

## ارزیابی کارایی عوامل اقلیمی و ژئومورفومتری در پیش‌بینی درصد پوشش گیاهی بر مبنای فرایندهای یادگیری ماشین

زینب میرشکاری<sup>۱</sup>، مجید صادقی‌نیا<sup>۲</sup>، مصطفی شیرمردی<sup>۳\*</sup>، مریم اسدی<sup>۴</sup>

تاریخ دریافت: ۱۳۹۷/۱۰/۳

تاریخ پذیرش: ۱۳۹۸/۴/۲۲

### چکیده

سطح قابل توجهی از مساحت مراتع کشور ایران، اختصاص به گیاهان بوته‌ای دارد. یکی از گیاهان مهم بوته‌ای، درمنه دشتی است. درمنه‌زارها از نظر حفاظت خاک و تأمین غذای دام‌های اهلی و وحشی نقش بسزایی ایفا می‌کنند. از این رو در این پژوهش، به بررسی وضعیت پراکنش گونه درمنه، بر مبنای عامل‌های ژئومورفومتری و اقلیمی و متغیر درصد پوشش گیاهی با استفاده از فرایند یادگیری ماشین پرداخته شده است. هدف از این مطالعه، ارزیابی کارایی مدل‌های نزدیک‌ترین همسایه، شبکه عصبی مصنوعی، فرایند گوسی، درخت تصمیم M5 و ماشین بردار پشتیبان به کمک عامل‌های ژئومورفومتری مستخرج از مدل رقومی ارتفاعی و همچنین عامل‌های اقلیمی برای پیش‌بینی درصد پوشش گیاهی است. پس از اجرای الگوریتم‌ها، وزن‌دهی عامل‌ها و تعیین میزان تأثیرشان در پیش‌بینی درصد پوشش انجام گرفت. ارزیابی نتایج مدل‌ها روی عامل‌های ژئومورفومتری نشان داد که در مجموع، برای داده‌های آموزش مدل فرایند گوسی با ریشه میانگین مربعات خطا ۲/۷۳ و ضریب تبیین ۰/۹۶ دارای بیشترین دقت است. در ارزیابی مدل نیز داده‌های آزمون فرایند گوسی با ریشه میانگین مربعات خطا ۱/۱۷ و ضریب تبیین ۰/۹۹ بهترین مدل است. همچنین ارزیابی نتایج مدل‌ها روی عامل‌های اقلیمی نشان داد که برای داده‌های آموزش مدل درخت تصمیم‌گیری با ریشه میانگین مربعات خطا ۹/۶۶ و ضریب تبیین ۰/۵۸ دارای بیشترین دقت است. در ارزیابی مدل نیز در مجموعه داده‌های آزمون، مدل درخت تصمیم‌گیری با ریشه میانگین مربعات خطا ۸/۶۰ و ضریب تبیین ۰/۵۷ بهترین مدل برآورد شد. نتایج حاصل از وزن‌دهی نیز نشان داد که از میان عوامل ژئومورفومتری، فاصله از آبراهه، سطح پایه آبراهه و ارتفاع دارای بیشترین تأثیر و از میان عامل‌های اقلیمی رطوبت دارای بیشترین تأثیر در پیش‌بینی درصد پوشش گیاهی است.

**واژه‌های کلیدی:** درصد پوشش، درمنه دشتی، عامل‌های اقلیمی، عامل‌های ژئومورفومتری، یادگیری ماشین.

۱. دانش‌آموخته کارشناسی ارشد مرتعداری، گروه مهندسی طبیعت، دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان

۲. استادیار گروه مهندسی طبیعت، دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان

۳. استادیار گروه علوم باغبانی، دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان؛ shirmardi@ardakan.ac.ir

۴. دانشجوی دکتری آبخیزداری، گروه مهندسی احیاء مناطق خشک و کوهستانی، دانشکده منابع طبیعی، دانشگاه تهران

## مقدمه

مراتع اکوسیستم‌های طبیعی هستند که دارای منابع ژنتیکی بزرگ و تنوع گونه‌ای گیاهی هستند که اثرات عمیقی بر ثبات آن دارد (یزدان‌شناس و همکاران، ۲۰۱۳). پوشش گیاهی از مهم‌ترین اجزای هر اکوسیستم به شمار می‌رود. از آنجا که پوشش گیاهی بستر حیات در کره زمین است و با عوامل زنده و غیرزنده اطراف خود ارتباط تنگاتنگی دارد، همچنان به‌طور ذاتی دارای پویایی است و تحت تأثیر عوامل محیطی اطراف خود تغییر می‌کند، تعیین میزان درصد پوشش گیاهی به‌منظور فهم تعاملات بین زمین و اتمسفر، تأثیر آن بر اقلیم، میزان فرسایش خاک، بررسی خشکسالی و مدیریت منابع طبیعی بسیار ضروری است (درویش‌زاده و همکاران، ۲۰۱۲). همچنین با بررسی عوامل مؤثر بر گیاهان با دقت بیشتری می‌توان در خصوص جنبه‌های مختلف بهره‌برداری از آن‌ها اظهار نظر کرد. از میان عوامل اکولوژیکی، عامل‌های اقلیمی مهم‌ترین عواملی هستند که در پراکنش پوشش گیاهی نقش عمده‌ای دارند (ساجیر، ۱۹۹۶). در واقع خصوصیات اقلیمی روی سایر عوامل محیطی تأثیر شگرفی دارند. همبستگی شدید و ارتباط تنگاتنگ بین پوشش گیاهی و عامل‌های اقلیمی به‌گونه‌ای است که تغییر در هر یک از آن‌ها، تأثیر شدیدی بر دیگر کارکردهای اکوسیستم می‌گذارد. میزان بارندگی، فصل بارش، تعداد روزهای بارانی و رژیم دمایی با توپوگرافی برهم‌کنش متقابل به وجود می‌آورند که حاصل آن ایجاد پوشش گیاهی ویژه در هر منطقه است. یکی دیگر از عوامل تأثیرگذار در پراکنش پوشش گیاهی مناطق، توپوگرافی است. شناخت روابط مقطعی بین توپوگرافی و پوشش گیاهی می‌تواند به درک بهتر و مدیریت منابع بیولوژیکی در هر منطقه منجر شود (جیانگ بو و تانگ، ۲۰۱۰). پستی و بلندی به‌طور مستقیم از طریق تأثیر بر عوامل محیطی و به‌طور غیرمستقیم از طریق اثرش در تشکیل خاک، بر رشد پوشش گیاهی مؤثر است. اختلاف فاحش در توپوگرافی موجب به وجود آمدن اقلیم و پوشش گیاهی

متفاوت می‌شود. دو عامل جهت جغرافیایی و زمین‌شناسی بر تغییرات درصد پوشش تاجی گیاهان تأثیرگذار است. تغییرات ارتفاع به‌دلیل تأثیر در اقلیم منطقه بر پراکنش گونه‌های گیاهی نقش مؤثری داشته و می‌تواند بسیاری از عوامل محیطی را تغییر دهد. با افزایش ارتفاع از سطح دریا، متوسط دمای هوا کاهش یافته و با توجه به سایر عوامل محیطی منجر به تشکیل نواحی اقلیمی شده، در نتیجه پوشش گیاهی با تنوع گونه‌ای خاص ایجاد می‌شود (خان‌به‌بینی و محمداسمعیلی، ۲۰۱۵). بنابراین عوامل اقلیمی و توپوگرافی بر هم اثر متقابل گذاشته و به‌کارگیری شاخص‌های مؤثر در ویژگی‌های پوشش گیاهی می‌تواند نقش مؤثری در برآورد درصد پوشش گیاهی داشته باشد. بررسی اثر هر یک از شاخص‌ها می‌تواند کمک شایانی در زمینه برآورد درصد پوشش در مناطق دشوار نمونه‌برداری داشته باشد. در مطالعات مختلف اثر عامل‌های مختلف همچون ارتفاع، جهت، درجه حرارت، بارندگی و مواردی دیگر مورد بررسی قرار گرفته است. نتایج مطالعه محتشم‌نیا (۲۰۱۱) نشان داد که عامل ارتفاع از سطح دریا یکی از عوامل اولیه اثرگذار در پراکنش گونه *Artemisia sieberi* در مراتع استپی استان فارس است. عبداللهی و نادری (۲۰۱۲) به بررسی اثر متغیرهای توپوگرافی و خصوصیات فیزیکی‌شیمیایی خاک بر نحوه عملکرد عامل‌های مؤثر بر رشد *Artemisia sieberi* در مراتع استپی ندوشن یزد پرداختند و به این نتیجه رسیدند که ارتفاع بیشترین تأثیرات مثبت را بر ویژگی‌های رویشی درمنه داشته است و با افزایش ارتفاع بر رشد و گسترش این گونه افزوده می‌شود. همچنین موعلی و همکاران (۲۰۱۳) در پژوهشی تأثیر عوامل اقلیمی بر پراکنش گونه مرتعی رمس را که گسترش وسیعی در مناطق گرم و خشک لارستان و طبس دارد، مورد بررسی قرار دادند. نتایج نشان داد که عامل درجه حرارت مهم‌ترین نقش را در پراکنش گونه مرتعی رمس در شهرستان‌های لارستان و طبس دارد. بهرامی و قربانی (۲۰۱۶) به بررسی و تعیین عوامل محیطی مؤثر بر پراکنش رویشگاه‌های مرتعی جنوب شرقی سبلان با استفاده از تجزیه و تحلیل آماری چندمتغیره پرداختند. نتایج آن‌ها نشان داد از

1. Sugier
2. Jiangbo & Tong

عامل مؤثر بر کاهش پوشش گیاهی، کاهش بارش سالیانه است.

فرایند یادگیری ماشین، از ابزارهای جدیدی است که می‌تواند با برقراری رابطه بین متغیرهای مختلف شرایط محیطی، درصد پوشش گیاهی را با دقت بالایی برآورد کند. در فرایند یادگیری ماشین، داده‌ها با استفاده از تکنیک‌ها و عملیات‌های داده‌کاوی مورد کاوش قرار می‌گیرند تا الگوهای مورد نظر کشف شود. با تعیین الگوی مناسب و مدل‌سازی روابط می‌توان با کمترین هزینه و نیروی کار، میزان درصد پوشش را برآورد کرد. این مدل‌سازی‌ها امروزه بر مبنای روابط بین متغیرها می‌تواند کمک شایانی در زمینه پیش‌بینی پوشش گیاهی داشته باشد؛ برای مثال، متکان<sup>۵</sup> و همکاران (۲۰۱۱) به بررسی کارایی استفاده از تصاویر ماهواره‌ای و شبکه‌های عصبی مصنوعی در تخمین درصد پوشش گیاهی در مناطق خشک بافق یزد پرداخته و به این نتیجه رسیدند که شبکه عصبی توانایی تخمین درصد پوشش گیاهی را با دقت مناسبی دارد ( $R^2 > 0/74$  و  $RMSE < 0/02$ ). کارپنتر<sup>۶</sup> و همکاران (۱۹۹۹) به منظور تهیه نقشه پوشش گیاهی از شبکه عصبی مصنوعی استفاده کرده و نشان داده‌اند که شبکه عصبی مصنوعی می‌تواند پوشش گیاهی را با دقت بالایی برآورد کند.

از آنجایی که یکی از مناطق مورد مطالعه (ندوشن) از قطب‌های دامداری استان یزد به شمار می‌آید و با توجه به اینکه گونه درمنه دشتی یکی از گونه‌های مهم مرتعی در استان می‌باشد که نقش مؤثری در تغذیه دام‌های استان داشته و از طرفی دارای پراکنش زیادی در سطح استان است و همچنین درمنه‌زارها تحت تأثیر فعالیت‌های تخریبی انسان و چرای دام‌ها قرار دارند، در زمره اکوسیستم‌هایی هستند که حفاظت از آنها ضروری است. از این رو، اولی‌ین گام برای جلوگیری از تخریب آنها، شناخت پوشش گیاهی، بررسی جوامع گیاهی و عوامل تأثیرگذار بر آنهاست تا با استفاده از این اطلاعات، راه‌های مدیریت صحیح برای حفظ و ارتقای

میان عوامل فیزیوگرافی، ارتفاع و جهت جغرافیایی بیشترین تأثیر را در پراکنش رویشگاه‌های مرتعی دارند. سلامی و همکاران (۲۰۱۷) تأثیر عوامل محیطی بر پوشش گیاهی مراتع اینچه برون در استان گلستان را با استفاده از رگرسیون بررسی کردند. نتایج نشان داد که بین ارتفاع و پوشش گیاهی ارتباط معنی‌داری وجود دارد و افزایش ارتفاع تغییر ترکیب گیاهی را سبب می‌شود. گورجل و فریرا<sup>۱</sup> (۲۰۰۳) به مطالعه تغییرات پوشش گیاهی در ارتباط با تغییرات اقلیمی در برزیل پرداختند. نتایج نشان داد که تغییرات پوشش گیاهی در واکنش به تغییرات اقلیمی بسته به نوع پوشش گیاهی منطقه متفاوت است و زمان بارندگی تأثیرات متفاوتی را بر پوشش گیاهی منطقه داشته است. بارندگی در فصل تابستان همبستگی ۷۰ درصد و در فصل بهار همبستگی ۹۱ درصد با تغییرات پوشش گیاهی نشان داد. بیتس<sup>۲</sup> و همکاران (۲۰۰۶) تأثیر زمان بارندگی را روی پوشش درمنه‌زارهای استپی آمریکا مورد تحقیق قرار دادند. نتایج تحقیق آنها نشان داد که بارندگی بهاره بیشترین تأثیر را روی پوشش گیاهی منطقه مورد مطالعه دارد. اشمیت<sup>۳</sup> و همکاران (۲۰۱۳) به بررسی رابطه بین پوشش گیاهی و اقلیم در مناطق گرمسیری شمال آفریقا با استفاده از تجزیه و تحلیل گام‌به‌گام پرداختند. نتایج نشان داد که علاوه بر بارش و درجه حرارت، تابش جهانی و رطوبت نسبی نیز نقش مهمی در پراکنش پوشش گیاهی منطقه دارند. نودهی و همکاران (۲۰۱۴) تأثیر عوامل توپوگرافی بر توزیع جوامع گیاهی در مراتع نیمه‌استپی منطقه قرخود خراسان شمالی را با استفاده از آنالیز تطبیقی متعارفی بررسی کرده و به این نتیجه رسیدند که ارتفاع به‌طور مستقیم و غیرمستقیم بر بارندگی و دما تأثیر گذاشته و مهم‌ترین عامل در توزیع جوامع گیاهی در منطقه است. یوسف النور<sup>۴</sup> و همکاران (۲۰۱۷) رابطه بین عوامل اقلیمی با پوشش گیاهی را در چین‌گهای چین مورد بررسی قرار دادند. نتایج نشان داد که پوشش گیاهی منطقه به‌صورت پراکنده بوده و مهم‌ترین

1. Gurgel & Ferreira
2. Bates
3. Schmidet
4. Yousif Elnour

5. Matkan  
6. Carpenter

مناطق و چگونگی نمونه برداری را نشان می دهد.

### آماده سازی داده ها

#### نمونه برداری و اندازه گیری درصد پوشش گیاهی

برای شناخت کلی مناطق نمونه برداری، در اواخر اسفند ۱۳۹۶ بازدیدی مقدماتی از محدوده های مورد نظر انجام شد. در فروردین ماه، بازدید دوم برای تعیین مناطق نمونه برداری صورت گرفته و در نهایت ده منطقه نمونه برداری مشخص گردید. سپس در اردیبهشت ۱۳۹۷ نمونه برداری از مناطق که شامل مناطق با درصد پوشش صفر تا درصد پوشش های متنوع بود انجام گرفت. در این مطالعه، از روش نمونه گیری تصادفی - سیستماتیک استفاده شد (ارزانی و عابدی، ۲۰۱۵). نقاط قرارگیری ترانسکت ها به صورت تصادفی انتخاب شد و پلات ها با فواصل مشخص در امتداد هر ترانسکت استقرار یافتند (مصادقی، ۲۰۱۵). در هر منطقه، ۴ ترانسکت ۱۵۰ متری که ۲ ترانسکت آن در جهت شیب و ۲ ترانسکت آن عمود بر جهت شیب بود مستقر شد. اندازه پلات های نمونه برداری ۱×۲ مترمربعی مشخص شد. برای تعیین تعداد پلات از روش های ترسیمی و آماری استفاده می شود (همان)؛ که در این پژوهش با استفاده از روش آماری، برای هر ترانسکت ۸ پلات در نظر گرفته شده و نام گونه های موجود و درصد پوشش هر گونه یادداشت گردیده است. همچنین با استفاده از GPS طول و عرض جغرافیایی محل پلات ها مشخص شد.

#### آماده سازی داده های ژئومورفومتری

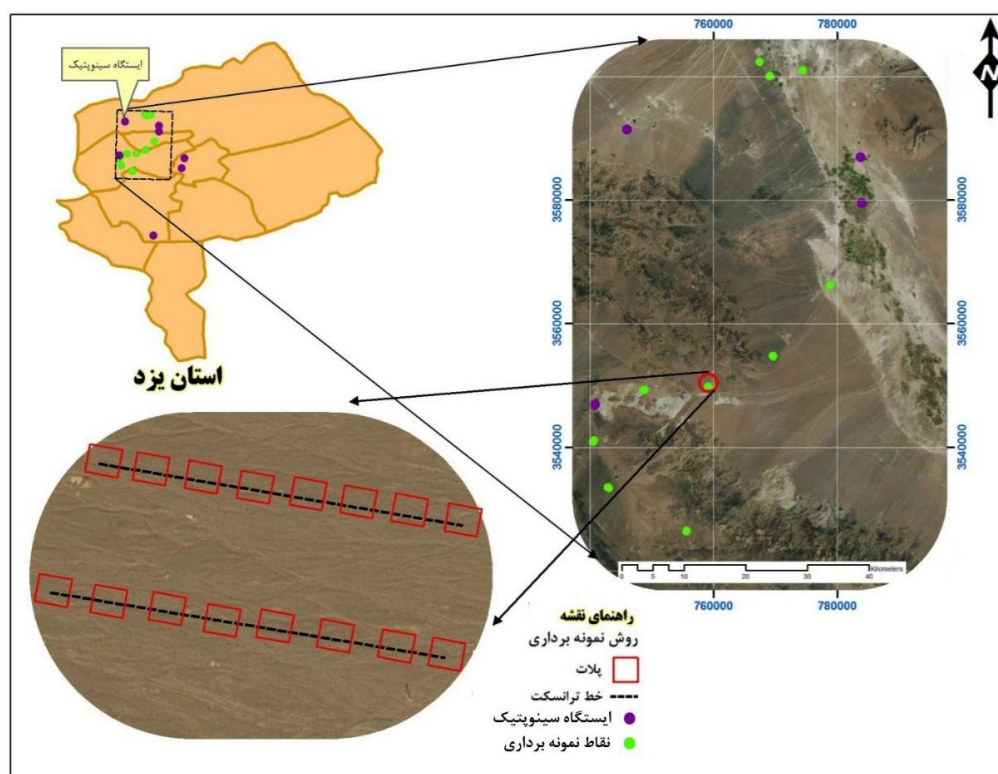
در ابتدا به منظور استخراج عامل های ژئومورفومتری از مدل رقومی ارتفاعی و نرم افزار SAGA با ورژن ۲.۱.۳ استفاده شد. مدل رقومی ارتفاعی به عنوان نقشه مبنا برای استخراج سایر عامل های ژئومورفومتری مورد استفاده قرار گرفت. پیکسل سائز مدل رقومی ارتفاعی بر مبنای قدرت تفکیک GPS تعیین شد (هنگل، ۲۰۰۶). رابطه (۱) محاسبات پیکسل را نشان می دهد.

این مراتع مهم کشور بررسی شود و تصمیم های صحیح حفاظتی اتخاذ گردد. همچنین با توجه به اهمیت برآورد درصد پوشش گیاهی و بررسی روند تغییرات آن در سال های مختلف لازم است تا محققان به دنبال روش های جدید و کارا برای برآورد درصد پوشش بدون استفاده از عملیات های میدانی باشند. لذا در این مطالعه، به پیش بینی درصد پوشش گیاهی بر مبنای عامل های ژئومورفومتری و اقلیمی با استفاده از تکنیک های داده کاوی پرداخته شد.

### مواد و روش ها

#### منطقه مورد مطالعه

منطقه مورد مطالعه از دشت یزد- اردکان شروع و در منطقه ندوشن خاتمه می یابد. این دشت یکی از وسیع ترین دشت های استان یزد است که در طول جغرافیایی ۵۳ درجه و ۱۵ دقیقه تا ۵۴ درجه و ۵۰ دقیقه شرقی و در عرض جغرافیایی ۳۱ درجه و ۱۵ دقیقه تا ۳۲ درجه و ۱۵ دقیقه شمالی قرار دارد. این منطقه به علت واقع بودن در زیر مرکز پرفشار جنب استوایی دارای بارندگی کم و نوسانات درجه حرارت زیاد است. جزء کمربند خشک فلات مرکزی ایران با متوسط دمای سالانه ۱۲ تا ۱۹ درجه، نوسان رطوبتی پایین بین ۳۰ تا ۵۰ درصد و تبخیر شدید بین ۲۲۰۰ تا ۳۲۰۰ میلی متر همراه با بارش های اندک و نامنظم است. همچنین منطقه ندوشن در محدوده جغرافیایی ۳۱ درجه و ۴۵ دقیقه تا ۳۲ درجه و ۳ دقیقه عرض شمالی و ۵۳ درجه و ۲۸ دقیقه تا ۵۳ درجه و ۴۷ دقیقه طول شرقی قرار دارد. مراتع این حوزه در ردیف مراتع استپی با پوشش غالب گیاهان بوته ای است و این منطقه یکی از قطب های مهم دامداری در استان است. بیشتر مساحت منطقه را مناطق تقریباً مسطح تشکیل می دهد که با حرکت به سمت کوهستان بر میزان شیب منطقه افزوده می شود، به گونه ای که شیب های بالای ۶۰ درجه در ستیغ های کوهستانی مشاهده می شود. متوسط دمای سالانه آن نیز با در نظر گرفتن کاهش دما در طول گرادیان ارتفاع در دامنه ۸/۷۵ تا ۱۴/۶ درجه سانتی گراد قرار دارد. سیمای پوشش گیاهی در محدوده مطالعاتی گیاهان درختچه ای و بوته ای و گونه گیاهی غالب این مناطق درمنه دشتی است. شکل (۱) محدوده



شکل (۱): موقعیت منطقه و ایستگاه‌های مورد مطالعه

Figure (1): Case study and studied stations

سایت <https://vertex.daac.asf.alaska.edu> با رزولوشن ۱۲ متری استفاده شد. این عامل‌ها با توجه به اینکه بیانگر ویژگی‌های سطح زمین هستند، به‌منظور افزایش دقت مدل‌های داده‌کاوی در پیش‌بینی درصد پوشش گیاهی به کار گرفته شدند. جدول (۱) عامل‌های استفاده‌شده را نشان می‌دهد.

### داده‌های اقلیمی

به‌منظور تهیه عامل‌های اقلیمی مورد نیاز این تحقیق، از داده‌های هواشناسی ۷ ایستگاه عقدا، میبد، یزد، اردکان، اشکذر، ندوشن و خضرآباد که نزدیک‌ترین ایستگاه‌ها به نقاط نمونه‌برداری بود و از داده‌های مطمئن‌تر و کامل‌تری نسبت به ایستگاه‌های دورتر برخوردار بود، در دوره آماری ۵ ساله از سال ۱۳۹۱ تا ۱۳۹۶ استفاده شد. عامل‌های اقلیمی استفاده شامل رطوبت<sup>۱</sup>، بارش<sup>۲</sup>، تبخیر<sup>۳</sup> و درجه‌حرارت<sup>۴</sup> است.

(۱)

$$P \leq \frac{MLD}{4} = \frac{\sqrt{SN^2 * 0/000025}}{2} = SN * 0.0025$$

که P اندازه پیکسل، SN مقیاس و MLD<sup>۱</sup> حداقل ناحیه تعریف قابل خواندن بر روی زمین است.

با تعیین رزولوشن مدل رقومی ارتفاعی، عامل‌های ژئومورفومتری با استفاده از نرم‌افزار ساگا استخراج شد. عامل‌های ژئومورفومتری اولین بار توسط ایوانس<sup>۲</sup> (۱۹۷۲) به‌صورت خلاصه در زمینه تحلیل ریاضی عوارض سطح زمین و در تفکیک عوارض سطح زمین به کار گرفته شد. الگوریتم‌های مورد استفاده در ژئومورفومتری شامل محاسباتی است که روی مشتق اول و دوم داده‌های ارتفاعی در پنجره‌ها یا فیلترهایی با ابعاد مختلف اعمال می‌شود (ایوانس، ۱۹۷۲؛ شری<sup>۳</sup>، ۱۹۹۵). حداکثر پیکسل سائز مدل رقومی ارتفاع با توجه به رابطه (۱) ۲۰ متر محاسبه شد. در این تحقیق از تصاویر

4. Humidity
5. Precipitation
6. Evaporation
7. Temperature

1. Minimum Legible Delineation
2. Evans
3. Shary

مشخصی از بردار داده‌ها و سپس باز نمونه‌گیری تصادفی از آن مجموعه برای شبیه‌سازی دوره زمانی متعاقب یک دوره معین است. از مزایای استفاده از این الگوریتم در پیش‌بینی‌ها می‌توان به اجرای ساده، عدم نیاز به مرحله تخمین عامل‌ها، قابلیت مدل‌سازی غیرخطی، مؤثر بودن و عملکرد با بازدهی بالا در برخورد با تعداد دسته‌های زیاد از داده‌ها اشاره کرد (عزومی و عراقی‌نژاد، ۲۰۱۱).

#### ب. شبکه‌های عصبی مصنوعی<sup>۵</sup> (ANN)

شبکه‌های عصبی مصنوعی از جمله مدل‌های محاسباتی هستند که می‌توانند رابطه بین ورودی‌ها و خروجی‌های یک سیستم فیزیکی را حتی اگر پیچیده و غیرخطی باشد، با شبکه‌ای از گره‌ها که همگی به هم متصل هستند تعیین کنند. ساختار شبکه‌های عصبی، به نحوی است که نرون‌ها در دسته‌هایی که لایه نامیده می‌شوند مرتب می‌گردند. معماری معمول برای این شبکه‌ها شامل سه لایه ورودی، پنهان و خروجی است. لایه ورودی، داده‌ها را در شبکه توزیع می‌کند، لایه پنهان داده‌ها را پردازش و لایه خروجی نتایج را به‌ازای ورودی‌های مشخص استخراج می‌کند.

#### ج. فرایند گوسی<sup>۶</sup> (GP)

فرایند گوسی یک روش یادگیری ماشینی غیرپارامتری قوی برای ایجاد مدل‌های احتمال‌گرایانه جامع از مسائل دنیای واقعی است. فرایند گوسی یک فرایند تصادفی است که متشکل از مقادیر تصادفی در هر نقطه در یک دامنه زمانی یا مکانی است، به‌طوری که هر یک از متغیرهای تصادفی دارای توزیع نرمال‌اند. هر مجموعه متناهی از این متغیرهای تصادفی دارای توزیع نرمال چندمتغیره است. یک فرایند گوسی تولید داده‌هایی می‌کند که در طول این طیف قرار دارد، به‌طوری که هر زیرمجموعه متناهی دامنه یک توزیع گوسی چندمتغیره را دنبال می‌کند (راسموسن و ویلیامز<sup>۷</sup>، ۲۰۰۶).

به‌منظور تهیه نقشه پارامترهای اقلیمی از روش‌های زمین‌آمار در محیط نرم‌افزار ArcGIS برای پهنه‌بندی مقادیر آن‌ها در نقاط ایستگاه استفاده شد که با توجه به کم بودن تعداد ایستگاه از دو روش تابع پایه شعاعی<sup>۱</sup> و عکس فاصله وزنی<sup>۲</sup> استفاده گردید (راهنمای روش‌های توزیع مکانی عوامل اقلیمی با استفاده از داده‌های نقطه‌ای<sup>۳</sup>، ۲۰۱۲). با توجه به بالاتر بودن دقت روش تابع پایه شعاعی، نقشه نهایی بر مبنای روش تابع پایه شعاعی تهیه شد. با تهیه نقشه عامل‌ها، مقادیر آن‌ها در نقاط نمونه‌برداری استخراج گردید. نتایج ارزیابی دو روش تابع پایه شعاعی و عکس فاصله وزنی در جدول (۲) ارائه شده است.

### یادگیری ماشین

علم یادگیری ماشین از علوم مختلف از جمله آمار، هوش مصنوعی، یادگیری ماشین، شناسایی الگو و پایگاه داده نشئت گرفته است. در واقع این علوم ریشه‌های علم یادگیری ماشین هستند. الگوریتم‌های موجود در هوش مصنوعی و علم آمار کمک شایانی به یادگیری ماشین می‌کنند. فرایندهای یادگیری ماشین شامل سه مرحله آماده‌سازی داده، یادگیری مدل، ارزیابی و تفسیر مدل است. مراحل چهارگانه همراه با زیر مراحل و ورودی و خروجی‌ها در شکل (۲) نشان داده شده است.

با تهیه مجموعه داده‌های اقلیمی و ژئومورفومتری در نقاط نمونه‌برداری پوشش گیاهی، مدل‌های داده‌کاوی به‌منظور ارزیابی دقت این مدل‌ها در برآورد درصد پوشش گیاهی بر روی مجموعه داده‌ها اعمال شد. هر مجموعه داده به‌منظور اعمال مدل‌های داده‌کاوی به دو دسته داده‌های آموزش و آزمون (۷۰ درصد آموزش و ۳۰ درصد آزمون) تقسیم شد. مدل‌های استفاده‌شده در این مطالعه، شامل مدل نزدیک‌ترین همسایه، شبکه عصبی مصنوعی، فرایند گوسی، درخت تصمیم M5 و ماشین بردار پشتیبان بدین شرح است:

#### الف. k نزدیک‌ترین همسایه<sup>۴</sup> (KNN)

الگوریتم k نزدیک‌ترین همسایه شامل انتخاب تعداد

1. Radial Basis Function (RBF)
2. Invers Distance Wiegthed (IDW)
3. Guide to spatial distribution of climatic factors using point data
4. K-Nearest Neighbors

5. Artificial Neural Network  
6. Gaussian Process  
7. Rasmussen & Williams

جدول (۱): عامل‌های ژئومورفومتری استفاده‌شده در این تحقیق

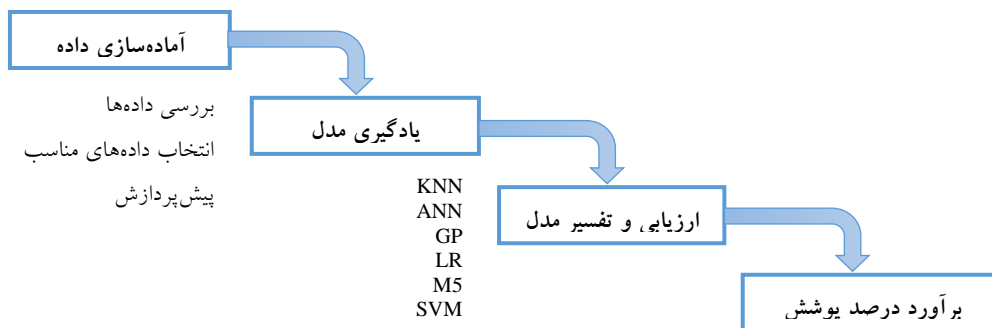
Table (1): The geomorphometric factors used in this research

Slope	شیب	Drainage Density	تراکم زهکشی
Profile Curvature	انحنای مقطع	Tangential Curvature	تانژانت انحنا
Topographic Position Index	شاخص موقعیت توپوگرافی	Longitudinal Curvature	انحنای طولی
Topographic Wetness Index	شاخص رطوبت توپوگرافی	Channel Network Base Level	سطح پایه آبراهه
Vertical Distance Channel Network	فاصله از آبراهه	Total Curvature	انحنای کل
Flow Directions	جهت جریان	Ls Factor	شاخص طول شیب
Plan Curvature	انحنای سطح	Elevation	ارتفاع
Aspects	جهت		

جدول (۲): بررسی دقت روش‌های پهنه‌بندی عامل‌های اقلیمی

Table (2): Assessment of the accuracy of zoning methods

عامل	روش پهنه‌بندی	میانگین	RMSE
رطوبت	RBF	۰/۱۵۴۵۴۵۱	۶/۳۹۶۹۷۱
	IDW	۰/۵۱۰۱۷۲۲	۷/۲۶۹۱۲۲
درجه حرارت	RBF	۰/۸۴۶۶	۲/۶۲
	IDW	۰/۵۴۱۲۲۳۳	۲/۵۸۹۶۱۴
تبخیر	RBF	-۵/۹۲۱۷۶۸	۵۲/۷۹۴۲۸
	IDW	-۱۷/۲۰۱۱۶	۵۷/۶۶۴۴
بارش	RBF	-۰/۲۹۴۶۰۷۹	۱/۱۷۶۹۰۴
	IDW	-۰/۵۲۱۰۶۰۳	۱/۲۷۹۰۰۵



شکل (۲): فلوچارت مراحل یادگیری ماشین (صنایع آباده و همکاران، ۲۰۱۴)

Figure (2): Flowchart of Machine Learning Steps (Sanaye Abadeh &amp; et al., 2014)

مستقل وجود دارد (رجایی و همکاران، ۲۰۰۹).

### ه درخت تصمیم‌گیری<sup>۲</sup> (M5)

ساختار یک مدل درختی شامل ریشه، گره‌های داخلی و برگ است. این مدل برخلاف مدل‌های درخت تصمیم معمول که کلاس یا رده‌های گسسته را به‌عنوان خروجی ارائه می‌کنند، یک مدل خطی چندمتغیره را برای داده‌ها در هر گره از مدل

### د. رگرسیون خطی<sup>۱</sup> (LR)

پیش‌بینی مقدار یک متغیر پیوسته بر اساس مقادیر سایر متغیرها بر مبنای یک مدل وابستگی خطی یا غیرخطی رگرسیون نامیده می‌شود. رگرسیون خطی برای مدل کردن مقدار یک متغیر کمی وابسته که بر رابطه خطی‌اش با یک یا چند متغیر مستقل بنا شده به کار می‌رود. در مدل رگرسیون خطی فرض بر این است که رابطه‌ای خطی (یا خط مستقیم) بین متغیر وابسته و هر متغیر

2. M5 rules

1. Liner Regression

درختی می‌سازد (ویتن و فرانک<sup>۱</sup>، ۲۰۰۷).

و. ماشین بردار پشتیبان<sup>۲</sup> (SVM)

ماشین‌های بردار پشتیبان دسته‌ای از روش‌های یادگیری با ناظر هستند که برای مسائل طبقه‌بندی و رگرسیون به کار می‌روند. این روش بر پایه نظریه یادگیری محاسباتی توسعه یافته است. مبنای کاری دسته‌بندی‌کننده SVM دسته‌بندی خطی داده‌هاست و در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب گردد که حاشیه اطمینان بیشتری داشته باشد. ماشین‌های بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرایی کمینه‌سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌شود (کریستینی و شوی-تیلور<sup>۳</sup>، ۲۰۰۰).

### ارزیابی عملکرد مدل‌ها

به منظور ارزیابی نتایج از ۳ معیار همبستگی، که جهت یک رابطه خطی را بین دو متغیر نشان می‌دهد و ضریب تبیین، که نوسان یک متغیر ارائه می‌کند و ریشه میانگین مربعات خطا، که میزان خطا را نشان می‌دهد استفاده شد.

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (۲)$$

$$R^2 = \left( \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \right)^2 \quad (۳)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \quad (۴)$$

که در این معادلات  $O_i$  مقادیر مشاهده‌ای،  $p_i$  داده‌های پیش‌بینی شده،  $\bar{p}$  میانگین داده‌های پیش‌بینی شده،  $\bar{O}$  میانگین داده‌های مشاهده شده،  $n$  تعداد داده‌هاست.

### وزن‌دهی عامل‌ها و شاخص‌ها

تمامی عامل‌های ورودی مدل، تأثیر و اهمیت یکسانی در پیش‌بینی ندارند. برخی از عامل‌ها دارای همبستگی و ارتباط بیشتری با خروجی مدل بوده و تأثیر بیشتری در پیش‌بینی‌ها داشته‌اند. وزن‌دهی عامل‌ها یکی از تکنیک‌های داده‌کاوی است.

در این مطالعه به منظور وزن‌دهی از الگوریتم ماشین‌برداری پشتیبان استفاده شده که این الگوریتم ضرایب بردار نرمال یک ماشین‌برداری پشتیبان خطی را به عنوان وزن ویژگی تعیین می‌کند (صنعی آباده و همکاران، ۲۰۱۴).

### نتایج

با توجه به نمونه‌برداری‌های انجام شده، متوسط پوشش ۶ درصد است. مهم‌ترین گونه‌های گیاهی مشاهده شده در منطقه شامل درمنه دشتی، اشنان، خارشتر، خارکوکو، قلم، پرند، پیچک درختچه‌ای، گون، گاو چاق کن و اروشیا می‌باشد.

با تعیین نقاط نمونه‌برداری و استخراج مقادیر عامل‌های ژئومورفومتری مقادیر آماری آن‌ها استخراج شد که در جدول (۳) ارائه شده است.

با انتخاب روش تابع پایه شعاعی که از میزان خطای پایین‌تری برخوردار بود و تهیه نقشه هریک از عامل‌ها مقادیر آماری آن‌ها استخراج شد. نتایج خلاصه آماری داده‌های اقلیمی در جدول (۴) ارائه شده است.

در بررسی عامل‌های مناسب برای برآورد درصد پوشش، نخست الگوریتم‌های مختلف روی داده‌های عامل‌های ژئومورفومتری و اقلیمی اعمال شد. با استفاده از معیارهای ارزیابی مناسب، دقت هریک از مدل‌ها در پیش‌بینی درصد پوشش با استفاده از عامل‌های ژئومورفومتری و اقلیمی ارزیابی شد. نتایج اجرای مدل‌ها روی داده‌های آموزش و آزمون بر مبنای عامل‌های ژئومورفومتری در جدول (۵) ارائه شده است. نتایج نشان داد که در مجموع داده‌های آموزش مدل فرایند گوسی با ریشه میانگین مربعات خطا ۲/۷۳ و ضریب تبیین ۰/۹۶ دارای بیشترین دقت و مدل رگرسیون خطی با ریشه میانگین مربعات خطا ۱۰/۱۹ و ضریب تبیین ۰/۶۷ دارای کمترین دقت هستند. همچنین در مجموع، داده‌های آزمون مدل فرایند گوسی با ریشه میانگین مربعات خطا ۱/۱۷ و ضریب تبیین ۰/۹۹ دارای بیشترین دقت و مدل رگرسیون خطی با ریشه میانگین مربعات خطا ۵/۳۳ و ضریب تبیین ۰/۵۹ دارای کمترین دقت در برآورد درصد پوشش هستند.

نتایج اجرای مدل‌ها روی داده‌های آموزش و آزمون بر مبنای عامل‌های اقلیمی در جدول (۶) ارائه شده است.

1. Witten & Frank
2. Support Vector Machine
3. Cristianini & Shawe-Taylor



نتایج نشان داد که در مجموع، داده‌های آموزش مدل درخت تصمیم‌گیری با ریشه میانگین مربعات خطا  $9/66$  و ضریب تبیین  $0/58$  دارای بیشترین دقت و مدل ماشین‌بردار پشتیبان با ریشه میانگین مربعات خطا  $12/28$  و ضریب تبیین  $0/34$  دارای کمترین دقت در برآورد درصد پوشش هستند.

جدول (۳): عامل‌های ژئومورفومتری مستخرج از DEM

Table (3): Geomorphometric factors extracted from DEM

عوامل	حداقل	حداکثر	میانگین	انحراف معیار
انحنای مقطع (متر)	-0/007	0/005	-0/09	0/002
تراکم زهکش	1	9	1/520	1/189
شیب (درصد)	0	0/151	0/043	0/033
تانژانت انحنا (متر)	-0/005	0/005	0	0/002
شاخص موقعیت توپوگرافی	-2/353	4/274	0/096	1/082
شاخص رطوبت توپوگرافی	7/627	24/774	10/754	2/796
فاصله از آبراهه (متر) اقلیدسی	-14/526	8/194	-2/027	4/93
جهت (درجه)	0	6/283	2/788	2/200
سطح پایه شبکه آبراهه	970/87	2423/24	1715/517	497/812
جهت جریان	0	7	2/272	2/305
انحنای کل (متر)	-0/019	0/020	0	0/008
انحنای طولی (متر)	-0/014	0/010	-0/08	0/005
شاخص طول شیب	0	2/511	0/666	0/586
انحنای سطح (متر)	-0/358	0/213	0/002	0/064
ارتفاع (متر)	980/518	2446/31	1727/654	495/687

جدول (۴): خلاصه آماری داده‌های اقلیمی

Table (4): A summary of the climate data

عوامل	حداقل	حداکثر	میانگین	انحراف معیار
تبخیر (mm)	361/492	386/97	373/055	6345
رطوبت (%)	21/310	24/579	23/488	1/021
بارندگی (mm)	4/258	6/174	5/304	0/656
دما (C)	21/290	27/003	24/017	2/093

جدول (۵): معیارهای ارزیابی داده‌های آموزش و آزمون بر مبنای عامل‌های ژئومورفومتری

Table (5): Assessment criteria for training and testing data based on Geomorphometric factors

مجموعه داده‌های آموزش			مجموعه داده‌های آزمون			روش‌های داده‌کاوی
R <sup>2</sup>	R	RMSE	R <sup>2</sup>	R	RMSE	
0/91	0/95	5/14	0/98	0/99	1/59	شبکه عصبی مصنوعی
0/81	0/90	6/05	0/63	0/80	4/78	k نزدیک‌ترین همسایه
0/96	0/98	2/73	0/99	0/99	1/17	فرایند گوسی
0/67	0/82	10/19	0/59	0/77	5/33	رگرسیون خطی
0/95	0/98	3/21	0/96	0/98	1/62	ماشین‌بردار پشتیبان
0/89	0/94	4/29	0/92	0/96	2/14	درخت تصمیم‌گیری

جدول (۶): معیارهای ارزیابی داده‌های آموزش و آزمون بر مبنای عامل‌های اقلیمی

Table (6): Assessment criteria for training and testing data based on climatic factors

مجموعه داده‌های آزمون			مجموعه داده‌های آموزش			روش‌های داده‌کاوی
R <sup>2</sup>	R	RMSE	R <sup>2</sup>	R	RMSE	
۰/۳۸	۰/۶۲	۱۰/۱۰	۰/۵۰	۰/۷۰	۱۲/۱۲	شبکه عصبی مصنوعی
۰/۵۶	۰/۷۵	۸/۸۵	۰/۵۶	۰/۷۵	۹/۷۷	k نزدیک‌ترین همسایه
۰/۵۷	۰/۷۶	۸/۸۳	۰/۵۸	۰/۷۶	۱۰/۴۳	فرایند گوسی
۰/۲۵	۰/۵۰	۱۰/۹۷	۰/۴۳	۰/۶۵	۱۲/۸۴	رگرسیون خطی
۰/۳۴	۰/۵۸	۱۲/۲۸	۰/۵۱	۰/۷۱	۱۳/۵۴	ماشین بردار پشتیبان
۰/۵۷	۰/۷۵	۸/۶۰	۰/۵۸	۰/۷۶	۹/۶۶	درخت تصمیم‌گیری

### وزن‌دهی پارامترها

درصد پوشش داشته و عامل انحنای طول دارای کمترین وزن و کمترین تأثیر در پیش‌بینی پوشش گیاهی در منطقه مورد مطالعه است. شکل (۳) نتایج وزن‌دهی عامل‌های ژئومورفومتری را نشان می‌دهد.

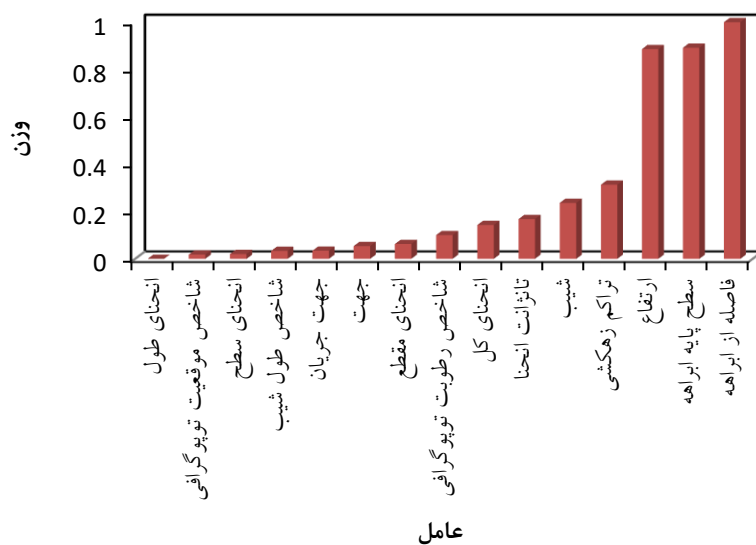
### نتایج وزن‌دهی عامل‌های اقلیمی

نتایج وزن‌دهی عامل‌های اقلیمی نشان داد که از میان این عامل‌ها، عامل رطوبت بیشترین وزن را دارد؛ یعنی بیشترین تأثیر را در پیش‌بینی درصد پوشش داشته است. عامل بارش کمترین وزن و کمترین تأثیر در پیش‌بینی پوشش گیاهی در منطقه مورد مطالعه را دارا بوده است. شکل (۴) نتایج وزن‌دهی عامل‌های اقلیمی را نشان می‌دهد.

تمامی پارامترهای ورودی مدل، تأثیر و اهمیت یکسانی در پیش‌بینی ندارند. برخی از پارامترها دارای همبستگی و ارتباط بیشتری با خروجی مدل بوده و تأثیر بیشتری در پیش‌بینی‌ها داشته‌اند. در این مطالعه به منظور وزن‌دهی از الگوریتم ماشین بردار پشتیبان استفاده شده است که این الگوریتم ضرایب بردار نرمال یک ماشین بردار پشتیبان خطی را به عنوان وزن ویژگی تعیین می‌کند.

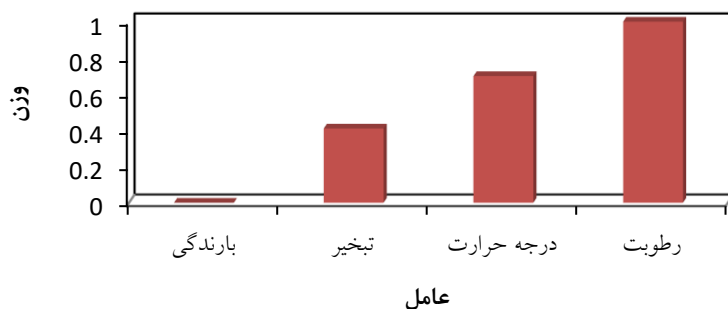
### نتایج وزن‌دهی عامل‌های ژئومورفومتری

نتایج وزن‌دهی عامل‌های ژئومورفومتری نشان داد که از میان این عوامل، عامل فاصله از آبراهه، سطح پایه آبراهه و ارتفاع بیشترین وزن را دارند؛ یعنی بیشترین تأثیر را در پیش‌بینی



شکل (۳): نمودار وزن‌دهی عامل‌های ژئومورفومتری در پیش‌بینی درصد پوشش

Figure (3): The weighting of Geomorphometric factors in predicting the percentage of coverage



شکل (۴): نمودار وزن‌دهی عامل‌های اقلیمی در پیش‌بینی درصد پوشش

Figure (4): The weighting of climatic factors in predicting the percentage of coverage

### بحث و نتیجه‌گیری

کمبود بارش و ظهور بارش‌های رگباری، هرزاب‌ها ایجاد می‌گردد که موجب می‌شود مسیل‌ها و فرورفتگی‌ها حاوی خاک‌های آبرفتی نرم‌تری بوده و درصد رطوبت بیشتری داشته و از حاصلخیزی بیشتری برخوردار باشند (جعفری و طویلی، ۲۰۱۰). لذا عامل فاصله از آبراهه و سطح پایه آبراهه می‌تواند از عوامل مؤثر روی درصد پوشش گیاهی این مناطق عنوان گردد. البته در اکثر مطالعات انجام‌شده در نواحی دیگر (محتشم‌نیا، ۲۰۱۱؛ عبدلهی و نادری، ۲۰۱۲؛ بهرامی و قربانی، ۲۰۱۶؛ سلامی و همکاران، ۲۰۱۷؛ نودهی و همکاران، ۲۰۱۴) نیز نشان داد که ارتفاع بیشترین تأثیر مثبت را بر ویژگی‌های رویشی درمنه داشته است. از میان عامل‌های اقلیمی، عامل رطوبت بیشترین تأثیر و بارش کمترین تأثیر در برآورد پوشش گیاهی داشته، که با نتایج مینگاد و یانگ (۲۰۱۳) و اشمیت و همکاران (۲۰۱۳) که رطوبت را یکی از مهم‌ترین عوامل مؤثر در پراکنش پوشش گیاهی می‌دانند مطابقت دارد؛ برای مثال، اشمیت و همکاران (۲۰۱۳) عامل رطوبت را یکی از عوامل مهم در پراکنش پوشش گیاهی در آفریقا می‌دانند. از آنجا که منطقه مورد مطالعه جزء مناطق نیمه‌خشک می‌باشد که متوسط بارندگی در آن کم است و اختلاف بارشی در نقاط نمونه‌برداری مشاهده نشده، عامل بارش تأثیر چندانی در برآورد پوشش در منطقه مورد مطالعه نداشته است. موغلی و همکاران (۲۰۱۳) نشان داده‌اند که عامل درجه‌حرارت مهم‌ترین نقش را در پراکنش گونه‌مرتعی رمس داشته است. البته تأثیر عوامل اقلیمی تابعی از شرایط هر منطقه بوده و در مناطق مختلف با توجه به

پوشش گیاهی به وسیله متغیرهای محیطی از جمله توپوگرافی و اقلیم کنترل می‌شود، به طوری که پراکنش هر گونه گیاهی در محدوده جغرافیایی خاص امکان‌پذیر است. از این رو، با تغییر تدریجی شرایط اکولوژیکی در یک منطقه و در طول گرادیان محیطی با گروهی از گونه‌های گیاهی با مشخصات گیاهی خاصی از جمله فرم رویشی، تیپ بیولوژیک و... مواجه خواهیم بود. لذا در این مطالعه، از عامل‌های ژئومورفومتری و اقلیمی به عنوان ورودی مدل به منظور ارزیابی دقت الگوریتم‌های یادگیری ماشین در پیش‌بینی درصد پوشش استفاده شد. نتایج حاصل از بررسی مدل‌های داده‌کاوی اعمال‌شده روی مجموعه داده‌ها نشان داد که مدل‌ها با ورودی عامل‌های ژئومورفومتری، دقت بیشتری نسبت به عامل‌های اقلیمی در پیش‌بینی درصد پوشش داشته‌اند؛ که می‌توان علت آن را در یکنواختی بیشتر عوامل اقلیمی در منطقه دانست. البته دقت بالای پارامترهای ژئومورفومتری استخراج‌شده نیز می‌تواند عامل مؤثری در بالا بودن دقت مدل‌ها در این مجموعه باشد. نتایج این قسمت از مطالعه نشان می‌دهد که در برخی از مناطق که نقص داده‌های اقلیمی به چشم می‌خورد، به‌کارگیری پارامترهای سهل‌الوصول ژئومورفومتری می‌تواند کمک شایانی در برآورد درصد پوشش گیاهی این مناطق داشته باشد. نتایج حاصل از وزن‌دهی عامل‌ها نیز نشان داد که از میان عامل‌های ژئومورفومتری، عامل فاصله از آبراهه، سطح پایه آبراهه و ارتفاع، بیشترین تأثیر را داشته‌اند. در مناطق بیابانی با توجه به

ساخت. همان طور که متکان و همکاران (۲۰۱۱) نیز در مطالعه خود با بررسی کارایی استفاده از تصاویر ماهواره‌ای و شبکه‌های عصبی مصنوعی در تخمین درصد پوشش گیاهی در مناطق خشک بافق یزد به این نتیجه رسیدند که شبکه عصبی، توانایی تخمین درصد پوشش گیاهی را با دقت مناسبی دارد.

از آنجا که گونه‌های درمنه و مطالعه اثر عوامل محیطی بر گسترش آن‌ها، اهمیت زیادی در برنامه‌ریزی و مدیریت عرصه‌های منابع طبیعی دارد و درمنه یکی از گیاهان بسیار مهم مراتع مرتفع کشور است، از نتایج این تحقیق می‌توان برای برنامه‌ریزی‌های کاربردی، حفظ و مدیریت این مراتع و همچنین احیا و بازسازی آن‌ها استفاده کرد. البته روابط بین گیاهان و عوامل محیطی بسیار پیچیده است و بررسی آن‌ها به مطالعات طولانی مدت و کامل نیاز دارد. بررسی بیشتر این مراتع در سایر نقاط کشور و بررسی عوامل زیست محیطی دیگر در پراکنش گیاهان این مراتع پیشنهاد می‌شود.

محدودیت‌های هر منطقه، نتایج یکسانی به دنبال نخواهد داشت. نتایج حاکی از آن است که هرچه نقاط نمونه‌برداری در فواصل نزدیک‌تری به آبراهه (به علت رطوبت بالای خاک اطراف ریشه) یا در ارتفاعات بالاتر باشد (زیرا با افزایش ارتفاع دمای محیط کاهش می‌یابد و میزان تبخیر و تعرق از سطح گیاه کم شده و میزان رطوبت در گیاه افزایش یافته و شرایط محیطی برای رشد گیاه به وجود می‌آید)، درصد پوشش بالاتری نیز دارند. از آنجا که منطقه مورد مطالعه جزء مناطق نیمه‌بیابانی می‌باشد که متوسط بارندگی در آن کم است و اختلاف بارشی در نقاط نمونه‌برداری مشاهده نشده، عامل بارش تأثیر چندانی در برآورد پوشش نداشته است. ارزیابی کلی نتایج حاکی از دقت فراوان مدل‌های داده‌کاوی برای پیش‌بینی درصد پوشش گیاهی است. بررسی نتایج این مطالعه و سایر مطالعات انجام شده نشان می‌دهد که می‌توان با به‌کارگیری مدل‌ها با دقت بیشتر و تعیین عوامل مؤثر در پوشش، تعیین درصد پوشش در مناطق مختلف را امکان‌پذیر

## منابع

1. Arzani, H., Abedi, M., 2015. Measure vegetation. Tehran University Press. Cover 2, P: 305.
2. Abdollahi, J., Naderi, H., 2012. Soil and Topographic variation influencing the growing factors of *Artemisia sieberi* in steppic rangeland, Nodoushan- Yazd. Watershed Management (Pajouhesh& Sazandegi), No. 97.
3. Azmi, M., Araghinejad, S., 2011. Development of K-Nearest Neighbour Regression Method in Forecasting River Stream Flow. Journal of Water and Wastewater. Vol. 23, Issue 2: 108- 119.
4. Bahrami, B., Ghorbani, A., 2016. The Influence Environmental Factors on the Distribution of Plant Species in the Southeast Rangelands of Sabalan. Journal of Natural Ecosystem of Iran, 7(1): 33-44.
5. Bates, J., Svejcar, T., Miller, R., Angell, A., 2006. The effect of precipitation timing on sagebrush steppe vegetation. J. Arid Environ. 64(4): 670-679.
6. Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C.E., Franklin, J., 1999. A Neural Network method for efficient vegetation mapping. Remote Sensing of Environment, Vol. 70, No. 3: 326– 338.
7. Cristianini, N., Shawe- Taylor, J., 2000. An introduction to support vector machines. Cambridge University Press.
8. Darvishzadeh, R. Matkan, A.A. Hosseiniasl, A., Ebrahimi Khusefi, M., 2012. Estimation of vegetation fraction in the Central arid region of Iran using satellite images (Case study: Sheitoor basin, Bafgh). Arid Biome Scientific and Research Journal. 2(1): 25- 38
9. Evans, I.S., 1972. General geomorphometry, derivatives of altitude and descriptive statistics. Spatial Analysis in Geomorphology, 17-90.
10. Guide to spatial distribution of climatic factors using point data, 2012. No. 585, Vice President of Strategic Planning and Control.
11. Gurgel, H. C., Ferreira, N. J., 2003. Annual and Interannual Variability of NDVI in Brazil and its Connections with Climate. International Journal of Remote Sensing, 24 (18): 3595–3609.
12. Hengl, H., 2006. Finding the right pixel size,

- Computers & Geosciences, 32(9): 1238-1298.
13. Jafari, M., Tavili, A., 2010. Restoration of arid and desert areas, University of Tehran.
  14. Jiangbo, X., Tong, L., 2010. Characterization of spatial scaling relationships between vegetation pattern and topography at different directions in Gurbantunggut desert, China, *Ecological Complexity*, 7(2): 234-242.
  15. Khan Behbini, M., 2015. The Relationship between Changes in Longitude and Soil Factors with Distribution of Vegetation in Velang Pasture in Golestan Province, range management graduate thesis, Gonbad-Kavoos University.
  16. Matkan, A.A., Darvishzadeh, R., Hosseiniasl, A., Ebrahimi Khusefi, M., 2011. Capability using satellite images and neural network in estimating vegetation percentage in arid region. *Journal of Environmental Erosion*, 1, 7- 27.
  17. Mesdaghi, M., 2015. Rangeland in Iran. P: 328.
  18. Minggagud, H., Yang, J., 2013. Wetland plant species diversity in sandy land of a semi-arid inland region of China, *Plant Biosystems*, Vol, 1. No, 147. pp: 25-32.
  19. Moghli, M., Afifi, M.A., Yousefi, A., 2013. Investigation of the role of climatic factors in the distribution of *Hammada salicornica* (Case study: Larestan and Tabas), *The Sustainable Agriculture and Natural Resources*.
  20. Mohtashamnia, S., 2011. Investigation of the most important environmental factors affecting the distribution of *Artemisia sieberi* in Fars province. *Journal of Natural Ecosystem of Iran*, 1(3): 75-86.
  21. Nodehi, N., Akbarlou, M., Sepehry, A., Vahid, H., 2014. Effects of Topographical Factors on Distribution of Plant Communities in Semi-Steppe Grasslands (Case Study: Ghorkhud Region, Northern Khorasan Province, Iran), *Journal of Rangeland Science*, Volume 4.
  22. Rajaei, T., Mirbagheri, S.A., Zounemat-Kermani, M., Nourani, V., 2009. Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Science of the total environment*, 407(17): 4916-4927.
  23. Rasmussen, C.E., Williams, C.K.I., 2006. *Gaussian Processes for Machine Learning*, the MIT Press, ISBN: 026218253X. Massachusetts Institute of Technology.
  24. Salami, A., Sepehri, A., Akbarlo, M., 2017. Effect of environmental factors on vegetation cover of Inche Bouron ranges in Golestan province, The First National and International Conference of Food Industry and Organic Products in Iran
  25. Sani Abade, M., Mahmoudi, S., Taherparvar, D., 2014. *Data Mining Applications (2nd Edition)*, Niaz-e-Danesh Publication, Tehran.
  26. Schmidet, M., Klein, D., Conrad, C., Dech, S., Paeth, H., 2013. On the relationship between vegetation and climate in tropical and northern Africa. *Theoretical and Applied Climatology*, 115 (1-2): 341-353.
  27. Shary, P.A., 1995. Land surface in gravity points classification by a complete system of curvatures. *Mathematical Geology* 27(3): 373- 390.
  28. Sugier, B., 1996. *Vegetation atmosphere.* (france: Dominos Flammarion), p: 107.
  29. Witten, I. H., Frank, E., 2005. *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann Publication.
  30. Yazdanshenas, H., jafari, M., Azarnivand, H., Arzani, H., Nasiri, M., 2013. Investigating the Effects of Soil Factors on Biodiversity in Plant Communities of Karvan Rangeland (Case Study: Isfahan Province, Iran), *Journal of Rangeland Science*, 4(1): 34- 42.
  31. Yousif Elnour, Y., Zhongqin, L., Omer Said, M., Muhammad Naveed, A., Feiteng, W., Yanqun, B., Bo, Z., 2017. Correlation between Climate Factors and Vegetation Cover in Qinghai Province, China. *Journal of Geographic Information System*, 9(4): 403-419.

# Assessment of the Efficiency of Climatic factors and geomorphometry in predicting vegetation percentages based on machine learning processes

Zaynab Mirashkari<sup>1</sup>, Majid Sadeghinia<sup>2</sup>, Mostafa Shirmandi<sup>3\*</sup>, Maryam Asadi<sup>4</sup>

Received: 24/12/2018

Accepted: 13/07/2019

## Expanded abstracts

**Introduction:** Rangelands are natural ecosystems having large genetic resources. Since plant vegetation is the bed of life on earth and changes under the influence of surrounding environmental elements, using environmental element can highly contribute to estimate vegetation percent more accurately. Two effective elements which can contribute to estimate the vegetation distribution are climatic elements and geomorphometric. Nowadays, one of new techniques which have attracted much attention to estimate vegetation percent is machine learning process which is able to establish a relationship between various variables of environmental conditions with the least costs and workforce. Therefore, in this study, geomorphometric and climatic elements and data mining techniques have been applied to address the vegetation percent.

**Materials and Methods:** The studied region is a part of Yazd-Ardakan plain and Nadoshan region. Sampling and measuring vegetation percent have been carried out in using transects and plots. In order to extract the geomorphometric elements, digital elevation models and SAGA software were utilized and also, seven meteorological stations were regarded to achieve the climatic elements. In the current research, to investigate the impact of different climatic and geomorphometric elements on vegetation percent estimate, data mining models such as artificial neural network, the nearest neighbor, support vector machine, decision tree, Gaussian process and linear regression were used.

Artificial neural network: is one of computational models which can determine the relationships between inputs and outputs of one physical system and a network of connected nodes even if they are complicated and nonlinear.

The nearest neighbor: involves selecting a certain number of data vectors and random sampling from the set in order to simulate the time interval followed by a certain period.

Support vector machine: is an efficient learning system based on theory of optimization applying the inductive principle of structural error minimization which leads to a total optimum response.

Decision tree: is a method to estimate the discrete functions which are strong against the confused data and are capable to learn the terminology with two different fields.

Gaussian process: is a random one consisted of random values in each point in a time or location domain so that each random variable has a normal distribution.

Linear regression: is applied to model the value of a dependent quantitative variable based on a linear relationship with one or more independent variables.

To assess the models and compare the results, such assessment criteria as RMSE, correlation coefficient and coefficient of determination have been used. Here, to weigh the input parameters of support vector machine algorithm, normal vector coefficients related to a linear support vector machine were specified as the weights.

**Results:** The study indicated that data mining models are able to estimate vegetation percent more accurately. Using geomorphometric elements, data mining models have shown that Gaussian process model had the most accuracy in the set of training and test data. As well, in applying the models on climatic data, it has been reported that decision tree had the most accuracy in the set of training and test data to estimate the vegetation percent.

**Discussion and Conclusion:** Vegetation is controlled by such environmental variables as geomorphometry and climate. The results have indicated that geomorphometric elements are of more impact on vegetation percent prediction as compared to climatic ones. Weighing results showed that such geomorphometric elements as distance to waterway, waterway baseline and elevation and such climatic ones as humidity affecting the vegetation growth rate were of the highest weigh and impact in vegetation percent prediction in the desired region.

**Keywords:** Coverage percentage, *Artemisia sieberi*, Climatic factors, Geomorphometric factors, machine learning.

1. M.Sc. Graduate of Range Management, Department of Nature Engineering, Faculty of Agriculture & Natural Resources, Ardakan University

2. Assistant Professor, Department of Nature Engineering, Faculty of Agriculture & Natural Resources, Ardakan University

3. Assistant Professor, Department of Horticultural Science, Faculty of Agriculture & Natural Resources, Ardakan University; shirmardi@ardakan.ac.ir

4. Ph.D. Student of Watershed Management, Department of Reclamation of Arid and Mountainous Regions Engineering, Faculty of Natural Resources, Tehran University

DOI: 10.22052/deej.2018.7.24.45