‌های مجاز و غیرمجاز در مناطق دارای عدم پتانسیل آب زیرزمینی و یا با پتانسیل کم جلوگیری شود.

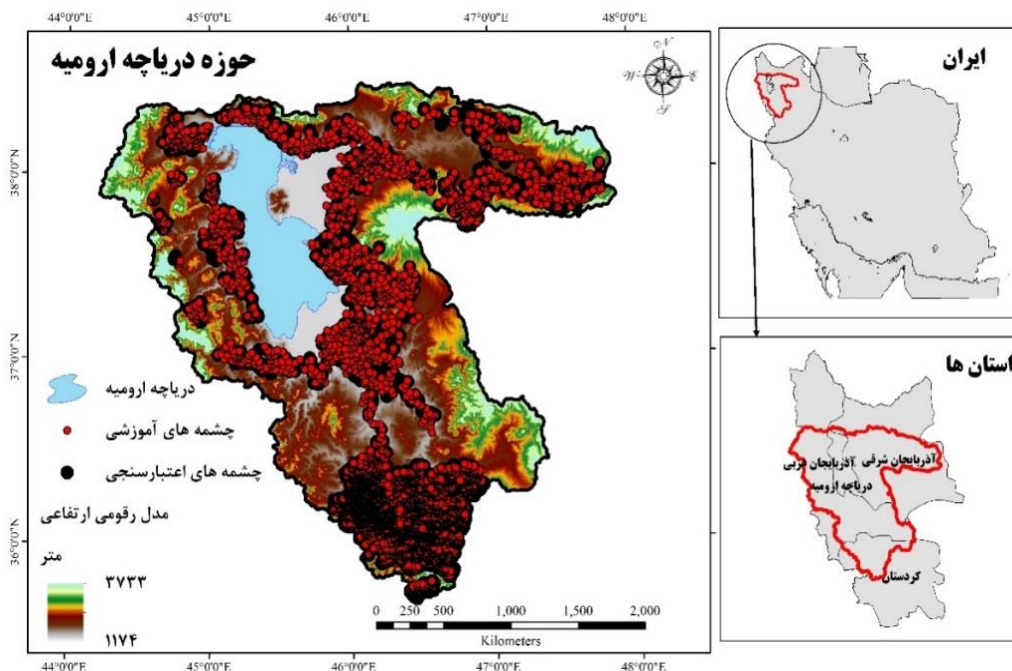
مواد و روش‌ها

منطقه مطالعه شده

حوضه آبخیز دریاچه ارومیه در استان‌های آذربایجان غربی، آذربایجان شرقی و کردستان در مختصات جغرافیایی (بین ۴۳ درجه و ۱۹ دقیقه تا ۴۹ درجه و ۱۷ دقیقه طول شرقی و ۳۵ درجه و ۴۹ دقیقه تا ۳۹ درجه و ۵۴ دقیقه عرض شمالی) و در شمال غربی ایران واقع شده است (شکل ۱). مساحت حوضه مطالعه شده ۵۱۸۲۶ کیلومتر مربع است. ارتفاع محدوده مطالعاتی بین ۱۱۷۴ تا ۳۷۳۳ متر از سطح دریا متغیر است. حداقل و حداکثر دما در این حوضه ۲ تا ۱۴ درجه سانتی‌گراد و همچنین، حداقل و حداکثر شیب حوضه نیز صفر تا ۱۶۸ درصد است. متوسط بارندگی سالانه حوضه آبخیز مورد نظر ۲۴۵ تا ۶۸۷ میلی‌متر در سال است. به لحاظ کاربری، بیشتر سطح حوضه را اراضی کشاورزی (۳۲/۸۴ درصد) و مراتع متوسط (۱۴/۸۸ درصد) تشکیل می‌دهد. شکل ۱ موقعیت حوضه آبخیز و همچنین، موقعیت چشمه‌های آموزش و اعتبارسنجی را نشان می‌دهد (شکل ۱).

این نتیجه رسیدند که روش یادشده دقت خیلی بیشتری نسبت به سایر روش‌های استفاده شده برای تعیین مناطق دارای پتانسیل آب زیرزمینی دارد [۱۷].

حوضه آبخیز دریاچه ارومیه در سال‌های اخیر به واسطه کاهش سطح آب دریاچه ارومیه، افزایش جمعیت و همچنین، تغییرات کاربری اراضی (تبدیل مراتع به اراضی کشاورزی) به یکی از حوضه‌های بحرانی کشور تبدیل شده است. شناسایی و پهنه‌بندی مناطق دارای پتانسیل آب زیرزمینی، گامی مهم و ضروری در مدیریت جامع حوضه آبخیز دریاچه ارومیه است. با توجه به مرور مطالعات انجام گرفته، تا کنون مطالعه جامعی در ارتباط با پهنه‌بندی پتانسیل آب زیرزمینی در کلان حوضه دریاچه ارومیه انجام نشده است. بنابراین، انجام پژوهش حاضر برای ارزیابی و مدیریت این حوضه به منظور پیشگیری و کاهش خسارت‌های ناشی از کاهش حجم آب زیرزمینی ضروری و حیاتی است. هدف از این تحقیق، تعیین مناطق دارای پتانسیل آب زیرزمینی با استفاده از مدل‌های شبکه عصبی مصنوعی، رگرسیون خطی، ماشین بردار پشتیبان و جنگل تصادفی و اولویت‌بندی پارامترهای مؤثر بر پتانسیل آب زیرزمینی با استفاده از مدل جنگل تصادفی است. نقشه نهایی می‌تواند به عنوان نقشه راهنما مورد استفاده مدیران



شکل ۱. موقعیت حوضه آبخیز دریاچه ارومیه در ایران و استان و چشمه‌های آموزش و اعتبارسنجی

داده‌های استفاده‌شده و روش تحقیق

در تحقیق حاضر ابتدا اطلاعات مربوط به موقعیت چشمه‌ها از شرکت آب منطقه‌ای استان‌های آذربایجان غربی و شرقی و کردستان دریافت شد. سپس، بر اساس مرور منابع گسترده در این زمینه از مجموع ۱۰۶۲۴ چشمه، به صورت تصادفی ۳۱۸۷ چشمه (۳۰ درصد) به عنوان داده‌های اعتبارسنجی و ۷۴۳۷ چشمه (۷۰ درصد) به عنوان داده‌های آزمون (آموزش) طبقه‌بندی شدند [۲-۶ و ۱۷-۲۲]. به منظور تعیین پارامترهای مؤثر در پتانسیل آب زیرزمینی و مطالعات مختلف در این زمینه ۱۴ پارامتر شناسایی شد و در مرحله بعد، تست هم‌خطی بین پارامترها با استفاده از نرم‌افزار SPSS انجام شد. با استفاده از شاخص‌های ضریب تحمل (Tolerance) و عامل تورم واریانس (VIF) در صورت وجود هم‌خطی بین پارامترها ($Tolerance < 0.1$ و $VIF > 5$) از روند مدل‌سازی حذف خواهند شد [۳].

ابتدا لایه مدل رقومی ارتفاعی با استفاده از تصاویر ماهواره Landsat-8 (ابعاد 30^*30 متر) تهیه شد. سپس، پارامترهای ارتفاع، شیب، جهت شیب، انحنای توپوگرافی، شاخص موقعیت توپوگرافیک و شاخص موقعیت شیب نسبی بر اساس مدل رقومی ارتفاعی و نرم‌افزار SAGA-GIS تهیه شدند و به عنوان شاخص‌های سطح به منظور ارزیابی پتانسیل آب زیرزمینی مورد استفاده قرار گرفتند [۱، ۳ و ۱۱]. طبقات ارتفاعی سبب ایجاد شرایط اقلیمی متفاوت شده و در نتیجه، سبب ایجاد خاک و تنوع پوشش گیاهی می‌شوند. لایه‌های شیب، جهت شیب و انحنای با استفاده از مدل رقومی ارتفاعی و به ترتیب از توابع Slope، Aspect و Curvature تهیه شده‌اند. شیب فرایند تغذیه آب‌های زیرزمینی، نفوذ و میزان رواناب را کنترل می‌کند و به این صورت بر میزان پتانسیل آب زیرزمینی تأثیرگذار است [۱ و ۱۵]. جهت شیب روی میزان بارندگی، تأثیرات باد و میزان دریافت نور خورشید تأثیرگذار است و به این ترتیب، بر میزان آب دریافتی و نوع پوشش گیاهی تأثیر می‌گذارد [۲۳]. همچنین، انحنای بر روی همگرایی، واگرایی و سرعت جریان تأثیر می‌گذارد [۱، ۷ و ۱۰]. لایه‌های موقعیت شیب نسبی، شاخص موقعیت توپوگرافیک و شاخص رطوبت توپوگرافیک از مدل رقومی ارتفاعی و با استفاده از نرم‌افزار SAGA GIS تهیه شدند. شاخص RSP نشان‌دهنده موقعیت شیب هر سلول و موقعیت نسبی آن بین کف دره (۰) تا

خطالرأس (۱) است (رابطه ۱). شاخص TPI ارتفاع هر یک از سلول‌ها در مدل رقومی ارتفاعی را با ارتفاع متوسط محدوده مشخص شده در اطراف این سلول مقایسه می‌کند [۲۴]. در واقع، این شاخص تفاوت بین پیکسل مرکزی و میانگین سلول‌های اطراف را نشان می‌دهد (رابطه ۲).

$$RSP = \left(\frac{z(s) - z(s)_v}{z(s)_r - z(s)_v} \right) \times 100 + 0.5 \quad (1)$$

که در آن $Z(s)$: ارتفاع؛ $z(s)_v$: ارتفاع دره؛ $z(s)_r$: ارتفاع خطالرأس.

$$TPI = Z_0 - \frac{\sum_{i=1}^n Z_n}{n} \quad (2)$$

که در آن Z_0 : ارتفاع نقطه ارزیابی شده؛ Z_n : ارتفاع شبکه اطراف نقطه ارزیابی شده؛ n : تعداد کل نقاط استفاده‌شده در ارزیابی. مقادیر مثبت مکان‌هایی را نشان می‌دهند که بالاتر از میانگین اطراف نقطه هستند مثل خطالرأس‌ها. مقادیر منفی مکان‌هایی را نشان می‌دهند که پایین‌تر از میانگین اطراف نقطه هستند مثل دره‌ها و چاله‌های محلی. مقادیر صفر مناطق مسطح و جایی که شیب صفر باشد را نشان می‌دهند. میزان بارندگی و نفوذ آن در اعماق زمین یکی از شاخص‌های اصلی وجود پتانسیل آب زیرزمینی است. نقشه بارندگی با استفاده از اطلاعات ۲۰۴ ایستگاه باران‌سنجی تهیه شده است. ابتدا میانگین بارندگی ۲۲ ساله (۱۳۷۶-۱۳۹۸) برای هر ایستگاه محاسبه شد و سپس، با استفاده از روش کریجینگ جهانی (به دلیل کمترین میزان RMSE) در نرم‌افزار ArcGIS پهنه‌بندی بارندگی انجام گرفت. پارامترهای هیدرولوژیک مانند شاخص رطوبت توپوگرافی، تراکم آبراهه و فاصله از آبراهه نقش مهمی در رطوبت خاک، جریان آب زیرزمینی و ثبات شیب دارند [۲۵]. لایه فاصله از آبراهه بر اساس نقشه آبراهه‌ها و با استفاده از تابع Euclidean distance در نرم‌افزار ArcGIS تهیه شد. لایه تراکم آبراهه بر اساس نقشه آبراهه‌ها و با استفاده از تابع Line Density در نرم‌افزار ArcGIS تهیه شد. زهکشی یک منطقه به ساختار سازندهای زمین‌شناسی، میزان نفوذپذیری و شیب بستگی دارد [۲۶]. تراکم زهکشی بالا سبب کاهش نفوذ و افزایش رواناب سطحی می‌شود و بنابراین، مناطق دارای تراکم زهکشی بالاتر پتانسیل کمتری دارند [۱]. شاخص رطوبت توپوگرافی به صورت رابطه ۳ تعریف می‌شود:

می‌توان با کمیت اسکالر S نمایش داد. در این شبکه‌ها وزن‌های داده‌شده به نرون به گونه‌ای تنظیم می‌شود که نزدیک‌ترین مقدار خروجی را نسبت به مقدار مشاهده‌ای تولید کند. این کار آموزش شبکه نام دارد. الگوریتم‌های آموزش، روندهایی هستند که توسط آنها وزن‌های شبکه تنظیم می‌شود. هدف از آموزش شبکه این است که شبکه قانون کار را یاد بگیرد و پس از آموزش به ازای هر ورودی، خروجی مناسب را ارائه دهد. یکی از الگوریتم‌های معروف آموزشی، الگوریتم پس انتشار خطا^۱ است (رابطه ۴).

$$S = \sum_{n=1}^N W_n \cdot x_n = W^T \cdot X \quad (4)$$

که در آن $W=(W_1, W_2, \dots, W_n)$ بردار وزن نرون‌ها است. کمیت s سپس وارد یک تابع غیرخطی f می‌شود تا خروجی را نتیجه دهد:

$$y = f(s) \quad (5)$$

تابع غیرخطی انتقال معمولاً به فرم یک تابع حلقوی به صورت رابطه ۶ تعریف می‌شود:

$$f(s) = (1 + \exp(-S))^{-1} \quad (6)$$

خروجی y می‌تواند نتیجه مدل و یا ورودی لایه بعدی در شبکه‌های چند لایه باشد. الگوریتم‌های مختلفی برای محاسبه وزن‌های بهینه ارائه شده است که در این میان، الگوریتم پس انتشار خطا پرکاربردترین آنهاست. شبکه‌هایی که در این تحقیق استفاده شده‌اند، از نوع شبکه‌های پرسپترون چندلایه هستند. این شبکه‌ها به صورت پیشرو عمل می‌کنند و در ساختار تمامی آنها سعی شده است تا جای ممکن طراحی بهینه با یک لایه میانی صورت پذیرد. آموزش شبکه‌های پرسپترون چندلایه با استفاده از الگوریتم آموزش پس انتشار خطا انجام می‌شود. فرایند عملکرد این شبکه‌ها به این شکل است که لایه ورودی، نسبت به پذیرش داده‌ها اقدام می‌کند و لایه (لایه‌های) میانی عمل فرآوری داده‌ها را انجام می‌دهد. در نهایت، لایه خروجی نیز به نمایش خروجی‌های حاصل از کاربرد مدل اقدام می‌کند. طی مرحله مدل‌سازی، ضرایب مربوط به خطای موجود در گره‌ها به صورت سعی و خطا تصحیح می‌شود که در اغلب موارد از شاخص خطای متوسط داده‌ها بهره برده می‌شود.

$$TWI = IN \left(\frac{A_s}{\tan \beta} \right) \quad (3)$$

که در آن A_s مساحت شیب‌های بالایی (کیلومتر مربع) که به وسیله یک سلول زهکشی می‌شود و β شیب به درجه است.

لایه لیتولوژی از نقشه زمین‌شناسی (مقیاس ۱:۱۰۰۰۰۰) منطقه استخراج شده است. لیتولوژی از طریق تأثیر بر نفوذپذیری و هدایت هیدرولیکی روی پتانسیل آب زیرزمینی تأثیرگذار است [۲۷]. گسل بیشتر پستی و بلندی، توزیع شبکه آبراهه و تجمع آب‌های زیرزمینی تحت تأثیر شیب را کنترل می‌کند [۲۸]. لایه‌های فاصله از گسل و تراکم گسل با استفاده از گسل‌های موجود در منطقه (نقشه زمین‌شناسی) و با استفاده از توابع Euclidean distance و Line Density تهیه شدند. لایه کاربری اراضی از اداره کل منابع طبیعی و آبخیزداری استان آذربایجان غربی دریافت شده است. کاربری اراضی جزء شاخص‌های اکولوژیک دسته‌بندی می‌شود و به صورت مستقیم و غیرمستقیم روی فرایندهای هیدرولوژیک مثل نفوذپذیری، تبخیر و تعرق و تولید رواناب تأثیرگذار است [۲۹].

مدل‌های استفاده شده در تحقیق حاضر

مدل شبکه عصبی مصنوعی

شبکه عصبی مصنوعی نوعی سازوکار محاسباتی است که با گرفتن اطلاعات و محاسبه کردن آنها، یک سری اطلاعات جدید را ارائه می‌دهد [۳۰]. شبکه‌های عصبی مصنوعی از تعدادی نرون تشکیل شده‌اند که به صورت لایه‌ای در کنار یکدیگر قرار می‌گیرند. هر شبکه حداقل از دو لایه تشکیل می‌شود، یک لایه ورودی و یک لایه خروجی. در عین حال می‌تواند چندین لایه مخفی بین لایه‌های ورودی و خروجی قرار گیرد. در این تحقیق از شبکه‌های عصبی پرسپترون چند لایه به دلیل کارایی بهتر نسبت به دیگر انواع شبکه عصبی مصنوعی [۳۱] استفاده شده است. عناصر اصلی یک شبکه عصبی، عصب‌ها یا نرون‌های مصنوعی هستند. الگوی ورودی به یک گره شبیه دندردیت یک سلول بیولوژیک است که می‌توان آن را با برداری با n متغیر یا نرون به صورت $X=(X_1, X_2, \dots, X_n)$ نشان داد. مجموع حاصل ضرب ورودی‌ها در وزن‌های نظیرشان را نیز

الگوریتم‌های یادگیری ماشین محسوب می‌شود. همان‌طور که از نام آن مشهود است، این الگوریتم جنگلی را به طور تصادفی می‌سازد. جنگل ساخته‌شده، در واقع گروهی از درخت‌های تصمیم (Decision Trees) است. کار ساخت جنگل با استفاده از درخت‌ها اغلب اوقات به روش کیسه‌گذاری (Bagging) انجام می‌شود. ایده اصلی روش کیسه‌گذاری آن است که ترکیبی از مدل‌های یادگیری، نتایج کلی مدل را افزایش می‌دهد. به بیان ساده، جنگل تصادفی چندین درخت تصمیم می‌سازد و آنها را با یکدیگر ادغام می‌کند تا پیش‌بینی‌های صحیح‌تر و پایدارتری به دست آیند [۳۴-۳۶]. این الگوریتم را نخستین بار برایمن و کاتلر ایجاد کردند و توسعه دادند. مدل پیش‌بینی‌کننده RF بر اساس میانگین‌گیری از نتایج حاصل از تمامی درخت‌های تصمیم مربوط استوار است و برای بسیاری از مجموعه داده‌ها، طبقه‌بندی با صحت بالایی را انجام می‌دهد [۳۴]. درختان تصادفی با در نظر گرفتن بردار ورودی، آن را با هر درخت در جنگل طبقه‌بندی کرده و خروجی، کلاس‌هایی هستند که از اکثریت آرا دریافت شده است. در این مدل از دو عامل میانگین‌گیری دقت و میانگین‌گیری کاهشی برای تعیین اولویت تأثیر هر یک از عوامل مؤثر استفاده می‌شود [۳۵].

مدل رگرسیون خطی

در آمار، رگرسیون خطی نوعی رویکرد مدل خطی بین متغیر پاسخ با یک یا چند متغیر توصیفی است. اغلب برای کشف مدل رابطه خطی بین متغیرها از رگرسیون استفاده می‌شود. در این حالت فرض بر این است که یک یا چند متغیر توصیفی که مقدار آنها مستقل از بقیه متغیرها است، می‌تواند در پیش‌بینی متغیر پاسخ که مقدارش وابسته به متغیرهای توصیفی نیست، مؤثر باشد. هدف از انجام تحلیل رگرسیون، شناسایی مدل خطی این رابطه است. اگر برای شناسایی و پیش‌بینی متغیر وابسته فقط از یک متغیر مستقل استفاده شود، مدل را رگرسیون خطی ساده می‌گویند. شیب خط در حالت رگرسیون خطی ساده، نشان می‌دهد که میزان حساسیت متغیر وابسته به متغیر مستقل چقدر است. به این معنا که با افزایش یک واحد به مقدار متغیر مستقل، چه میزان متغیر وابسته تغییر خواهد کرد. عرض از مبدأ نیز بیانگر مقداری از متغیر وابسته است که به ازای مقدار متغیر مستقل برابر با صفر محاسبه می‌شود. به شکل دیگر، می‌توان مقدار

این امر از طریق مقایسه خروجی‌های مدل با داده‌های ورودی مشاهداتی انجام می‌پذیرد [۳۱].

مدل ماشین بردار پشتیبان

ماشین بردار پشتیبان یکی از مدل‌های یادگیری ماشین نظارت‌شده است که به منظور طبقه‌بندی و تفکیک داده‌ها به کار می‌رود. به بیان دیگر، پس از مشخص شدن داده‌های ورودی مدل (متغیرهای مستقل) و داده‌های هدف (متغیرهای وابسته)، مدل ماشین بردار پشتیبان پس از تجزیه و تحلیل بین متغیرهای مستقل و وابسته (واسنجی)، داده‌ها را به گروه‌های متمایزی تقسیم می‌کند. در الگوریتم SVM، هر نمونه داده را به عنوان یک نقطه در فضای n بعدی روی نمودار پراکنده داده‌ها ترسیم کرده (n تعداد ویژگی‌هایی است که یک نمونه داده دارد) و مقدار هر ویژگی مربوط به داده‌ها، یکی از مؤلفه‌های مختصات نقطه روی نمودار را مشخص می‌کند. سپس، با ترسیم یک خط راست، داده‌های مختلف و متمایز از یکدیگر دسته‌بندی می‌شوند. ماشین بردار پشتیبان مرزی است که به بهترین شکل دسته‌های داده‌ها را از یکدیگر جدا می‌کند. ایده اصلی این الگوریتم به صورت یک طبقه‌بندی دوتایی با استفاده از نقاط آموزشی است که فضای ورودی اصلی را به یک فضای با ابعاد بالاتر، برای یافتن یک ابرصفحه مطلوب تبدیل می‌کند [۳۲]. نقاط آموزشی که نزدیک به صفحه مطلوب هستند، پشتیبان بردار نامیده می‌شوند. هنگامی که سطح تصمیم به دست آمد، می‌توان از آن برای برآورد داده‌های جدید استفاده کرد [۳۳]. رابطه γ مجموعه‌ای از سلول‌های آموزشی را نشان می‌دهد.

$$X_i = (i=1, 2, \dots, n) \quad (7)$$

سلول‌های آموزشی شامل دو کلاس $X_i = +1$ (بودن یا نبودن چشمه) بوده که به عنوان هدف مدل SVM مشخص می‌شوند.

مدل جنگل تصادفی

جنگل تصادفی، نوعی الگوریتم یادگیری ماشین و نظارت‌شده با قابلیت استفاده آسان است که اغلب اوقات نتایج بسیار خوبی را حتی بدون تنظیم پارامترهای آن، فراهم می‌کند [۳۴]. این الگوریتم به دلیل سادگی و قابلیت استفاده، هم برای دسته‌بندی و هم رگرسیون، از پرکاربردترین

برای هر مقدار احتمالی از برش‌هاست. سطح زیر منحنی، بیانگر مقدار پیش‌بینی سیستم از طریق توصیف توانایی آن در تخمین درست وقایع رخ داده (وجود چشمه) و رخ ندادن آن (نبود چشمه) است. ایده‌آل‌ترین مدل، بیشترین سطح زیرمنحنی را دارد و مقادیر AUC از ۰/۵ تا ۱ متغیر است. هر چه سطح زیرمنحنی به یک نزدیک‌تر باشد، بیانگر بهترین دقت از نقشه‌های پهنه‌بندی تهیه شده است [۲ و ۵]. شکل ۲ فلوجارت مراحل جریان تحقیق را نشان می‌دهد.

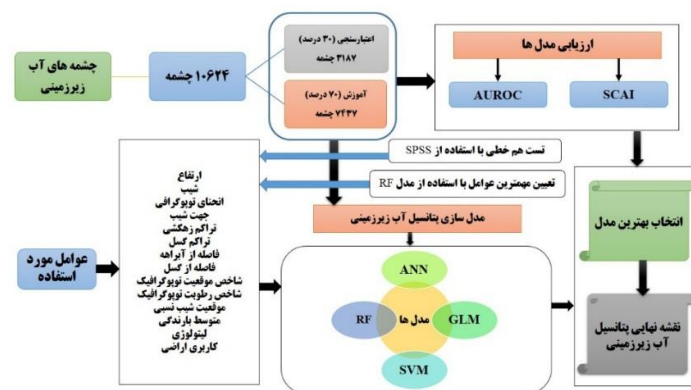
نتایج و بحث

همان‌طور که در بخش مواد و روش‌ها اشاره شد، ابتدا تست هم‌خطی بین پارامترها انجام گرفت (جدول ۱). نتایج نشان داد بین پارامترها هم‌خطی وجود ندارد، بنابراین کلیه پارامترها در روند مدل‌سازی استفاده شدند.

ثابت یا عرض از مبدأ را مقدار متوسط متغیر وابسته به ازای حذف متغیر مستقل در نظر گرفت. در رگرسیون خطی سعی می‌شود، به کمک معادله خطی که توسط روش رگرسیون معرفی می‌شود، برآورد مقدار متغیر وابسته به ازای مقدارهای مختلف متغیر مستقل توسط خط رگرسیون به دست آید. به منظور برآورد پارامترهای مناسب برای مدل، کوشش می‌شود بر اساس داده‌های موجود، مدلی انتخاب شود که کمترین خطا را داشته باشد. معیاری که در مدل رگرسیون خطی ساده به کار می‌رود، کمینه کردن مجموع مربعات خطا است [۳۶].

ارزیابی مدل‌ها

به منظور ارزیابی مدل‌ها، از منحنی تشخیص عملکرد نسبی (ROC) برای قدرت پیش‌بینی مدل‌ها و همچنین، شاخص سطح سلول هسته (SCAI) به منظور دقت تفکیک بین طبقات استفاده شد. منحنی ROC یک نمایش گرافیکی از موازنه بین میزان خطای منفی و مثبت



شکل ۲. فلوجارت تعیین مناطق دارای پتانسیل آب زیرزمینی در حوضه آبخیز دریاچه ارومیه

جدول ۱. نتایج آنالیز تست هم‌خطی بین پارامترهای استفاده‌شده در تحقیق

VIF	Tolerance	پارامترها
۱/۰۷۳	۰/۹۲۱	کاربری اراضی
۱/۰۸۸	۰/۹۱۰	تراکم زهکشی
۱/۱۱۴	۰/۹۰۴	فاصله از آبراهه
۱/۱۸۱	۰/۸۳۴	TPI
۱/۳۱۷	۰/۷۲۵	TWI
۱/۳۵۶	۰/۷۱۳	لیتولوژی
۱/۵۳۹	۰/۶۳۴	بارندگی
۱/۵۵۴	۰/۶۲۵	تراکم گسل
۱/۹۳۵	۰/۵۲۸	RSP
۲/۱۵۳	۰/۴۶۸	جهت شیب
۲/۱۹۶	۰/۴۵۱	فاصله از گسل
۲/۳۶۸	۰/۴۲۱	شیب
۲/۶۲۴	۰/۳۷۴	انحنای
۲/۸۲۷	۰/۳۳۵	ارتفاع

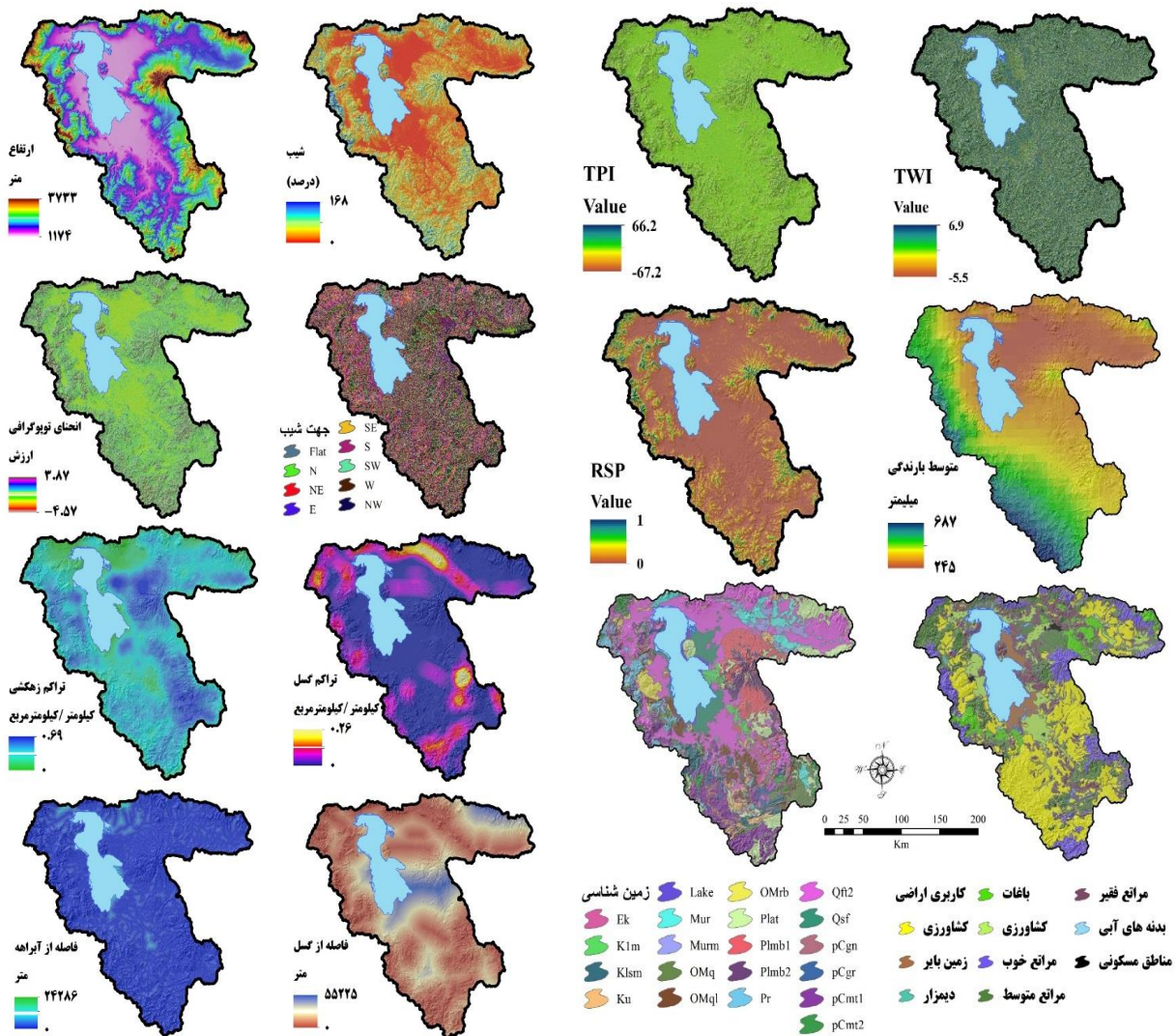
شکل ۳ نقشه پارامترهای استفاده شده در تعیین مناطق دارای پتانسیل آب زیرزمینی را نشان می‌دهد. بر اساس این شکل، حوضه آبخیز دریاچه ارومیه دارای دامنه ارتفاعی ۱۱۷۴ تا ۳۷۳۳ متر و شیب صفر تا ۱۶۸ درصد است که در مناطق جنوبی و غربی شیب بیشترین مقدار و در مرکز حوضه و به سمت دریاچه ارومیه شیب کمتر می‌شود. پارامتر انحنای توپوگرافی دامنه تغییرات بین ۴/۵۷- تا ۳/۸۷ را دارد و همچنین، دامنه تغییرات پارامترهای TPI، TWI و RSP به ترتیب بین ۶۷/۲- تا ۶۶/۲، ۵/۵- تا ۶/۹ و صفر تا یک متغیر است. تراکم زهکشی و تراکم گسل به ترتیب از صفر تا ۰/۶۹ و صفر تا ۰/۲۶ کیلومتر در کیلومترمربع است که تراکم پایینی محسوب می‌شوند. میزان بارندگی حوضه از ۲۴۵ تا ۶۸۷ میلی‌متر متغیر است که بخش جنوب غربی و غرب حوضه بیشترین میزان بارندگی را دارد. حوضه آبخیز مطالعه شده ۱۰ نوع کاربری دارد که میزان اراضی کشاورزی به خصوص در بخش جنوب غربی حوضه بیشترین مساحت را به خود اختصاص داده است. با توجه به وسعت زیاد این حوضه، تنوع سازندهای زمین‌شناسی (دارای ۲۰ نوع واحد لیتولوژی) در آن بسیار زیاد است.

نتایج نشان داد ارتفاعات پایین‌تر نسبت به ارتفاعات بالاتر پتانسیل بیشتری دارند و از دلایل این امر می‌توان به ایجاد رواناب بیشتر در ارتفاعات بالاتر و تغذیه و نفوذ آب زیرزمینی در ارتفاعات پایین‌تر اشاره کرد که با نتایج [۳، ۴ و ۳۷] همخوانی دارد. پارامتر شیب نشان داد در شیب‌های پایین‌تر به دلیل کاهش رواناب میزان پتانسیل آب زیرزمینی بیشتر است که نتیجه نفوذ بیشتر آب در این مناطق است و با نتایج [۱۴ و ۲۰] همخوانی دارد. بر اساس پارامتر جهت شیب، جهات شمال، غرب و جنوب غربی حوضه به دلیل دریافت بیشتر میزان بارش و همچنین، رطوبت بیشتر پتانسیل بیشتری نسبت به سایر جهات دارند. نتایج نشان داد با افزایش شاخص رطوبت توپوگرافی، پتانسیل آب زیرزمینی افزایش پیدا کرده و در بعضی موارد با افزایش TWI کاهش پتانسیل اتفاق افتاده است که افزایش تراکم آبراهه در این مناطق و خروج آب به صورت رواناب را نشان می‌دهد که با تحقیقات [۱۷ و ۳۷] مطابقت دارد. نتایج پارامتر انحنای توپوگرافیک نشان داد سطوح مقعر به دلیل جمع‌آوری آب (حتی به صورت موقت) و شرایط مناسب توپوگرافیک (کاهش شیب) بیشترین میزان تأثیر را بر

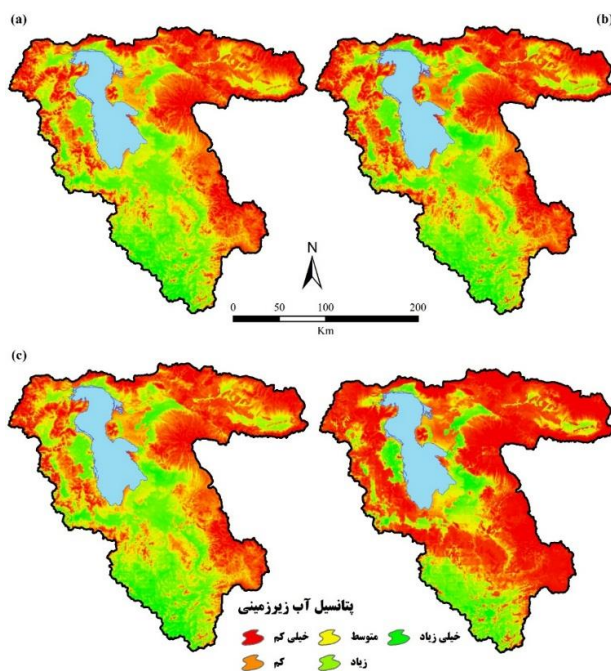
پتانسیل آب زیرزمینی داشته‌اند که با نتایج [۳ و ۳۸] همخوانی دارد. پارامتر بارندگی نشان داد کلاس‌های متوسط و رو به پایین بیشترین تأثیر را در وقوع سفره‌های آب زیرزمینی در حوضه مطالعه شده داشته است. مهم‌ترین دلیل این امر این است که در مناطق مرتفع با شیب بالا میزان بارندگی بیشتر است، ولی به دلیل شیب بالا و شرایط زمین‌شناسی (نبود خاک مناسب برای نفوذ آب و حالت صخره‌ای)، بارندگی تبدیل به رواناب شده و نفوذ بسیار کم است. نتایج نشان داد مناطق دارای تراکم آبراهه بالاتر پتانسیل کمتری نسبت به مناطق دارای تراکم پایین‌تر دارند، زیرا تراکم آبراهه بالاتر رواناب را افزایش می‌دهد و سبب خروج آب بیشتر و سریع‌تر از منطقه می‌شود و از طرف دیگر، نفوذ آب کمتر می‌شود که با نتایج [۶ و ۲۲] در یک راستا است. این در حالی است که در بعضی از تحقیقات بیان داشتند که بین تراکم آبراهه و پتانسیل آب رابطه‌ای وجود ندارد [۳]. همچنین، با فاصله گرفتن از آبراهه‌ها، میزان پتانسیل آب زیرزمینی منطقه مطالعه شده کاهش پیدا کرده است که ارتباط بین آب‌های زیرزمینی با آبراهه‌ها را نشان می‌دهد [۳، ۱۵ و ۲۲]. نتایج نشان داد در منطقه مطالعه شده با افزایش TPI پتانسیل آب زیرزمینی کاهش پیدا کرده است. افزایش مقدار TPI نشان‌دهنده مناطق تپه‌مانند با شیب بالا است که سبب افزایش رواناب می‌شود که در راستای نتایج [۳۷] است. بین تراکم گسل و پتانسیل آب زیرزمینی رابطه مستقیم برقرار است که اهمیت گسل‌ها در تغذیه آب زیرزمینی را نشان می‌دهد [۲، ۱۵، ۲۰ و ۳۷]. این در حالی است که ذبیحی و همکاران [۲۲] رابطه معکوس بین این دو را بیان کردند و دلیل آن را نفوذ آب به اعماق بر اثر تراکم زیاد گسل و عدم ظهور آب در سطح زمین دانستند که با نتایج تحقیق حاضر مطابقت دارد و گسل‌های موجود در منطقه اثر محسوس بر پتانسیل آب زیرزمینی نشان نداده‌اند. نتایج نشان داد با افزایش فاصله از گسل، پتانسیل آب زیرزمینی به صورت نامنظم افزایش و کاهش پیدا کرده است و در حالت کلی روند نزولی دارد. یکی از دلایل این امر، افزایش میزان نفوذ آب زیرزمینی در مناطق نزدیک به گسل است که با نتایج محققان دیگر [۳ و ۲۲] مطابقت دارد. نتایج نشان داد کاربری باغ، مراتع متوسط، مناطق مسکونی و کشاورزی-باغ بیشترین تأثیر را بر پتانسیل آب زیرزمینی داشته‌اند که از دلایل آن می‌توان

پتانسیل آب زیرزمینی است. افزایش مقدار TPI نشان‌دهنده مناطق با شیب بالا است که سبب افزایش رواناب می‌شود و با نتایج [۳۷] مطابقت دارد. شکل ۴ نتایج حاصل از مدل‌سازی پتانسیل آب زیرزمینی در حوضه آبخیز دریاچه ارومیه را با استفاده از مدل‌های یادشده نشان می‌دهد. جدول ۲ مساحت و درصد مساحت هر یک از کلاس‌های پتانسیل آب زیرزمینی را در مدل‌های مختلف نشان می‌دهد. بر اساس جدول یادشده، در همه مدل‌ها طبقه با پتانسیل خیلی کم آب زیرزمینی بیشترین مساحت را به خود اختصاص داده است و طبقه متوسط پتانسیل آب زیرزمینی به غیر از مدل GLM کمترین مساحت را دارد.

به استفاده بیش از حد آب برای مصارف مختلف کشاورزی، باغی و مصارف انسانی و بارندگی مناسب در مراتع متوسط اشاره کرد. در بخش لیتولوژی رسوبات تراس دره‌ای و دشت رسوبی ضخیم، کنگلومرا و سنگ آهک بیشترین میزان مشارکت در پتانسیل آب زیرزمینی را داشته‌اند، چون این سنگ‌ها توانایی زیادی در جذب و ذخیره آب دارند [۶ و ۱۵]. همچنین، نتایج نشان داد با کاهش موقعیت شیب نسبی پتانسیل آب زیرزمینی افزایش پیدا کرده است. به دلیل اینکه با افزایش RSP به مناطق خط‌الرأس نزدیک می‌شویم که در این مناطق بیشتر حالت صخره‌سنگ دارند که در حوضه مطالعه شده دیده می‌شود و کاهش پتانسیل امری بدیهی است. نتایج نشان‌دهنده افزایش پتانسیل آب زیرزمینی با افزایش TPI تا مقدار صفر و سپس، کاهش



شکل ۳. نقشه پارامترهای استفاده‌شده در تعیین پتانسیل آب زیرزمینی در حوضه آبخیز دریاچه ارومیه



شکل ۴. نقشه پتانسیل آب زیرزمینی در حوضه آبخیز ارومیه (a ANN، b RF، c SVM و d GLM)

جدول ۲. مساحت و درصد مساحت کلاس‌های مختلف پتانسیل آب زیرزمینی

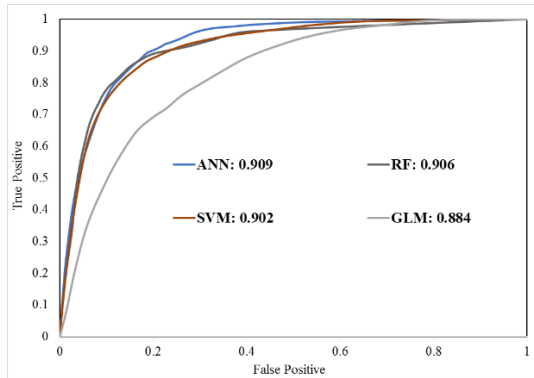
مدل‌ها	کلاس پتانسیل	مقدار کمی	مساحت (کیلومتر مربع)	مساحت (%)
ANN	خیلی کم	۰/۲-۰	۱۸۹۱۳/۳	۳۶/۵
	کم	۰/۴ - ۰/۲	۹۴۷۳/۱	۱۸/۲۸
	متوسط	۰/۶ - ۰/۴	۷۱۲۵/۱	۱۳/۷۵
	زیاد	۰/۸ - ۰/۶	۷۳۰۶/۶	۱۴/۱
	خیلی زیاد	۱ - ۰/۸	۹۰۰۸/۵	۱۷/۳۸
RF	خیلی کم	۰/۲-۰	۱۷۲۰۰/۸	۳۶/۶۹
	کم	۰/۴ - ۰/۲	۱۰۵۴۹/۶	۱۸/۸۲
	متوسط	۰/۶ - ۰/۴	۷۹۶۳/۵	۱۳/۷۲
	زیاد	۰/۸ - ۰/۶	۷۵۸۳/۱	۱۳/۹۵
	خیلی زیاد	۱ - ۰/۸	۸۵۲۹/۶	۱۶/۸۳
SVM	خیلی کم	۰/۲-۰	۱۸۰۷۰	۳۴/۸۷
	کم	۰/۴ - ۰/۲	۹۵۲۳	۱۸/۳۷
	متوسط	۰/۶ - ۰/۴	۷۵۷۶/۴	۱۴/۶۲
	زیاد	۰/۸ - ۰/۶	۷۶۱۰/۸	۱۴/۶۹
	خیلی زیاد	۱ - ۰/۸	۹۰۴۶/۵	۱۷/۴۶
GLM	خیلی کم	۰/۲-۰	۲۸۰۸۳/۴	۵۴/۱۹
	کم	۰/۴ - ۰/۲	۸۶۳۷/۲	۱۶/۶۷
	متوسط	۰/۶ - ۰/۴	۵۲۶۷/۹	۱۰/۱۶
	زیاد	۰/۸ - ۰/۶	۴۵۱۵/۵	۸/۷۱
	خیلی زیاد	۱ - ۰/۸	۵۳۲۲/۶	۱۰/۲۷

شکل‌های ۵ و ۶ نتایج حاصل از قدرت پیش‌بینی مدل‌ها بر اساس مقادیر ROC را در دو بخش آموزش و اعتبارسنجی نشان می‌دهد. بر اساس شکل‌های یادشده در هر دو بخش میزان پیش‌بینی دقت مدل‌های ANN, SVM

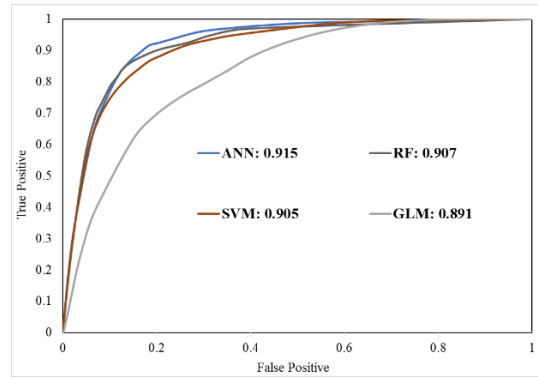
RF و GLM عالی و دقت مدل GLM خیلی خوب بوده است. همچنین، جدول ۳ اولویت‌بندی مدل‌ها را بر اساس میزان AUC نشان می‌دهد. بر این اساس، مدل ANN بیشترین دقت و مدل GLM دقت کمتری (دقت خیلی خوب)

جدول ۵ نتایج وزن پارامترها را حاصل از اجرای مدل RF نشان می‌دهد. بر اساس مدل RF پارامتر ارتفاع مهم‌ترین عامل در تعیین مناطق پتانسیل آب زیرزمینی داشته و پارامتر جهت شیب کمترین وزن را داشته است.

نسبت به سایر مدل‌ها دارد. جدول ۴ نیز نتایج به دست آمده از دقت تفکیک طبقات با استفاده از شاخص سطح سلول هسته را نشان می‌دهد. بر اساس این شاخص هر چهار مدل با دقت مناسبی به تفکیک طبقات پرداخته‌اند.



شکل ۶. نتایج مقادیر ROC در بخش اعتبارسنجی مدل‌ها



شکل ۵. نتایج مقادیر ROC در بخش آموزش مدل‌ها

جدول ۳. اولویت‌بندی مدل‌ها بر اساس نتایج مقدار AUC در دو مرحله آموزش و اعتبارسنجی

مقادیر TSS	اولویت‌بندی		ROC		مدل
	TSS	اولویت بر اساس آموزش و اعتبارسنجی	اعتبارسنجی	آموزش	
۰/۷۸۱	۱	۰/۹۰۹	۰/۹۱۵	ANN	
۰/۷۷۸	۲	۰/۹۰۶	۰/۹۰۷	RF	
۰/۷۷۵	۳	۰/۹۰۲	۰/۹۰۵	SVM	
۰/۷۵۶	۴	۰/۸۸۴	۰/۸۹۱	GLM	

جدول ۴. کلاس‌های پتانسیل آب زیرزمینی و شاخص سطح سلول هسته

SCAI	چشمه‌های اعتبارسنجی			چشمه‌های آموزشی			کلاس‌ها	مدل
	کل چشمه‌ها	درصد	تعداد چشمه	درصد	تعداد چشمه			
۱۴/۷۷	۲/۶۶	۲/۴۵	۷۸	۲/۸۶	۲۱۳	خیلی کم	ANN	
۲/۹۹	۴/۵۹	۴/۵۲	۱۴۴	۴/۶۷	۳۴۷	کم		
۱/۵۶	۴/۸۸	۴/۶۴	۱۴۸	۵/۱۱	۳۸۰	متوسط		
۰/۴۸	۲۸/۵۱	۲۸/۵۵	۹۱۰	۲۸/۴۷	۲۱۱۷	زیاد		
۰/۴۴	۵۹/۳۷	۵۹/۸۴	۱۹۰۷	۵۸/۸۹	۴۳۸۰	خیلی زیاد		
۱۱/۶۲	۲/۱۹	۲/۲	۷۰	۲/۱۸	۱۶۲	خیلی کم	RF	
۲/۳۴	۱۰/۱۵	۹/۵۴	۳۰۴	۱۰/۷۶	۸۰۰	کم		
۱/۲۵	۱۵/۱۴	۱۵/۲۸	۴۸۷	۱۴/۹۹	۱۱۱۵	متوسط		
۰/۷۲	۲۲/۷۸	۲۲/۵۹	۷۲۰	۲۲/۹۷	۱۷۰۸	زیاد		
۰/۳۱	۴۹/۷۵	۵۰/۳۹	۱۶۰۶	۴۹/۱۱	۳۶۵۲	خیلی زیاد		
۱۴/۶۸	۲/۴۹	۲/۳۲	۷۴	۲/۶۵	۱۹۷	خیلی کم	SVM	
۳/۷۴	۴/۸۹	۵/۰۲	۱۶۰	۴/۷۶	۳۵۴	کم		
۱/۴۷	۹/۳۲	۹/۴۸	۳۰۲	۹/۱۷	۶۸۲	متوسط		
۰/۷۴	۱۹/۰۱	۱۹/۲۳	۶۱۳	۱۸/۷۸	۱۳۹۷	زیاد		
۰/۲۷	۶۴/۲۹	۶۳/۹۵	۲۰۳۸	۶۴/۶۴	۴۸۰۷	خیلی زیاد		
۴/۵۱	۱۲/۰۲	۱۲/۰۸	۳۸۵	۱۱/۹۷	۸۹۰	خیلی کم	GLM	
۱/۱۹	۱۴/۰۲	۱۵/۳۴	۴۸۹	۱۲/۶۹	۹۴۴	کم		
۰/۶۸	۱۴/۹۴	۱۶/۷۶	۵۳۴	۱۳/۱۲	۹۷۶	متوسط		
۰/۴	۲۱/۹۹	۲۴/۴۷	۷۸۰	۱۹/۵	۱۴۵۰	زیاد		
۰/۲۸	۳۷/۰۳	۳۱/۳۵	۹۹۹	۴۲/۷۲	۳۱۷۷	خیلی زیاد		

جدول ۵. وزن پارامترهای استفاده شده برای تعیین مهم ترین پارامترها بر اساس مدل RF

وزن	پارامتر
۶۹/۲۹	ارتفاع
۴۶/۰۲	کاربری اراضی
۴۵/۴۹	شیب
۴۴/۳۶	فاصله از گسل
۷۳/۹۳	TWI
۲۴/۸۸	لیتولوژی
۲۳/۴۱	بارندگی
۲۳/۱۴	تراکم آبراهه
۲۰/۴۸	TPI
۱۹/۹۵	انحنای توپوگرافیک
۱۸/۶۶	فاصله از آبراهه
۱۸/۴۲	RSP
۱۳/۷۶	تراکم گسل
۸/۱۱	جهت شیب

نتیجه گیری

با توجه به هزینه بر و زمان بر بودن آزمایش های اکتشافی منابع آب زیرزمینی، تهیه نقشه پتانسیل آب زیرزمینی با استفاده از روش های نوین یادگیری ماشینی با هدف حفاظت، مدیریت و بهره برداری صحیح از منابع آب امری مهم است. با توجه به افزایش جمعیت و نیاز غذایی بیشتر و ضرورت بررسی وجود منابع آب زیرزمینی تحقیق حاضر انجام شد. در این تحقیق با در نظر گرفتن پارامترهای تأثیرگذار و با استفاده از مدل های شبکه عصبی مصنوعی، جنگل تصادفی، ماشین بردار پشتیبان و رگرسیون خطی نقشه پتانسیل آب زیرزمینی حوضه آبخیز ارومیه تهیه شد. نتایج مدل RF نشان داد لایه های ارتفاع، کاربری اراضی، شیب، فاصله از گسل، TWI و لیتولوژی به ترتیب مهم ترین عوامل مؤثر بر پتانسیل آب زیرزمینی هستند که در تحقیقات زیادی [۲، ۳، ۱۵، ۲۰، ۲۲، ۳۷ و ۴۰] مورد تأکید قرار گرفته اند و پارامترهای بارندگی، تراکم آبراهه، TPI، انحنای توپوگرافیک، فاصله از آبراهه، RSP، تراکم گسل و جهت شیب در رده های بعدی هستند [۲۲].

منحنی ROC رسم شده نشان داد دقت مدل ها در برآورد مناطق دارای پتانسیل آب زیرزمینی هم در مرحله آزمون و هم در مرحله اعتبارسنجی عالی بوده است که به معنای عملکرد عالی مدل ها است. با توجه به نتایج

به دست آمده، می توان گفت که مدل های به کاررفته در این تحقیق توانایی زیادی در تعیین مناطق دارای پتانسیل آب زیرزمینی دارند و با توجه به سرعت و دقت زیاد مدل ها، پیشنهاد می شود که در تحقیقات مشابه به خصوص در کشورهای در حال توسعه با هدف بهره برداری مطمئن از این منابع استفاده شود. نتایج به دست آمده از تحقیق حاضر، توجه مدیران، دست اندرکاران و مسئولان امر در اتخاذ تصمیم های آتی مدیریت منابع آب و همچنین، بهره برداری از آنها در منطقه مطالعاتی را می طلبد.

منابع

- [1]. Jothibas A, Anbazhagan S. Modeling groundwater probability index in Ponnaiyar River basin of South India using analytic hierarchy process. *Model. Earth Syst. Environ.* 2016; (2): 109.
- [2]. Rahmati O, Pourghasemi HR, Melesse AM. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region, Iran. *Catena.* 2016; (137): 360–372.
- [3]. Arabameri AR, Sohrabi M, Rezaei KH, Yamani M, Shirani K. Simulation of Najaf-Abad watershed groundwater using data driven ensemble model EBF-Index of entropy. *Journal of Water and Soil Conservation.* 2018; 25(2), 25-48. [Persian]
- [4]. Davoodi Moghaddam D, Rezaei M, Pourghasemi HR, Pourtaghie ZS, and Pradhan B. Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan watershed, Iran. *Arabian Journal Geoscience.* 2015; 8(2), 913-929.
- [5]. Chen W, Li H, Hou E, Wang S, Wang G, Panahi M, et al. GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Science of the Total Environment.* 2018; (634): 853–867.
- [6]. Agarwal R, Garg PK. Remote Sensing and GIS based groundwater potential & recharge zones mapping using multi criteria decision analysis making technique. *Water Resources Management.* 2016 ; (30): 243-260.
- [7]. Rahmati O, Naghibi SA, Shahabi H, Bui DT, Pradhan B, Azareh A, et al. Groundwater spring potential modeling: Comprising the capability and robustness of three different modeling approaches. *Journal of Hydrology.* 2018; (565): 248–261.

- [8]. Chen W, Pradhan B, Li S, Shahabi H, Rizaei H, Hou E, Wang S. Novel Hybrid Integration Approach of Bagging-Based Fishers Linear Discriminate Function for Groundwater Potential Analysis. *Natural Resources Research*. <https://doi.org/10.1007/s11053-019-09465-w>.
- [9]. Mogaji KA, Lim HS. Application of Dempster-Shafer theory of evidence model to geoelectric and hydraulic parameters for groundwater potential zonation. *Journal of Astronomy and Geophysics*. 2018; 7(1): 134-148.
- [10]. Naghibi SA, Moghaddam D, Kalantari B, Pradhan B, Kisi O. A comparative assessment of GIS-based data mining models and a novel ensemble model in groundwater well potential mapping. *Journal of Hydrology*. 2017; (548): 471-483.
- [11]. Lee S, Song KY, Kim Y, Park I. Regional groundwater productivity potential mapping using a geographic information system (GIS) based artificial neural network model. *Hydrogeology Journal*. 2012; (20): 1511-1527.
- [12]. Naghibi SA, Pourghasemi HR. A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods for groundwater potential mapping in Iran, *Water Resources Management*. 2015; 29(14): 5217-5236.
- [13]. Mousavi SM, Golkarian A, Naghibi SA, Kalantari B, and Pradhan B. GIS-based Groundwater Spring Potential Mapping Using Data Mining Boosted Regression Tree and Probabilistic Frequency Ratio Models in Iran. *Geosciences*. 2017; 3(1): 91-115.
- [14]. Lee S, Hyun Y, Lee M. Groundwater Potential Mapping Using Data Mining Models of Big Data Analysis in Goyang-si, South Korea. *Sustainability*. 2019; (11): 1678.
- [15]. Haghizade A, Moghaddam D, Pourghasemi H. GIS-based bivariate statistical techniques for groundwater potential analysis. *Journal of Earth System Science*. 2017; (126): 109.
- [16]. Thapa R, Gupta S, Guin S, Kaur H. Assessment of groundwater potential zones using multi-influencing factor (MIF) and GIS: a case study from Birbhum district, West Bengal. *Appl. Water Sci*. 2017; (7): 4117-4131.
- [17]. Miraki SH, Zanganeh S, Chapi K, Singh VP, Shirzadi A, Shahabi H, Pham BT. Mapping Groundwater Potential Using a Novel Hybrid Intelligence Approach. *Water Resources Management*. 2019; (33): 281-302.
- [18]. Kordestani MD, Naghibi SA, Hashemi H, Ahmadi K, Kalantar B, Pradhan B. Groundwater potential mapping using a novel data-mining ensemble model. *Hydrogeology*. 2018; (27): 211-224.
- [19]. Golkarian A, Naghibi SA, Kalantar B, Pradhan B. Groundwater potential mapping using C5.0, random forest, and multivariate adaptive regression spline models in GIS. *Environmental Monitoring and Assessment*. 2018; 190(3): 149.
- [20]. Golkarian A, Rahmati O. Use of a maximum entropy model to identify the key factors that influence groundwater availability on the Gonabad Plain, Iran. *Environmental Earth Sciences*. 2018; (77): 369.
- [21]. Shafiei N, AmirAhmadi A, Rahmani A. Finding potential sources of underground water catchment model Shannon entropy Nour plain mamasani. *Journal of Geographic space*. 2019; 19(66): 291-304. [Persian]
- [22]. Zabihi M, Pourghasemi HR, Behzadfar M. Groundwater Potential Mapping using Shannon's Entropy and Random Forest Models in the Bojnourd Township. *Journal of EcoHydrology*. 2015; 2(2): 221-232. [Persian]
- [23]. Chowdhury A, Madan C, Bimal C. Integrated remote sensing and GIS-based approach for assessing groundwater potential in West Medinipur district, West Bengal, India. *International Journal of Remote Sensing*. 2009; (30): 231-250.
- [24]. Yeh HF, Cheng YS, Lin H, Lee C. Mapping groundwater recharge potential zone using a GIS approach in Hualian River, Taiwan. *Sustainable Environment Research*. 2016; (26): 33-43.
- [25]. Gaur S, Chahar R, Graillot D. Combined use of groundwater modeling and potential zone analysis for management of groundwater. *International Journal of Applied Earth Observation and Geoinformation*. 2011; (13): 127-139.
- [26]. Manap MA, Nampak H, Pradhan B, Lee S, Soleiman WNA, Ramli MF. Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. *Arab. J. Geosci*. 2012; 7(2): 711-724.
- [27]. Ayazi MH, Pirasteh S, Arvin AKP, Pradhan B, Nikouravan B, Mansor S. Disasters and risk reduction in groundwater: Zagros mountain southwest Iran using geo-informatics techniques. *Dis. Adv*. 2010; (3): 51-57.

- [28]. Samy I, Shattri M, Bujang BK, Ahmad R.M. Structural geologic control with the limestone bedrock associated with piling problems using remote sensing and GIS: a modified geomorphological method. *Environ. Earth Sci.* 2011; 66: 8. 2185-2195.
- [29]. Tehrany MS, Pradhan B, Jebur MN. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol.* 2013; (504): 69-79.
- [30]. Lee S, Ryu JH, Lee MJ, Won JS. The application of artificial neural network to landslide susceptibility mapping at Janghung, Korea. *Mathematical Geology.* 2006; 38(2): 199-220.
- [31]. Caniani D, Pascale S, Sdao F, and Sole A. Neural networks and landslide susceptibility: a case study of the urban area of Potenza. *Natural hazards.* 2008; (45): 55-72.
- [32]. Peng L, Ruiqing NBH, Xueling W, Yannan Z, Runqing Y. Landslide Susceptibility Mapping Based on Rough Set Theory and Support Vector Machines: A Case of the Three Gorges Area, China. *Geomorphology.* 2014; (204): 287-301.
- [33]. Tien Bui D, Pradhan B, Owe L, Inge R. Landslide Susceptibility Assessment in Vietnam Using Support Vector Machines, Decision Tree, ANaive Bayes Models. *Mathematical Problems in Engineering.* 2012; 1-26.
- [34]. Catani F, Lagomarsino D, Segoni S, Tofani V. Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Natural Hazards Earth System Science.* 2013; 13: 2815-2831.
- [35]. Nicodemus KK. Letter to the Editor: On the stability and ranking of predictors from random forest variable importance measures. *Briefings in Bioinformatics.* 2011; (12): 369-373.
- [36]. Youssef AM, Pourghasemi H.R, Pourtaghi Z.S, Al-Katheeri M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region. *Saudi Arabia Landslides.* 2015; DOI 10.1007/s10346-015-0614-1.
- [37]. Arulbalaji P, Padmalal D, Sreelash K. GIS and AHP techniques Based Delineation of Groundwater potential Zones: a case study from southern Western Ghats, India. *Scientific Reports,* 2019; 9: 2082.
- [38]. Al-Abadi A, Al-Temmeme A, Al-Ghanimy A. A GIS-based combining of frequency ratio and index of entropy approaches for mapping groundwater availability zones at Badra-Al-Gharbi-Teeb areas, Iraq. *Sustain. Water Resources Management.* 2016; 2(3): 265-283.
- [39]. Deng F, Deng Z, Lv D, Wang D, Duan H, Xing Z. Application of remote sensing and GIS analysis in groundwater potential estimation in west Liaoning Province, China. *Journal of Engineering Research.* 2016; 4(3): 1-17.
- [40]. Muniraj K, Jesudhas CJ, Chinnasamy A. Delineating the Groundwater Potential Zone in Tirunelveli Taluk, South Tamil Nadu, India, Using Remote Sensing, Geographical Information System (GIS) and Analytic Hierarchy Process (AHP) Techniques. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences.* 2019; Doi.org/10.1007/s40010-019-00608-5.