

شناسایی کانون های مستعد گردشگری براساس عوامل محیطی در حوضه کمه استان اصفهان با استفاده از الگوریتم جنگل تصادفی

شهره مرادپور*: دانش آموخته دکترای ژئومورفولوژی، دانشکده علوم جغرافیایی و برنامه ریزی، دانشگاه

اصفهان، ایران Sh.moradpour@geo.ui.ac.ir

شماره تماس ۰۹۱۳۸۸۶۰۰۷۱

مجید غیاث: استادیار گروه گردشگری، دانشکده علوم جغرافیایی و برنامه ریزی، دانشگاه اصفهان، ایران

m.ghias@geo.ui.ac.ir

چکیده

در این پژوهش، از روش جنگل تصادفی، یکی از الگوریتم های یادگیری ماشین، برای شناسایی و پیش بینی مناطق مستعد گردشگری در حوضه کمه سمیرم استان اصفهان استفاده شده است. بیست منطقه شامل چشمه ها، آبشارها، امامزاده ها و بناهای تاریخی به عنوان داده های ورودی مدل در نظر گرفته شده اند. متغیرهای ارتفاع، شیب، جهت شیب، پوشش گیاهی، فاصله از رودخانه و جاده، تراکم رودخانه و جاده و موقعیت شیب نسبی، پس از انتخاب بر اساس ضریب تورم واریانس (VIF)، به عنوان متغیرهای پیش بینی کننده مورد استفاده قرار گرفتند. نتایج ارزیابی مدل با استفاده از معیارهای R^2 ، RMSE و MAE نشان دهنده عملکرد بسیار مطلوب مدل جنگل تصادفی با R^2 برابر با ۰/۸۹، RMSE برابر با ۱/۳۳ و MAE برابر با ۱/۱۳ می باشد. فاصله از جاده ها (۰/۱۰)، جهت شیب (۰/۸۴/۳۳)، ارتفاع (۰/۶۲/۸۳) و پوشش گیاهی (۰/۱۴/۷۲) به عنوان مهم ترین متغیرهای پیش بینی کننده شناسایی شدند. نقشه حاصل از مدل، پتانسیل بالای مناطق مرکزی و شمالی حوضه کمه را برای جذب گردشگر نشان می دهد. این مطالعه نشان می دهد که روش های یادگیری ماشین ابزار قدرتمندی برای پیش بینی مناطق مستعد گردشگری هستند.

کلیدواژه: حوضه کمه شهرستان سمیرم، هوش مصنوعی، الگوریتم یادگیری ماشین، آزمون عامل تورم واریانس

مقدمه

یکی از راهبردهای که برای تقویت نواحی محروم و دارای قابلیت توسعه مطرح شده است، توسعه و گسترش گردشگری در مناطقی است که دارای پتانسیل های لازم برای توسعه گردشگری هستند. از طرف دیگر رشد و گسترش شهرنشینی و نیاز جوامع شهرنشین به بازدید از جاذبه های طبیعی، سبب سرازیر شدن خیل عظیمی از جمعیت شهر نشین به نواحی و مکان های دارای جاذبه های گردشگری شده است.

هوش مصنوعی و الگوریتم های یاد گیری ماشین از پدیده های نوظهور هستند که توانسته اند به شدت در زندگی بشری تاثیر گذار باشند. هوش مصنوعی دگرگونی های عجیبی را به وجود آورده است. صنعت گردشگری نیز از این تغییرات مستثنی نیست. همه فرایند سفر توسط هوشمندسازی و هوش مصنوعی از قبیل رزرو اسکان، رزرو برای وسیله حمل و نقل هوایی، دریایی، ریلی و زمینی، نقشه راهنمای مجازی، رستوران های بین شهری، وضعیت شهرهای مقاصد سفر و دیدنی هایی که در مقاصد وجود دارد از قبیل منابع طبیعی، منابع تاریخی و فرهنگی، ارزش ها و هنجارهای اجتماعی، هنرهای هفتگانه و تهیه بلیط های اماکن دیدنی، اپلیکیشن های راهنمایی کننده برنامه ریزی خواهد شد (لوفان، ۲۰۱۴: ۲) و حتی شناسایی و پیش بینی مکانی مناطق گردشگری پذیر ناشناخته و مستعد توسط هوش مصنوعی امکانپذیر خواهد شد.

روش های یادگیری ماشین مانند ماشین بردار پشتیبان^۱ و مدل های شبکه های عصبی^۲ توجه فزاینده ای را برای پیش بینی سری های زمانی اقتصادی به خود جلب کرده اند (Huang et al. 2005: 2513؛ Kim, 2003: 307)، اما تلاش های کمی برای پیش بینی تقاضای گردشگری انجام شده است (Chen & Wang, 2007: 215 Hong) (et al. 2011: 1881؛ Wu et al., 2012: 4769؛ Akin, 2015: 64) خلاصه ای کامل از پیش بینی با شبکه های عصبی را می توان در Zhang و همکاران (۱۹۹۸) یافت. در حالی که شبکه پرسپترون چندلایه (MLP) معمولاً برای پیش بینی تقاضای گردشگری به کار رفته است (Padhi and Aggarwal, 2011: 725؛ Lin et al., 2011: 14؛ Claveria and Torra, 2014: 220)، در حوزه گردشگری، ماشین یادگیری نقش مهمی در تقویت جنبه های مختلف صنعت ایفا می کند. الگوریتم های یادگیری از ماشین برای تجزیه و تحلیل حجم وسیعی از داده ها و استخراج بینش های ارزشمندی که به بهبود تجربه کلی سفر برای افراد کمک می کند (Şeker, 2023: 54).

^۱ Support vector machines

^۲ Artificial Neural Networks

در مورد شناسایی و پیش بینی مکانی، مکان های گردشگرپذیر با استفاده از هوش مصنوعی مطالعات گسترده ای در جهان و ایران صورت نگرفته است و در زیر به تعدادی که مطالعاتی که در این زمینه صورت گرفته است اشاره می گردد.

(Bernard, 2010) در پژوهشی به بررسی پتانسیل های ژئوتوریسم و بالابردن سطح آگاهی از چشم انداز در مورد میراث زمین شناختی و محیط زیست در جزیره در شرق ماداگاسکار پرداختند و به این نتیجه رسیدند که تهیه طرح توسعه ژئوتوریسم و حفاظت ضروری است. آموزش و تفسیر که از اجزای اساسی ژئوتوریسم می باشد، باید تأکید بر اطلاعات جغرافیایی در طرح مشخص شود.

(Petrović, 2020) به پیش بینی تعداد گردشگران با استفاده از یادگیری ماشین " پرداختند و بیان کردند که هنگامی که می خواهیم مقدار چندین گام زمانی آینده را پیش بینی کنیم، مدل باید یاد بگیرد که طیف وسیعی از مقادیر آینده را پیش بینی کند. بنابراین، برخلاف مدل هایی که یک گام زمانی آینده را پیش بینی می کنند و تنها یک مقدار آینده را پیش بینی می کنند، این مدل ها دنباله ای از مقادیر آینده را پیش بینی می کنند (Petrović, 2020:28).

(Ullah Khan et. Al., 2023) در پیش بینی و طبقه بندی فعالیت های کاربر با استفاده از ماشین مدل های یادگیری از داده های شبکه اجتماعی مبتنی بر مکان بیان کردند که مدل های یادگیری عمیق در طبقه بندی داده های مرتبط با گردشگری از سایر روش ها بهتر عمل می کنند، که نشان دهنده این است که این تکنیک های پیشرفته می توانند به پیش بینی های دقیق تری از مقاصد و رفتارهای گردشگری منجر شوند.

نوری و نوروزی (۱۳۸۶) نیز در پژوهشی با عنوان ارزیابی توان محیطی برای توسعه گردشگری در دهستان چغاخور، از لایه های اطلاعاتی نظیر شیب، پوشش گیاهی، خاک و غیره استفاده و در نهایت با روی هم گذاری لایه های اطلاعاتی موجود، دهستان چغاخور را به چهار واحد توریستی از لحاظ نوع تفرج گسترده و متمرکز تقسیم بندی کرده اند (نوری و نوروزی، ۱۳۸۶: ۱۳).

تقوایی و همکاران (۱۳۹۰) در بررسی تحت عنوان مکان یابی دهکده های گردشگری با استفاده از سیستم اطلاعات جغرافیایی و مدل SWOT (نمونه موردی: ساحل دریاچه کافترا) از طریق تلفیق ۲۳ لایه اطلاعاتی از نقشه های با مقیاس ۱:۲۵۰۰۰ سازمان نقشه برداری کشور و نیز نقشه مقیاس ۱:۱۰۰۰۰۰ سازمان زمین شناسی و اکتشافات معدنی، در محیط GIS و نیز بهره گیری از تکنیک SWOT، بهترین مکان ایجاد دهکده گردشگری در ارتفاعات جنوبی مشرف بر دریاچه تعیین گردید. در نهایت، راهبردهایی برای، رشد و توسعه صنعت گردشگری در ساحل دریاچه کافترا ارائه شد (تقوایی و همکاران، ۱۳۹۰: ۹۹)

مومنی و همکاران (۱۳۹۲) در مکانیابی بهینه مناطق نمونه گردشگری در استان سیستان و بلوچستان نتایج این تحقیق نشان می دهد که مناطق گردشگری دره سرباز، تیس و تنگه فنوج مناسب ترین گزینه ها برای

انتخاب به عنوان مناطق نمونه گردشگری و تهیه طرح جامع گردشگری هستند (مومنی و همکاران، ۱۳۹۲:۱۱۳).

سلیقه و جاوید (۱۳۹۳) در پهنه بندی مکان گزینی گردشگری فصلی شهرستان سمیرم بر پایه مقایسه روش تحلیل سلسله مراتبی و شاخص اقلیمی به منظور تعیین پهنه های مناسب برای گردشگری در چهارفصل از دو مدل بیوکلیماتیک و تحلیل سلسله مراتبی AHP استفاده نمودند. نتایج حاصل از تطبیق دو مدل مذکور مشخص نمود که خروجی مدل تحلیل سلسله مراتبی نسبت به مدل بیوکلیماتیک پهنه های واقعی تری در اختیار گردشگران قرار می دهد (سلیقه و جاوید، ۱۳۹۳: ۲۳۹).

فروش نیا و کرم (۱۴۰۱) برای پهنه بندی پتانسیل های ژئوتوریسمی شهرستان سمیرم از پارامترهای طبیعی و انسانی به عنوان شاخص های مدل آنتروپی شانون استفاده کردند که نتایج این مطالعه نشان داد که غار دنگزلو و تخت سلیمان به عنوان باارزش ترین منابع زمین شناختی منطقه برای توسعه ژئوتوریسم، باید مورد توجه بیشتری قرار گیرند (فروش نیا و کرم، ۱۴۰۱: ۱۲۹).

غیاث و مرادپور (۱۴۰۳) در شناسایی کانون های مستعد گردشگری براساس عوامل محیطی در غرب استان اصفهان با استفاده از الگوریتم جنگل تصادفی به این نتیجه دست یافتند که مدل جنگل تصادفی عالی ترین عملکرد پیش بینی را با بالاترین R^2 (۰/۹۱)، RMSE پایین (۱/۰۶) و MAE (۱/۱۳) را دارا می باشد. شیب، جهت شیب، پوشش گیاهی، فاصله از رودخانه ها و جاده ها به عنوان پیش بینی کننده های مهم شناسایی شدند.

نقطه مثبت پژوهش حاضر نسبت به سایر پژوهش های انجام شده، به کار گیری سیستم اطلاعات جغرافیایی، داده های سنجش از دور، داده های نقشه های موضوعی (فاصله از جاده ها، آبراهه ها، تراکم جاده ها و آبراهه ها) تلفیق کلیه لایه های اطلاعاتی موجود در زمینه گردشگری، جهت انتخاب محل های جدید مستعد گردشگری است .

حوضه کمه که در شهرستان سمیرم واقع شده است باتوجه به جاذبه های تاریخی و طبیعی و پتانسیل های جاذب گردشگری، متأسفانه به دلیل عدم توجه به زیر ساخت ها و عدم برنامه ریزی و مدیریت جامع، شاهد کم بودن گردشگر در شهر های خارج از مرکز استان بوده ایم و لازم است شهرهای دیگر این استان را که دارای پتانسیل های بالای گردشگری هستند را شناسایی و معرفی نمود.

روش شناسی پژوهش

توسعه روزافزون فناوری های GIS و سنجش از دور، امکان استفاده از مدل های آماری و یادگیری ماشین را برای پیش بینی مکانی و به ویژه در حوزه گردشگری فراهم کرده است. این مدل ها با تحلیل ارتباط بین عوامل طبیعی و توسعه گردشگری، به تصمیم گیران کمک می کنند تا مکان های مناسب برای سرمایه گذاری در صنعت

گردشگری را شناسایی کنند. هرچند مدل‌های آماری و یادگیری ماشین هر دو برای این منظور قابل استفاده هستند، اما مدل‌های یادگیری ماشین به دلیل توانایی آن‌ها در مدیریت داده‌های پیچیده و غیرخطی، معمولاً دقت بالاتری دارند.

روند تولید نقشه‌های پیش‌بینی مکانی مناطق مستعد گردشگری به طور کلی شامل چهار مرحله اصلی است: (۱) جمع‌آوری اطلاعات دقیق در مورد مکان‌های گردشگری موجود در منطقه مورد مطالعه، (۲) شناسایی و بررسی عوامل جغرافیایی موثر بر انتخاب مکان‌های گردشگری، (۳) ساخت و توسعه مدل‌های مناسب برای پیش‌بینی و تولید نقشه، و (۴) ارزیابی و اعتبارسنجی نتایج حاصل از مدل‌ها.

در گام نخست این پژوهش، با هدف تهیه نقشه مناطق مستعد گردشگری، اطلاعات دقیقی از ۲۰ نقطه گردشگری جمع‌آوری شد. پس از آن، عوامل موثر بر جذب گردشگر به مناطق مختلف شناسایی و مورد بررسی قرار گرفت. در ادامه، با استفاده از مدل‌های مناسب، نقشه پهنه‌بندی مناطق مستعد گردشگری تهیه شد و عوامل موثر در شکل‌گیری این مناطق تحلیل گردید. برای انتخاب مهم‌ترین عوامل موثر بر گردشگری، از شاخص تورم واریانس (VIF) استفاده شد. این شاخص به ما کمک می‌کند تا ارتباط بین عوامل مختلف را بررسی کنیم و از وجود چندخطی بودن بین آن‌ها اطمینان حاصل کنیم. بر اساس نتایج حاصل از این تحلیل، ۹ عامل اصلی شامل ارتفاع، شیب، جهت شیب، فاصله از رودخانه‌ها، تراکم رودخانه‌ها، فاصله از جاده‌ها، تراکم جاده‌ها، شاخص پوشش گیاهی^۱ (NDVI) و موقعیت شیب نسبی به عنوان عوامل موثر شناسایی شدند. در مرحله بعد، با ترکیب اطلاعات مربوط به این عوامل محیطی با موقعیت مکانی نقاط گردشگری، مدل‌های پیش‌بینی توسعه داده شدند. برای آموزش و ارزیابی این مدل‌ها، از نرم‌افزار R استفاده شد. به این ترتیب که ابتدا داده‌ها به صورت تصادفی به دو دسته آموزشی (۷۰٪) و آزمایشی (۳۰٪) تقسیم شدند. سپس مدل بر روی داده‌های آموزشی آموزش داده شد و عملکرد آن بر روی داده‌های آزمایشی ارزیابی گردید.

در گام سوم، پس از طراحی سناریوهای مختلف، با بهره‌گیری از روش‌های یادگیری ماشین در محیط برنامه‌نویسی R، مدل‌های پیش‌بینی توسعه یافتند. در مرحله چهارم، دقت مدل‌های ساخته‌شده با استفاده از معیارهای آماری همچون ضریب تعیین (R^2)، ریشه میانگین مربعات خطا (RMSE) و میانگین خطای مطلق (MAE) ارزیابی شد.

برای پیش‌بینی مناطق مستعد گردشگری، از مجموعه‌ای از داده‌های محیطی استفاده گردید. این داده‌ها شامل تصاویر ماهواره‌ای (سنجش از دور)، اطلاعات استخراج شده از مدل رقومی ارتفاع (DEM)، داده‌های هیدرولوژیکی و فاصله از راه‌های ارتباطی بود. این متغیرها به عنوان ورودی مدل‌های یادگیری ماشین عمل کرده و در فرآیند مدل‌سازی به کار گرفته شدند (شکل ۱).

متغیرهای کمکی سنجش از دور

^۱ Normalized Difference Vegetation Index

برای ایجاد شاخص‌های پوشش گیاهی، تصاویر سنتینل ۲ با کمتر از ۱۰٪ پوشش ابر، از وبسایت گوگل ارث انجین دانلود شد. از نرم‌افزار ساگاسی آاس^۱ برای استخراج متغیر شاخص گیاهی تفاضلی نرمال شده (NDVI)، استفاده شد (شکل ۲). این شاخص دارای بهترین توان دینامیک می‌باشد. شاخص NDVI با فرمول داده شده توسط روز و همکاران (۱۹۷۴) محاسبه شد:

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \quad \text{رابطه (۱)}$$

که در آن NIR، بازتاب در باند مادون قرمز نزدیک و RED بازتاب در باند قرمز است. این شاخص در محدوده ۱- تا ۱+ می‌باشد ولی در عمل کمتر از ۱ و بیشتر از ۱- می‌باشد.

متغیرهای کمکی ویژگی‌های توپوگرافی

توپوگرافی را می‌توان به عنوان یکی از عوامل جذب گردشگر در نظر گرفت. از پایگاه داده Alaska Satellite Facility و از ماهواره ALOS برای دستیابی به نقشه رقومی ارتفاع با قدرت تفکیک ۱۲/۵ متر استفاده شد. ویژگی‌های زمین که از نقشه رقومی ارتفاع تهیه گردید عبارتند از: ارتفاع، جهت شیب^۲، شیب^۳ و موقعیت شیب نسبی^۴ می‌باشد در محیط ساگاسی آاس تهیه گردید.

فاصله از جاده و تراکم جاده ها

توسعه زیرساخت‌ها، از جمله راه، جریان حمل و جابه جایی مسافر، انگیزه ی گردشگران را برای بازدید از مکان های مختلف افزایش می‌دهد (Sharply, 2002).

فاصله از آبراهه ها و تراکم آبراهه ها

آبراهه های منطقه مورد مطالعاتی از روی تصویر مدل رقومی ارتفاع با استفاده از الحاقی ArcHydro در محیط نرم افزار ArcGIS استخراج گردید و در نهایت با استفاده از تابع فاصله اقلیدسی نقشه فاصله از آبراهه تهیه شد.

جدول ۱: متغیرهای محیطی مورد استفاده برای پیش‌بینی مکان‌های مستعد گردشگری

ردیف	معیار	روش مورد استفاده	ردیف	معیار	روش مورد استفاده
۱	ارتفاع	ماهواره ALOS	۵	فاصله تا رودخانه‌ها	تهیه شده در نرم افزار Arc GIS10.8
۲	زاویه شیب	زوربرگن و تورن	۶	تراکم آبراهه	تهیه شده در نرم افزار Arc GIS10.8
۳	جهت شیب	زوربرگن و تورن	۷	فاصله تا جاده‌ها	تهیه شده در نرم افزار Arc GIS10.8

¹ System for Automated Geoscientific Analysis

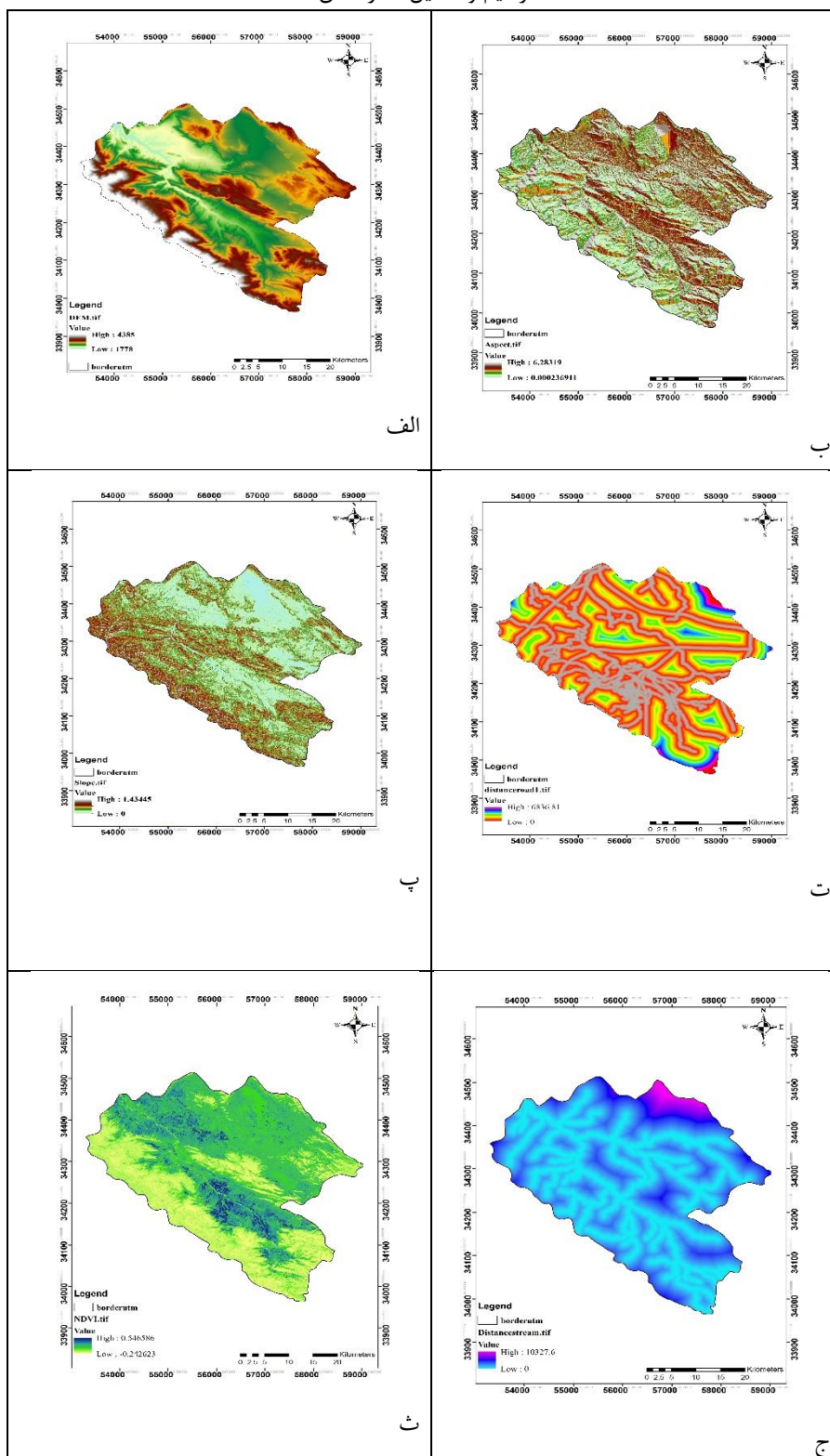
² Aspect

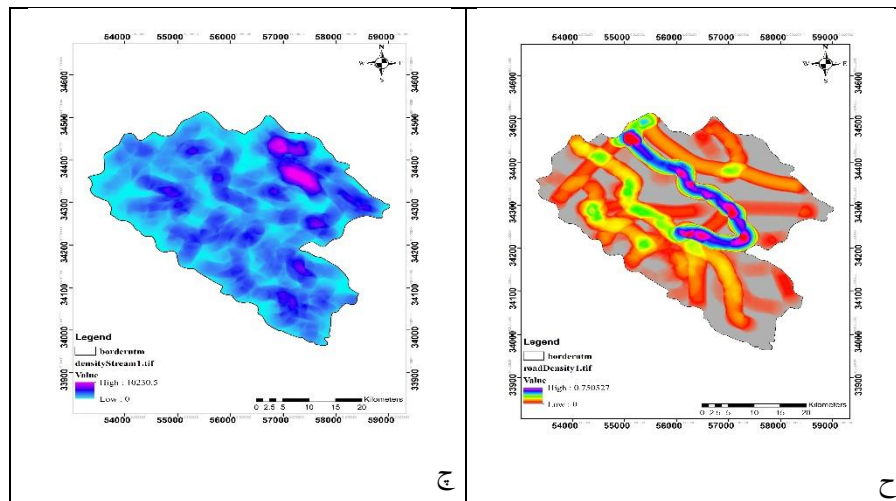
³ Slope

⁴ Relative Slope Position

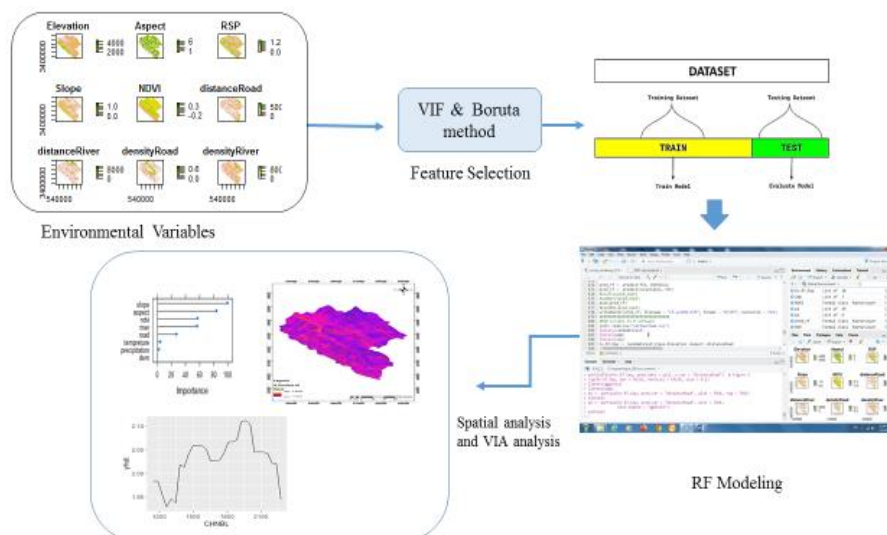
۴	تراکم رودخانه ها	تهیه شده در نرم افزار Arc GIS10.8	۸ شاخص پوشش گیاهی (NDVI)	تصاویر ماهواره سنتینل ۲، در نرم افزار SsgaGis
۹	موقعیت شیب نسبی	تصاویر ماهواره سنتینل ۲، در نرم افزار SsgaGis		

ترسیم و تحلیل: نگارندگان





شکل ۱: نقشه پارامترهای مورد ارزیابی در مدل جنگل تصادفی (الف: نقشه مدل رقومی ارتفاع، ب: نقشه جهت شیب، پ: نقشه شیب ت: نقشه فاصله از جاده ها، ث: شاخص NDVI، ج: فاصله از رودخانه ها، چ: نقشه تراکم رودخانه ها ح: نقشه تراکم جاده ها. ترسیم و تحلیل: نگارندگان



شکل ۲: فرآیند کلی و مراحل مختلف پژوهش. ترسیم و تحلیل: نگارندگان

چارچوب مفهومی

نرم افزار برنامه نویسی Rstudio:

R یک ابزار قدرتمند برنامه نویسی و محیط محاسباتی است که به طور ویژه برای تحلیل آماری و داده کاوی طراحی شده است. این نرم افزار متن باز، بر پایه زبان های برنامه نویسی S و Scheme دارد و طیف وسیعی از روش های آماری، از جمله مدل سازی پیشرفته، آزمون های آماری، تحلیل سری های زمانی و تکنیک های یادگیری ماشین را در اختیار کاربران قرار می دهد (طالبی و همکاران، ۱۳۹۷: ۴۵).

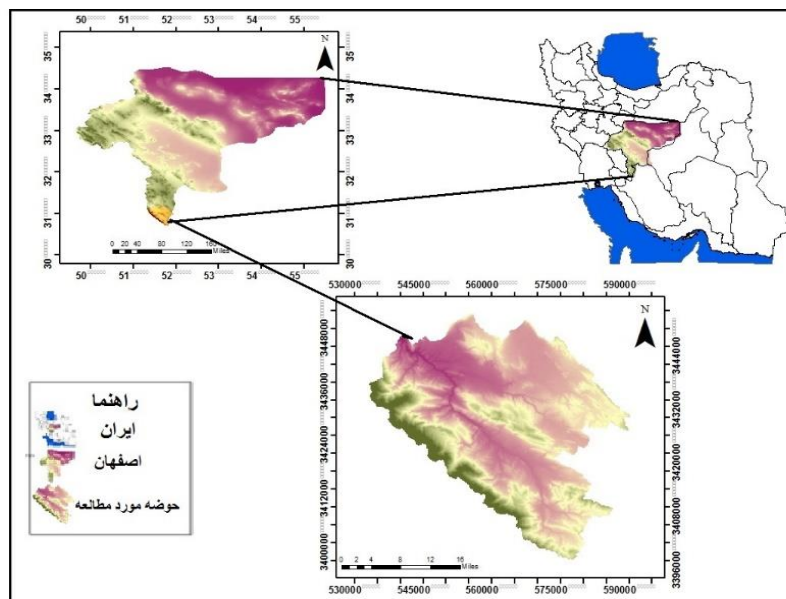
الگوریتم یادگیری ماشین جنگل تصادفی!

شامل مجموعه ای تصادفی شده از رگرسیون درختی می باشد که با اجرای دسته های متعددی از رگرسیون های درختی و نهایتاً تجمیع نتایج همه این مدل ها، تعیین و ارائه یک پیش بینی واحد تحت عنوان روش جنگل تصادفی می باشد. در این روش، داده ها به طور تکراری برای بدست آوردن ارتباط بین متغیر پاسخ و متغیرهای مستقل و انجام تخمین، جداسازی می شوند. در روش جنگل تصادفی به جای رشد یک درخت، صدها یا هزاران درخت تولید می شود (Allbed et al. 2014:1). در این روش چندین درخت تصمیم آموزش داده می شود تا دقت مدل را افزایش دهند. نتیجه حاصل پیش بینی گروهی از درختان تصمیم گیری است (Breiman, 2001: 5). در روش یادگیری جنگل تصادفی، هر درخت تصمیم گیری با استفاده از یک نمونه تصادفی از داده های آموزشی، آموزش می بیند. همچنین انتخاب متغیرهای پیش بینی کننده نیز که برای تقسیم بندی گره ها استفاده می شود، به صورت تصادفی انجام می گیرد. در این روش، تعداد متغیرهای کمکی (mtry) (از یک تا کل متغیرها) در هر زیر مجموعه و تعداد درختان مورد استفاده (ntree) از ۵۰۰ تا ۱۰۰۰ تعیین و در مدل لحاظ می شود (Brungard, 2015:63).

معرفی منطقه مورد مطالعه:

منطقه مورد مطالعه، طبق نام گذاری جاماب در واحد هیدرولوژیک ۲۳۱۴ به نام کمه در حدفاصل ۳۰ درجه و ۴۲ دقیقه تا ۳۱ درجه و ۱۱ دقیقه عرض شمالی و ۵۱ درجه و ۲۱ دقیقه تا ۵۱ درجه و ۵۶ دقیقه طول شرقی در جنوب اصفهان و در مرز با استان های فارس و کهگیلویه و بویراحمد قرار دارد که از نظر تقسیمات سیاسی بخشی از شهرستان سمیرم در استان اصفهان است (شکل ۳) (سازمان جغرافیایی نیروهای مسلح). مساحت این حوضه ۱۶۶۳/۵ کیلومتر مربع است و قسمت اعظم آن در دهستان های پادنا، علیا، پادنا و وسطی و قسمتی نیز از دهستان پادنا و سفلی که در جنوب شهرستان سمیرم قرار دارد، واقع شده است. همچنین این حوضه دهستان حنا را نیز در بر دارد. این حوضه جزئی از حوضه کارون می باشد (نوجوان، ۱۳۹۸: ۱۵۰)،^۲ در کوهپایه های شرقی کوهستان دنا واقع شده و از نظر آب و هوایی سردسیر است. ارتفاع متوسط این حوضه ۲۵۰۰ متر از سطح دریا و شیب متوسط آن ۲۸ درصد است (عمو شاهی، ۱۳۹۹).

^۱ Random forest



شکل ۳: موقعیت حوضه کمه ترسیم و تحلیل: نگارندگان

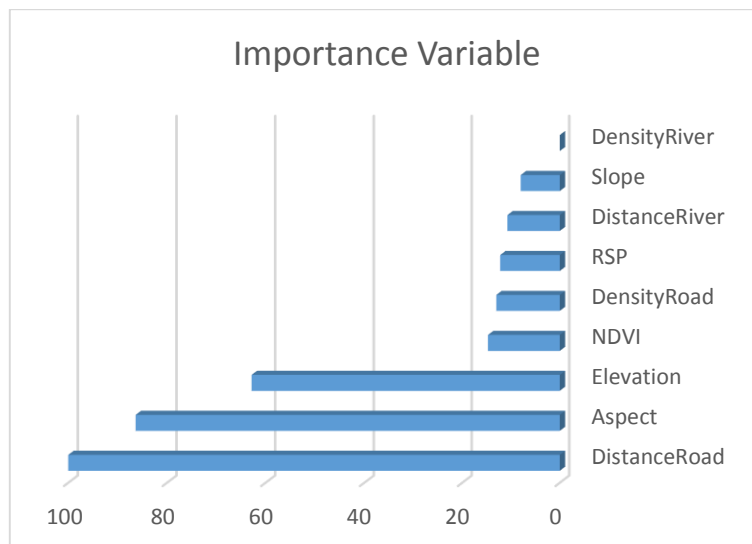
نتایج

انتخاب متغیرها

برای ارزیابی تأثیر هم‌خطی چندگانه متغیرها بر مدل، از شاخص تورم واریانس^۱ (VIF) استفاده شد. بدین ترتیب، متغیرهایی که VIF آن‌ها بیش از ۱۰ بود (به عنوان آستانه رایج هم‌خطی بالا در نظر گرفته می‌شود) از مدل حذف شدند. از مجموع ۱۵ متغیر اولیه، ۹ متغیر شامل ارتفاع، شیب، جهت شیب، فاصله تا رودخانه‌ها، تراکم رودخانه‌ها، تراکم جاده‌ها، فاصله تا جاده‌ها، شاخص پوشش گیاهی (NDVI)، موقعیت شیب نسبی برای مدل‌سازی انتخاب شدند. این انتخاب بر اساس تحلیل اهمیت متغیرها با استفاده از بسته caret در محیط R انجام شد. با توجه به تحلیل اهمیت متغیر (VIA)^۲، به ترتیب فاصله از جاده‌ها، جهت شیب، ارتفاع، پوشش گیاهی و تراکم جاده‌ها مهم‌ترین متغیرها برای انتخاب مکان‌های گردشگری پیش‌بینی شدند (شکل ۴).

^۱ variance inflation factor

^۲ variable importance analysis



شکل ۴: اهمیت نسبی هر یک از متغیرهای کمکی محیطی (%). ترسیم و تحلیل: نگارندگان

مدلسازی و ارزیابی مدل

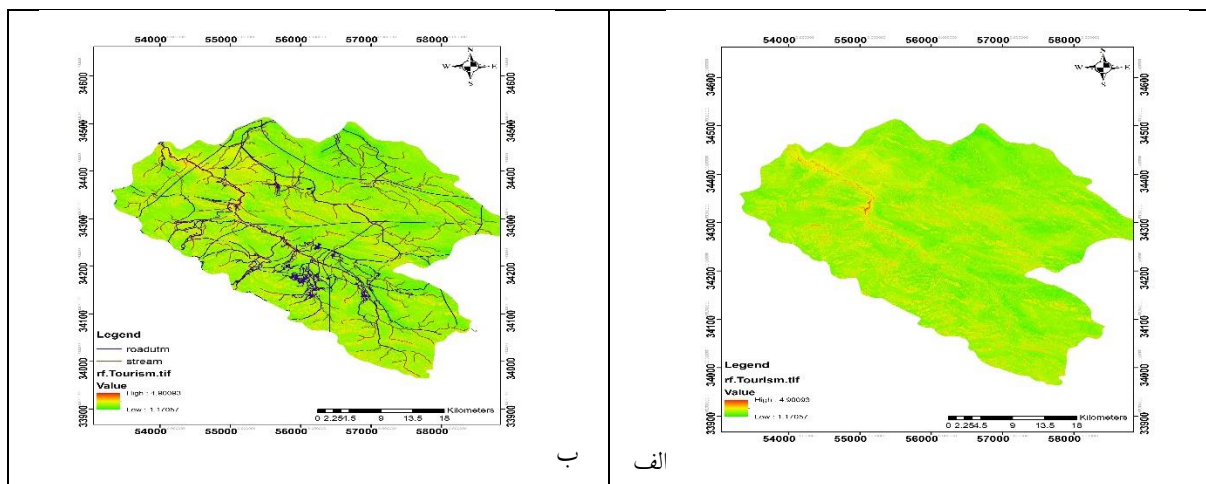
پس از انتخاب مهم ترین متغیرها محیطی از طریق آزمون VIF ، اقدام به تلفیق متغیرهای محیطی با مناطق گردشگری نموده (شکل)، در این رویکرد، داده های کامل را به صورت تصادفی به مجموعه های آموزشی و آزمایشی در محیط نرم افزار R تقسیم می کنیم. سپس آموزش مدل را روی مجموعه آموزشی انجام داده و از مجموعه تست برای اعتبارسنجی استفاده می شود (شکل ۲). در حالت ایده آل، داده ها را به ۷۰:۳۰ تقسیم می کنند مدل های یادگیری ماشین که در این مطالعه جنگل تصادفی می باشند، برای آموزش استفاده می شوند. سه شاخص آماری که برای ارزیابی عملکرد پیش بینی مدل در نظر گرفته شده است شامل ضریب تعیین (R^2)، ریشه میانگین مربعات خطا (RMSE)، میانگین خطای مطلق (MAE) می باشد. براساس نتایج به دست آمده، همان طور که در جدول ۲ مشاهده می شود. برای پیش بینی مکان های مستعد گردشگری مدل جنگل تصادفی با بالاترین ضریب تعیین ($R^2=0/89$)، و پایین ترین ریشه میانگین مربعات خطا ($1/33$) RMSE و میانگین قدرمطلق خطا ($1/13$) MAE را می توان قوی ترین مدل برای پیش بینی مکان های مستعد گردشگری در نظر گرفت. ضریب تعیین نزدیک به یک نشان دهنده عملکرد بهتر مدل و برآورد دقیق با خطای کمتر است (Zeraatpisheh et al., 2022: 6 ; Naimi et al., 2021: ۱۵).

جدول ۲: معیارهای اعتبارسنجی (۱۰ برابر اعتبار متقاطع) برای پیش بینی مکان های مستعد گردشگری (برای اعتبارسنجی ۳۰ درصد داده ها به صورت تصادفی کنار گذاشته شده). ترسیم و تحلیل: نگارندگان

Location	Model	سناریو		
		R^2	RMSE	MAE
	RF	۰/۸۹	۱/۳۳	۱/۱۳

توزیع مکانی مکان های مستعد گردشگری

نقشه پیش بینی مکانی با استفاده از ۲۰ نقطه گردشگری، متغیرهای محیطی و مدل جنگل تصادفی تهیه گردید. همانطور که مشاهده می شود مناطق مرکزی و شمالی منطقه بهترین مکان ها برای گردشگری انتخاب شده اند (شکل ۵). در این مدل مکان های به عنوان مکان های گردشگر پذیر پیشنهاد شده است که دارای بیشترین جاده ها و جریان های رودخانه ای می باشد (شکل ۶).



شکل ۵: الف: نقشه پیش بینی مکانی مربوط به مکان های مستعد برای گردشگری ب: نقشه پیش بینی مکانی مربوط به مکان های مستعد برای گردشگری به همراه جاده ها و جریان های آب. ترسیم و تحلیل: نگارندگان

بحث و نتیجه گیری

رونق صنعت گردشگری، بویژه در نواحی کم برخوردار، می تواند تحولی شگرف در ارتقای سطح معیشت، کاهش بیکاری، شکوفایی اقتصادی و بهبود کیفیت زندگی ساکنان پدید آورد. شناسایی و معرفی ظرفیت های بالقوه گردشگری، پیش نیاز دستیابی به چنین آرمانی است. این مطالعه نیز به همین مهم اختصاص یافته است.

در مطالعه حاضر اهمیت برخی از عوامل محیطی در شناسایی مکان های مستعد گردشگری در جنوب غرب استان اصفهان با استفاده از الگوریتم جنگل تصادفی مورد بررسی قرار گرفت که به ترتیب: فاصله از جاده ها (۱۰۰٪)، جهت شیب (۸۴/۳۳٪)، ارتفاع (۶۲/۸۳٪)، شاخص پوشش گیاهی (۱۴/۷۲٪)، مهمترین متغیرها برای پیش بینی مکان های مستعد گردشگری شناسایی شدند. همچنین نتایج نقشه پیش بینی مکانی نشان می دهد که مکان های مستعد برای پذیرش گردشگر مرکز و نقاط شمالی منطقه می باشد زیرا دارای جاده اصلی، بیشترین تراکم رودخانه ها، ارتفاع پایین تر نسبت به سایر مناطق حوضه (۱۷۷۰ متر)، جهت شیب به سمت جنوب و جنوب غرب و پوشش گیاهی متوسط می باشند. بررسی نتایج این پژوهش با پژوهش های

گذشته نشان می دهد که تقوایی و همکاران در مکانیابی دهکده های گردشگری با استفاده از سیستم اطلاعات جغرافیایی از طریق تلفیق ۲۳ لایه و با بهره گیری از تکنیک SWOT بهترین مکان ایجاد دهکده گردشگری در ارتفاعات جنوبی مشرف بر دریاچه تعیین گردید.

احمدی در مطالعه ای به پهنه بندی مکان های مستعد گردشگری در استان ایلام با استفاده از داده های اقلیمی به وسیله مدل AHP نموده و به این نتیجه دست یافتند که برای تعیین مکان های مناسب گردشگری نمی توان به عوامل اقلیمی اکتفا کرد.

فروش نیا و کرم در ارزیابی و پهنه بندی پتانسیل ژئواکوتوریسمی شهرستان سمیرم برای مدل سازی از پارمترهای طبیعی و انسانی موجود در منطقه به عنوان شاخص هایی برای آنالیز شانون استفاده شد. این شاخص ها شامل ارتفاع، شیب، فاصله از رودخانه، فاصله از ژئوسایت، فاصله از قله، فاصله از امامزاده، فاصله از دشت، فاصله از کاربری اراضی، فاصله از واحد اراضی، فاصله از دریاچه، فاصله از آبشار، فاصله از غار، فاصله از چشمه، فاصله از گسل، فاصله از جاده، فاصله از روستا، مناطق حفاظت شده، اقلیم می باشد که نتایج نشان داد که بزرگ ترین موانع توسعه گردشگری و زمین گردشگری در شهرستان سمیرم، ناشناخته بودن این منطقه و منابع بالقوه آن و عدم وجود امکانات در کمترین سطح برای گردشگران است که در صورت رفع این موانع می توان از توانمندی این شهرستان برای جذب گردشگران استفاده نمود.

داداش پورمقدم و همکاران در ارزیابی پتانسیل گردشگری شهری، رویکرد مدل سازی ترکیبی جدید با ترکیب K-mean، فازی منطق، و یک شبکه عصبی مصنوعی (ANN) استفاده کردند. در همبستگی بین عوامل شرطی سازی گردشگری شهری و نقشه های گردشگری، مشخص شد که عواملی از جمله فاصله از مراکز پذیرایی، کیفیت مصالح ساختمانی، فاصله از مراکز تاریخی فاصله از مراکز بهداشتی و درمانی، حداکثر دما و فاصله از پارکینگ در افزایش گردشگر شهری موثر هستند.

براساس نتایج به دست آمده می توان بیان کرد که روش های یادگیری ماشین در بسیاری از زمینه های مطالعاتی، از جمله بخش گردشگری، به طور گسترده می تواند مورد استفاده قرار بگیرد.

منابع:

- احمدی، م. ۱۳۸۹. شناسایی پهنه بندی مناسب برای توسعه ژئوتوریسم در استان ایلام با استفاده از سیستم اطلاعات جغرافیایی. پایان نامه کارشناسی ارشد، دانشکده ادبیات دانشگاه تربیت معلم تهران. ۱۶۰.
- تقوایی، م.، تقی زاده، م.، کیومرثی، ح. ۱۳۹۰. مکان یابی دهکده های گردشگری با استفاده از سیستم اطلاعات جغرافیایی و مدل SWOT (نمونه موردی: ساحل دریاچه کافت). مجله برنامه ریزی محیطی، ۲۲ (۲): ۹۹-۱۲۰.
- سلیقه، م.، بهشتی جاوید، ا. ۱۳۹۳. پهنه بندی مکان گزینی گردشگری فصلی شهرستان سمیرم بر پایه مقایسه روش تحلیل سلسله مراتبی و شاخص اقلیمی. جغرافیا (فصلنامه علمی- پژوهشی و بین المللی انجمن جغرافیایی ایران). سال دوازدهم، شماره ۴۱: ۲۳۹-۲۵۵.
- طالبی، ع.، گودرزی، س.، پور قاسمی، ح. (۱۳۹۷). بررسی امکان تهیه نقشه خطر زمین لغزش با استفاده از الگوریتم جنگل تصادفی (محدوده مورد مطالعه: حوضه آبخیز سردار آباد، استان لرستان). مجله مخاطرات محیط طبیعی، دوره هفتم، شماره ۱۶، ۴۵-۶۴. doi: [10.22111/JNEH.2017.3213](https://doi.org/10.22111/JNEH.2017.3213)
- عموشاهی، ناهید، ۱۳۹۹. پهنه بندی خطر زمین لغزش با استفاده از مدل حداکثر آنتروپی (مطالعه موردی: حوضه ی آبخیز کمه)، پایان نامه کارشناسی ارشد، دانشگاه اصفهان، گروه جغرافیای طبیعی، استاد راهنما رضا ذاکری نژاد
- غیاث، م.، مرادپور، ش. (۱۴۰۳). شناسایی کانون های مستعد گردشگری براساس عوامل محیطی در غرب استان اصفهان با استفاده از الگوریتم جنگل تصادفی. مطالعات مدیریت گردشگری، آماده انتشار. <https://doi.org/10.22054/tms.2024.18247>
- فروش نیا، س.، کرم، ا. ۱۴۰۱. ارزیابی و پهنه بندی پتانسیل ژئواکتوریسمی شهرستان سمیرم. سنجش از دور و سامانه اطلاعات جغرافیایی در منابع طبیعی، ۱۳ (۲): ۱۰۷-۱۲۹.
- لوفان، ع. ۱۴۰۲. واکاوی اثربخشی رویکرد هوشمندسازی و به ویژه رویکرد هوش مصنوعی بر صنعت گردشگری جوامع. چهاردهمین کنفرانس بین المللی پژوهش های مدیریت و علوم انسانی در ایران، تهران، ۲۹-۳۰. ۱-۱۰.
- مومنی، م.، تقی پور جاوی، م.، مستغاثی، ش. ۱۳۹۲. مکان یابی بهینه مناطق نمونه گردشگری (مطالعه موردی: استان سیستان و بلوچستان). جغرافیایی فضای گردشگری، ۲ (۸)، ۱۱۳-۱۴۱. SID. <https://sid.ir/paper/214269/fa>
- نوری، ه.، نوروزی آورگانی، ا. ۱۳۸۶. ارزیابی توان محیطی برای توسعه گردشگری در دهستان چغاخور، مجله پژوهشی دانشگاه اصفهان، ۲۲ (۱): ۱۳.

نوجوان، م.، شاه زیدی، س.، داودی، م.، امین الرعایایی، ه. ۱۳۹۸، پهنه بندی خطر زمین لغزش با استفاده از تلفیق دو مدل فرآیند تحلیل سلسله مراتبی و فازی (مطالعه موردی: حوضه آبخیز کمه، استان اصفهان)، پژوهش‌های ژئومورفولوژی کمی، سال هفتم، شماره ۴، ۱۵۹ – ۱۴۷.

Allbed, A., Kumar, L. and Aldakheel, Y. Y. 2014. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region. *Geoderma*. 230–231: 1–8. doi: 10.1016/j.geoderma.2014.03.025.

Akin, M. (2015). A novel approach to model selection in tourism demand modeling. *Tourism Management*, 48, 64–72. <http://dx.doi.org/10.1016/j.tourman.2014.11.004>.

Breiman, L. 2001. Random Forests. *Mach. Learn.* 45: 5–32. doi: 10.1023/A:1010933404324.

Bernard JE. 2010. Australia's Geoheritage: History of Study, A New Inventory of Geosites and Applications to Geotourism and Geoparks. *Geoheritage*, 2(1): 39-56. doi:<https://doi.org/10.1007/s12371-010-0011-z>.

Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A. and Edwards, T. C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*. 239–240: 68–83. doi: 10.1016/j.geoderma.2014.09.019.

Chen, K. Y., Wang, C. H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, 28: 215–226. <https://doi.org/10.1016/j.tourman.2005.12.018>.

Claveria, O., Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling*, 36, 220–228. <https://doi.org/10.1016/j.econmod.2013.09.024>.

Dadashpour Moghaddam, M., Ahmadzadeh, H., Valizadeh, R. (2022). A GIS-Based Assessment of Urban Tourism Potential with a Branding Approach Utilizing Hybrid Modeling. *Spat. Inf. Res.* (2022) 30(3):399–416. <https://doi.org/10.1007/s41324-022-00439-4>.

Huang, W., Nakamori, Y., Wang, S. Y. 2005. Forecasting stock market movement direction with support vector machine. *Computers & Operations Research* 32, 2513–2522.

Kim K. 2003. Financial Time Series Forecasting using support vector machines. *Neurocomputing* 55, 307-319.

Lin, C., Chen, H., Lee, T. (2011). Forecasting tourism demand using time series, artificial neural networks and multivariate adaptive regression splines: Evidence from Taiwan. *International Journal of Business Administration*, 2, 14–24. <http://dx.doi.org/10.5430/ijba.v2n2p14>.

Naimi, S. , Ayoubi, S. , Zeraatpisheh, M. and Dematte, J. A. M. (2021). Ground observations and environmental covariates integration for mapping of soil salinity: A machine learning-based approach. *Remote Sens.* 13:4825. doi: 10.3390/rs13234825.

Petrović, D., Čebić, B., Beljić, D. (2020). Predicting the number of tourist using machine learning. *Turističko poslovanje*. 28-41. DOI 10.5937/turpos0-31845.

Padhi, S. S., Aggarwal, V. (2011). Competitive revenue management for fixing quota and price of hotel commodities under uncertainty. *International Journal of Hospitality Management*, 30, 725–734. <http://dx.doi.org/10.1016/j.ijhm.2010.12.007>.

Şeker, F. 2023. Evolution of Machine Learning in Tourism: A Comprehensive Review of Seminal Research. *Journal of Artificial Intelligence and Data Science (JAIDA)*. Vol. 3, No. 2, pp. 54-79 2023. <https://dergipark.org.tr/pub/jaida>.

Sharply, Richard. 2002. Managing Sustainable and the challenges of tourism diversification, tourism management, Vol. 23.

Ullah Khan, N., Wan, W, Riaz, R., Jiang, SH., Wang, X. 2023. Prediction and Classification of User Activities Using Machine Learning Models from Location-Based Social Network Data. *Appl. Sci.* 2023, 13, 3517. <https://doi.org/10.3390/app13063517>.

Wu, Q., Law, R., Xu, X. (2012). A sparse Gaussian process regression model for tourism demand forecasting in Hong Kong. *Expert Systems with Applications*, 39, 4769–4774. <https://doi.org/10.1016/j.eswa.2011.09.159>.

Hong, W., Dong, Y. Chen, L., Wei S. (2011). SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Applied Soft Computing* 11, 1881–1890. <https://doi.org/10.1016/j.asoc.2010.06.003>.

Zeraatpisheh, M. , Garosi, Y. , Owliaie, H. R. , Ayoubi, S. , Taghizadeh-Mehrjardi, R. , Scholten, T. , Xu, M. (2022). Improving the spatial prediction of soil organic carbon using environmental covariates selection: A comparison of a group of environmental covariates. *Catena* 208:105723. <http://dx.doi.org/10.1016/j.catena.2021.105723>.