



دانشگاه علوم کشاورزی و منابع طبیعی گزن

نشریه پژوهش‌های علوم و فناوری چوب و جنگل
جلد بیست و سوم، شماره چهارم، ۱۳۹۵
<http://jwfst.gau.ac.ir>

تهیه نقشه خطر وقوع آتش‌سوزی با استفاده از الگوریتم‌های ماشین بردار پشتیبان، جنگل تصادفی و شبکه عصبی مصنوعی (مطالعه موردی: پارک ملی گلستان، شمال شرقی ایران)

*محمدامین اسحاقی^۱ و شعبان شنتایی جویباری^۲

^۱دانش‌آموخته کارشناسی‌ارشد، گروه جنگلداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان،
^۲دانشیار گروه جنگلداری، دانشکده علوم جنگل، دانشگاه علوم کشاورزی و منابع طبیعی گرگان
تاریخ دریافت: ۱۳۹۴/۰۳/۱۸؛ تاریخ پذیرش: ۱۳۹۴/۱۰/۰۱

چکیده

سابقه و هدف: پیش‌بینی مکانی احتمال وقوع آتش‌سوزی و تهیه نقشه احتمال خطر وقوع آتش در عرصه‌های طبیعی یکی از راه‌های جلوگیری و مدیریت آتش‌سوزی است. هدف از این تحقیق پهنه‌بندی خطر احتمال وقوع آتش‌سوزی در پارک ملی گلستان با استفاده از الگوریتم‌های ناپارامتریک شبکه عصبی مصنوعی، ماشین بردار پشتیبان و جنگل تصادفی می‌باشد.

مواد و روش‌ها: در این مطالعه ۱۰۰ نقطه از محل وقوع آتش‌سوزی جهت مدل‌سازی احتمال خطر وقوع آتش به کار برده شد. عوامل مؤثر بر وقوع آتش‌سوزی شامل پوشش گیاهی، فیزیوگرافی، اقلیمی و انسانی شناسایی و نقشه‌های عوامل فوق تهیه شدند. به منظور مدل‌سازی، منطقه به سطوح تصمیم‌گیری و مدل‌سازی یک هکتاری تقسیم‌بندی شد و ارزش رقومی لایه‌های عوامل تأثیرگذار بر وقوع آتش‌سوزی در محل سطوح یک هکتاری استخراج و استاندارد گردید. مدل‌سازی احتمال خطر وقوع آتش‌سوزی با الگوریتم‌های ناپارامتریک با استفاده از ۷۰ درصد نقاط آتش‌سوزی به عنوان نمونه‌های تعلیمی انجام شد. نقشه احتمال خطر وقوع آتش‌سوزی به دست آمده به ۴ طبقه کم خطر،

*مسئول مکاتبه: amin.eshaghi@gmail.com

متوسط خطر، پرخطر و خطرناک پهنه‌بندی گردید. ارزیابی صحت طبقه‌بندی نقشه‌های حاصل از مدل‌سازی با معیار صحت کلی با استفاده از ۳۰ درصد نقاط آتش‌سوزی باقیمانده صورت پذیرفت. یافته‌ها و نتیجه‌گیری کلی: نتایج نشان داد الگوریتم جنگل تصادفی با صحت کلی ۷۵ درصد توانست احتمال خطر وقوع آتش‌سوزی را بهتر از سایر الگوریتم‌ها پیش‌بینی کند. هم‌چنین از نظر میزان مطابقت وقوع آتش‌سوزی با نتایج الگوریتم‌ها مشخص شد که همه الگوریتم‌ها توانستند منطقه را به خوبی از نظر احتمال وقوع آتش‌سوزی طبقه‌بندی نمایند به طوری که بیش از ۸۰ درصد نقاط آتش‌سوزی در مناطق با کلاسه‌های پرخطر و خطرناک واقع شدند.

واژه‌های کلیدی: نقشه خطر آتش‌سوزی، جنگل تصادفی، شبکه عصبی مصنوعی، ماشین‌بردار پشتیبان، پارک ملی گلستان

مقدمه

جنگل‌های مرطوب هیرکانی در شمال ایران با برآورد سطح تقریبی ۹۶۵۰۰ هکتار و تراکم ناچیز با کمتر از ۱ درصد مساحت کل کشور یادگار دوران سوم زمین‌شناسی هستند (۳۶). از جمله مهمترین عوامل تهدید کننده این جنگل‌ها احتمال بالای وقوع آتش‌سوزی در آن است. به‌طور متوسط وقوع ۱۳۰ آتش‌سوزی در هر سال باعث سوختن سطحی برابر ۵۴۰۰ هکتار در این جنگل‌ها شده است (۱)، (۱۱، ۵۰). از جمله مناطق در معرض خطر وقوع آتش‌سوزی در جنگل‌های هیرکانی پارک ملی گلستان در شمال شرقی ایران است. پارک ملی گلستان با وسعت ۹۱۸۹۵ هکتار، نخستین پارکی است که در ایران عنوان پارک ملی را به خود اختصاص داده است (۳۵). در سالیان اخیر حدود ۱۰۸۴۵ مورد آتش‌سوزی در جنگل‌های کشور به ثبت رسیده است (۲۱). بنابراین لزوم پیش‌گیری و مقابله با این بحران امری ضروری است. به‌منظور به حداقل رساندن آسیب‌های وارده به جنگل در اثر آتش‌سوزی لازم است شناخت کامل از فرآیند آتش‌سوزی‌های گذشته داشته و پس از شناخت کامل فرآیندهای زمانی و مکانی می‌بایست تحلیل مکانی وقوع آتش‌سوزی‌ها در هر منطقه به‌منظور طراحی استراتژی‌های حفاظتی صورت گیرد. لازمه این طراحی توانایی در پیش‌بینی، تحلیل مکانی و تهیه نقشه پهنه‌بندی احتمال خطر آتش‌سوزی است (۳۲). مناطق خطر به مناطقی گفته می‌شود که احتمال شروع آتش‌سوزی و انتشار آن در این مکان‌ها بیشتر است. بنابراین از طریق تهیه نقشه احتمال خطر

آتش‌سوزی و شناسایی مناطق با خطر ریسک آتش‌سوزی بالا و اجرای تمهیدات مدیریت پیشگیری و اطفای حریق می‌توان خسارات ناشی از آتش‌سوزی را به حداقل رساند (۳۲). لازمه شناسایی مناطق در معرض خطر آتش‌سوزی، مشخص کردن عوامل مؤثر بر وقوع و گسترش آتش از قبیل مواد سوختی، شرایط توپوگرافی، شرایط آب و هوایی و عامل انسانی می‌باشد (۵۶). لذا باید مشخص شود چه رابطه‌ای بین عوامل و میزان وقوع و انتشار آتش وجود دارد. برای این منظور نیاز است تا مناطق آتش‌سوزی شده در دوره‌های گذشته، ثبت و این مناطق را با لایه‌های مربوط به عوامل اثرگذار روی آتش تطبیق و در نهایت تحلیل‌های فراوانی نقاط و پهنه‌های آتش‌سوزی را در هر یک از عوامل و شرایط مختلف آن عامل به‌منظور مدل‌سازی انجام شود (۳۱). به‌طور کلی مدل‌سازی آماری به دو روش پارامتریک و ناپارامتریک انجام می‌شود. روش‌های پارامتریک برای مدل‌سازی روابط بین متغیرها دارای تعدادی پیش‌فرض و محدودیت می‌باشند. در نظر گرفتن یک توزیع پیش‌فرض مانند توزیع نرمال برای متغیرهای پاسخ، خطی بودن رابطه‌ی پیشنهادی، یکسان بودن واریانس خطاها و مستقل بودن متغیرها از جمله محدودیت‌های روش‌های پارامتریک هستند که هنگام استفاده‌ی عملی از این روش‌ها، اگر داده‌های واقعی، شرایط مفروض مدل را نداشته باشند استفاده از این روش‌ها امکان‌پذیر نبوده و یا با خطای قابل توجه همراه است. به‌علاوه، هیچ‌یک از این روش‌ها قابلیت مدل‌سازی روابط پیچیده‌ی غیرخطی و اثر متقابل درجه بالا را ندارند. محدود بودن تعداد داده‌های ورودی، حساس بودن بیشتر این مدل‌ها به داده‌های گم‌شده و داده‌های پرت از دیگر محدودیت‌های این روش‌ها به‌شمار می‌آید (۴۶). از این روش‌های مدل‌سازی پارامتریک نظیر رگرسیون لجستیک^۱ و روش‌های ترکیب خطی و رتبه‌ای^۲ مورد استفاده در تحقیقات گذشته (۵۶، ۵)، به‌دلیل پیچیدگی روابط بین وقوع آتش‌سوزی و عوامل مؤثر بر آن دقت و کارایی لازم را ندارند. بنابراین روش‌های دقیق و نوین ناپارامتریک جهت پیش‌بینی و مدل‌سازی مکانی خطر آتش‌سوزی از جمله شبکه‌عصبی مصنوعی^۳ و روش‌های داده‌کاوی با الگوریتم یادگیری ماشین مانند ماشین پشتیبان بردار^۴ و درخت تصمیم مانند جنگل تصادفی^۵ قادرند ارتباطات پیچیده بین عوامل مؤثر بر وقوع آتش‌سوزی را بهتر مدل‌سازی

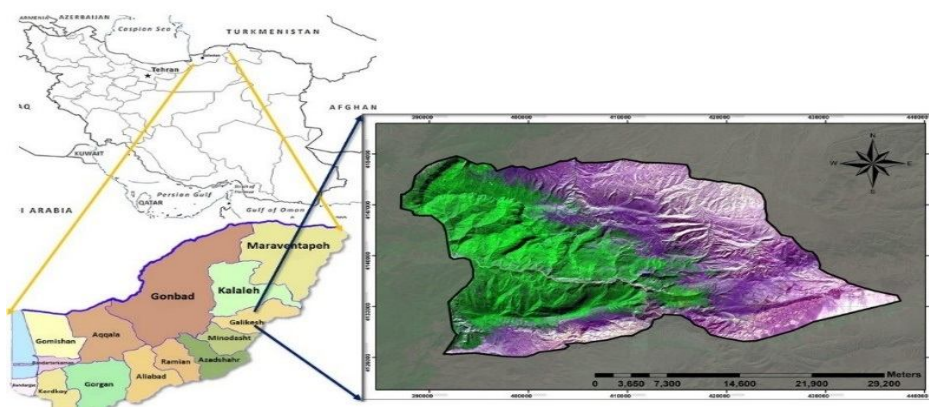
-
- 1- Logistic regression
 - 2- Linear combination
 - 3- Artificial Neural Network
 - 4- Support Vector Machine
 - 5- Random Forest

نمایند. هلدن و همکاران (۲۰۰۹)، با استفاده از الگوریتم جنگل تصادفی به پیش‌بینی خطر وقوع آتش‌سوزی در جنگل‌های ملی جیلا در جنوب غربی آمریکا پرداختند. آن‌ها از داده‌های توپوگرافی و میزان رطوبت نسبی به‌عنوان متغیر مستقل و از ۱۱۴ نقطه آتش‌سوزی به‌عنوان متغیر وابسته استفاده کردند. نتایج ارزیابی صحت نشان داد که الگوریتم جنگل تصادفی با صحت کلی ۷۹/۵ درصد توانست خطر وقوع آتش‌سوزی را پیش‌بینی کند (۲۸). الیورا و همکاران (۲۰۱۲)، به مدل‌سازی مکانی وقوع آتش‌سوزی با استفاده از الگوریتم‌های رگرسیون لجستیک و جنگل تصادفی پرداختند. پس از انجام مدل‌سازی نقشه احتمال وقوع با دقت ۱۰ متر تهیه شد. نتایج نشان داد الگوریتم جنگل تصادفی توانایی بالاتری در پیش‌بینی وقوع آتش‌سوزی دارد (۴۲). کرتز و مورایس (۲۰۰۷)، با استفاده از تکنیک‌های داده کاوی شامل ماشین‌بردار پشتیبان، جنگل تصادفی، شبکه عصبی پرسپترون تک لایه^۱ و رگرسیون چند متغیره به پیش‌بینی خطر آتش‌سوزی جنگل در منطقه‌ای در شمال شرقی پرتغال پرداختند. آن‌ها دریافتند ماشین‌بردار پشتیبان با چهار داده‌ی اقلیمی ورودی به مدل شامل بارش، دما، رطوبت نسبی و باد با صحت بالاتری خطر آتش‌سوزی را پیش‌بینی می‌کند (۱۲). سکار و همکاران (۲۰۱۱)، با استفاده از الگوریتم ماشین‌بردار پشتیبان و شبکه عصبی مصنوعی به پیش‌بینی خطر آتش‌سوزی در لبنان پرداختند. نتایج نشان داد ماشین‌بردار پشتیبان با دقت بالاتری قادر به پیش‌بینی خطر وقوع آتش‌سوزی با دو پارامتر ورودی اقلیمی می‌باشد (۴۹). واسکانسل و همکاران (۲۰۰۱)، با استفاده از رگرسیون لجستیک و شبکه عصبی پرسپترون چند لایه با الگوریتم یادگیری پس انتشار خطا اقدام به پیش‌بینی احتمال خطر وقوع آتش‌سوزی کردند و سپس نقشه حاصله در ۵ طبقه طبقه‌بندی شد. نتایج نشان داد شبکه عصبی مصنوعی با صحت کلی ۷۵ درصد توانایی بالاتری در پیش‌بینی احتمال خطر وقوع آتش‌سوزی در منطقه مورد مطالعه دارد (۵۷). آلونسو و همکاران (۲۰۰۳)، خطر آتش‌سوزی جنگل را در منطقه‌ای در شمال غرب اسپانیا با استفاده از شبکه عصبی مصنوعی پیش‌بینی کردند. نقشه خطر در چهار طبقه کم خطر، متوسط خطر، پرخطر و خطرناک طبقه‌بندی شد. نتایج ارزیابی صحت نشان داد مدل موردنظر با صحت کلی ۹۳ درصد احتمال خطر وقوع آتش‌سوزی را پیش‌بینی می‌کند (۳). سافی و برومی (۲۰۱۳)، به پیش‌بینی آتش‌سوزی جنگل با استفاده از شبکه عصبی پرسپترون چند لایه در یک پارک طبیعت واقع در پرتغال پرداختند. آن‌ها با استفاده از داده‌های اقلیمی منطقه و نقاط آتش‌سوزی به

پیش‌بینی خطر آتش‌سوزی پرداختند. آن‌ها دریافتند شبکه‌ای با دو لایه مخفی شامل ۶ و ۱۲ نرون در هر لایه و استفاده از الگوریتم تابع یادگیری پس انتشار کم‌ترین خطا را در منطقه مورد مطالعه دارد (۴۷). از آنجایی که در کشور مطالعه جامعی در زمینه استفاده از روش‌های دقیق و نوین ناپارامتریک جهت پیش‌بینی و مدل‌سازی مکانی خطر آتش‌سوزی صورت نگرفته است این مطالعه با هدف بررسی احتمال خطر وقوع آتش‌سوزی با استفاده از سه الگوریتم رایج ناپارامتریک ماشین بردار پشتیبان، جنگل تصادفی و شبکه عصبی مصنوعی از نوع پرسپترون چند لایه، در پارک ملی گلستان انجام شد و اقدام به مدل‌سازی و تهیه نقشه احتمال خطر وقوع آتش‌سوزی با استفاده از الگوریتم‌های فوق گردید و در نهایت ارزیابی صحت نقشه‌های حاصل از مدل‌سازی به‌منظور معرفی بهترین الگوریتم جهت تهیه نقشه احتمال خطر وقوع آتش‌سوزی در منطقه مورد مطالعه انجام شد.

مواد و روش‌ها

منطقه مورد مطالعه: پارک ملی گلستان منطقه‌ای کوهستانی است که در شرق جنگل‌های شمال کشور واقع شده است. این پارک از نظر موقعیت جغرافیایی در حد فاصل ۳۷ درجه ۱۶ دقیقه ۳۴ ثانیه تا ۳۷ درجه و ۳۷ دقیقه ۰۰ ثانیه عرض شمالی و ۵۵ درجه و ۴۳ دقیقه ۰۰ ثانیه تا ۵۶ درجه و ۱۷ دقیقه ۵۵ ثانیه طول شرقی بین شهرستان‌های گنبد کابوس و بجنورد قرار گرفته است (شکل ۱). مساحت پارک حدود ۹۱۸۹۵ هکتار و محیط آن ۱۴۷ کیلومتر است. پارک ملی گلستان یک منطقه کوهستانی است که در شرقی‌ترین امتداد رشته کوه البرز قرار گرفته و سیمای آن از کوه‌های خشک صخره‌ای، دره‌ها، تپه ماهورها، جنگل‌های کوهستانی، کوه‌های استپی تا دشت هموار و خشک محدودی در شرق منطقه متغیر است. دامنه تغییرات ارتفاعی در این منطقه بین ۴۵۰ متر و ۲۴۱۱ متر از سطح دریا متغیر است. دمای متوسط سالانه پارک بین ۱۱/۵ و ۱۷/۵ درجه سانتی‌گراد متغیر است. دامنه نوسانات بارندگی در پارک ملی گلستان تغییراتی از ۱۵۰ میلی‌متر تا ۷۵۰ میلی‌متر را نشان می‌دهد (۳۶).

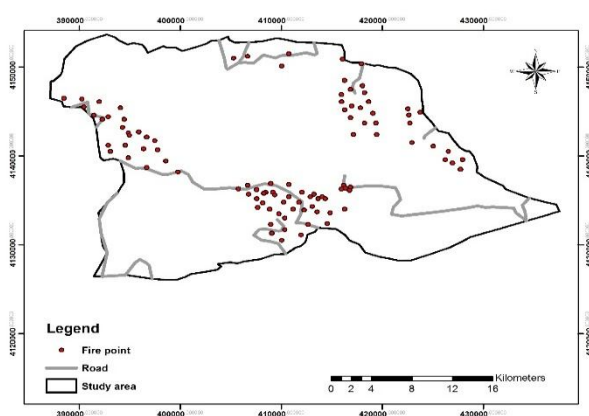


شکل ۱- موقعیت منطقه مورد مطالعه.

Figure 1. The position of the Golestan National Park in the study area.

روش تحقیق

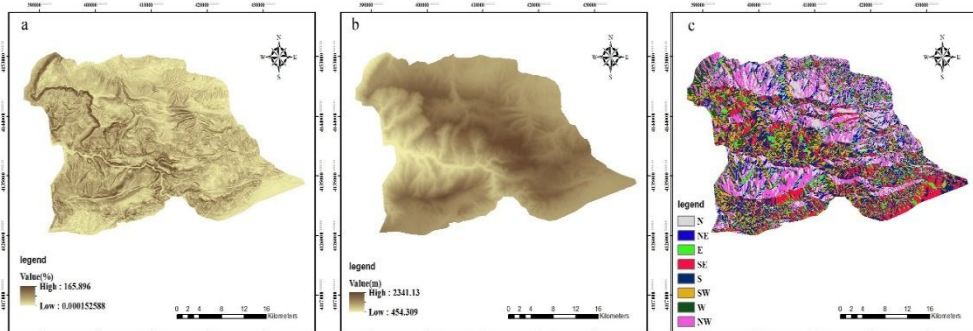
تهیه نقشه آتش‌سوزی‌های بوقوع پیوسته سنوات گذشته: ابتدا مختصات نقاط آتش‌سوزی از اداره کل محیط‌زیست استان گلستان که بین سال‌های ۱۳۸۲ تا ۱۳۹۲ به وقوع پیوسته بود، دریافت شد. سپس نقشه واقعیت زمینی این نقاط تهیه گردید (شکل ۲). اکثر آتش‌سوزی‌ها در پارک ملی گلستان از ماه فروردین شروع شده و تا اواخر آذرماه ادامه دارد.



شکل ۲- نقاط آتش‌سوزی و جاده‌های موجود در پارک ملی گلستان.

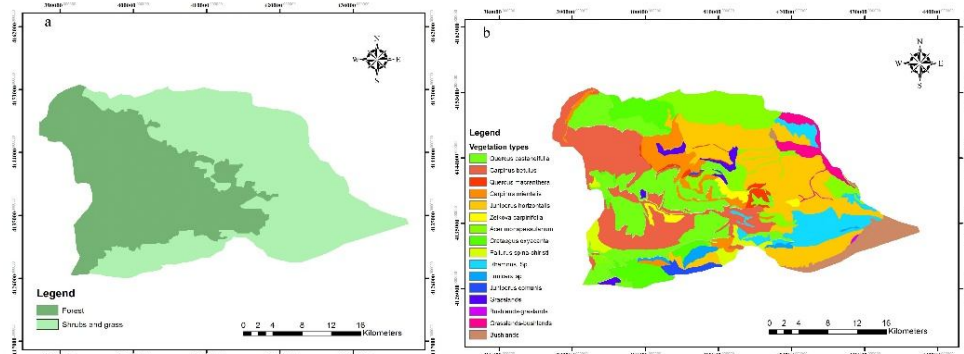
Figure 2. Fire points and roads in the Golestan National Park.

شناسایی عوامل مکانی مؤثر بر وقوع آتش سوزی: برای تهیه نقشه پهنه‌بندی خطر آتش سوزی در ابتدا عوامل مؤثر بر وقوع آتش سوزی در منطقه شناسایی شدند. این عوامل عبارتند از: مشخصه‌های فیزیوگرافی شامل ارتفاع از سطح دریا، شیب و جهت شیب (شکل ۳)؛ مشخصه‌های پوشش گیاهی شامل تیپ‌بندی و کاربری اراضی (شکل ۴)؛ مشخصه‌های اقلیمی شامل متوسط حداکثر دما، بارش متوسط ماهانه، رطوبت نسبی، متوسط تبخیر ماهانه (شکل ۵)؛ مشخصه‌های انسانی شامل فاصله از جاده، فاصله از اراضی زراعی و مشخصه فاصله از رودخانه (شکل ۶) می‌باشد. پس از شناسایی این مشخصه‌ها برای هر یک نقشه مکانی در محیط GIS تهیه شد.



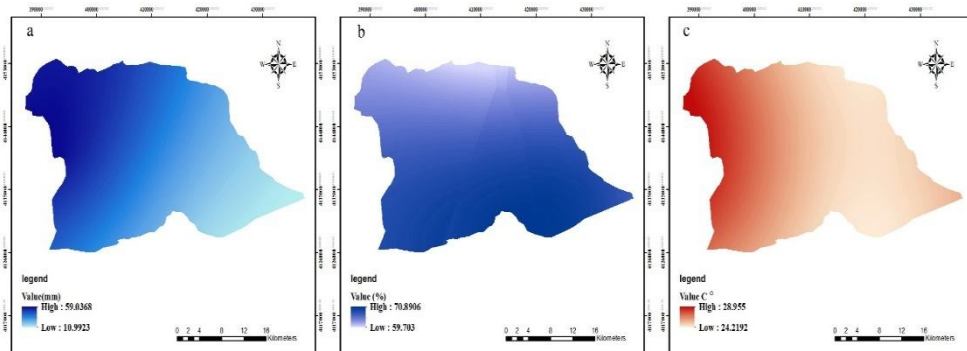
شکل ۳- نقشه عوامل فیزیوگرافی در پارک ملی گلستان (a) شیب (b) ارتفاع (c) جهت شیب.

Figure 3. Physiographic factors' map in the Golestan National Park. a) slope; b) elevation, c) aspect.



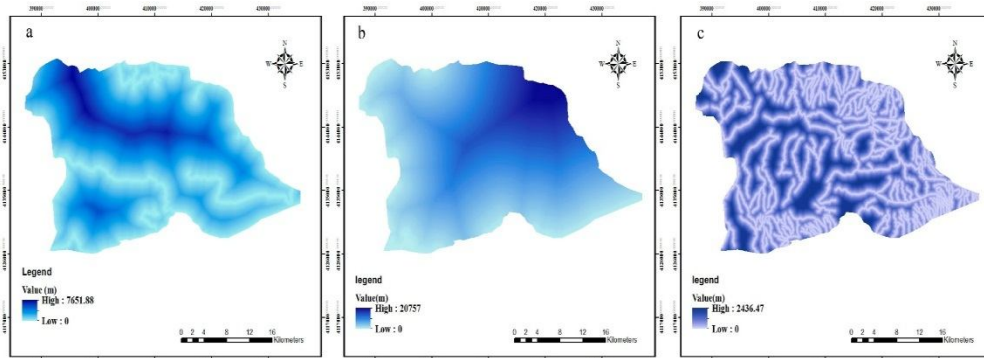
شکل ۴- (a) نقشه کاربری پوشش گیاهی (b) نقشه تیپ پوشش گیاهی.

Figure 4. a) Land use, b) vegetation type maps.



شکل ۵- نقشه عوامل اقلیمی در مهرماه (a) بارندگی (b) رطوبت نسبی (c) متوسط حداکثر دما.

Figure 5. Map of ecological factors in October; a) rainfall; b) the relative humidity; and c) the average maximum temperature.



شکل ۶- نقشه عوامل انسانی (a) فاصله از جاده‌ها (b) فاصله از اراضی زراعی (c) فاصله از رودخانه‌ها.

Figure 6. Map of human factors a) distance from roads, b) distance from arable lands, and c) distance from rivers.

تعیین میزان اثر نسبی هر یک از عوامل بر وقوع آتش‌سوزی: به منظور مدل‌سازی، منطقه به سطوح تصمیم‌گیری یک هکتاری تقسیم‌بندی شد (۱۶)، و ارزش رقومی لایه‌های عوامل تأثیرگذار بر وقوع آتش‌سوزی در محل سطوح یک هکتاری استخراج و سپس مقادیر نمونه‌گیری شده با استفاده از رابطه (۶) (۲۵)، استاندارد شدند و در بازه ۰ تا ۱ قرار گرفتند سپس جداول آماری جهت ورود به نرم‌افزار جهت مدل‌سازی با الگوریتم‌های ناپارامتریک تشکیل شد.

$$X_{norm} = \left(\frac{X - \bar{X}}{X_{min} - X_{max}} \right) \quad \text{رابطه ۶}$$

X_{norm} : داده نرمال \bar{X} : میانگین متغیرها X_{min} : حداقل مقدار متغیر مورد نظر X_{max} : حداکثر مقدار متغیر مورد نظر

مدل‌سازی و پیش‌بینی خطر احتمال وقوع آتش‌سوزی با استفاده از الگوریتم‌های ناپارامتریک: در این مطالعه ۱۰۰ نقطه از محل وقوع آتش‌سوزی جهت مدل‌سازی خطر احتمال وقوع آتش‌سوزی به‌کار برده شد. جهت یکسان‌سازی اثر انتخاب نمونه‌های تعلیمی در دقت مدل‌سازی (۵۲)، در هر سه الگوریتم از ۱۰۰ نمونه، ۷۰ درصد نمونه‌ها به‌صورت کاملاً تصادفی به‌عنوان نمونه‌های تعلیمی و ۳۰ درصد باقیمانده به‌عنوان داده‌های تست مورد استفاده قرار گرفت.

پیش‌بینی خطر احتمال وقوع آتش‌سوزی جنگل با الگوریتم شبکه عصبی مصنوعی: در این پژوهش از شبکه عصبی پرسپترون چند لایه با الگوریتم آموزش پس انتشار خطا (۴۷)، برای پیش‌بینی خطر احتمال وقوع آتش‌سوزی استفاده شد. از ۱۰۰ نقطه آتش‌سوزی، ۷۰ نقطه جهت آموزش شبکه و ۳۰ نقطه جهت بررسی کارایی مدل به‌کار برده شد. طی فرآیند آموزش شبکه عصبی مصنوعی با افزایش تعداد تکرارها، خطای مدل در پیش‌بینی داده‌های مرحله آموزش کاهش می‌یابد تا جایی که در تکرارهای بالا، شبکه به‌صورتی آموزش می‌بیند که فقط می‌تواند داده‌های مرحله آموزش را به خوبی تخمین زده و از پیش‌بینی داده‌های خارج از این محدوده عاجز می‌ماند به‌منظور حل این مشکل، درصدی از داده‌ها در ابتدای کار برای مرحله اعتبارسنجی در نظر گرفته می‌شود. با این کار در هر تکرار، میزان خطای مربوط به مرحله اعتبارسنجی نیز همزمان با مرحله آموزش محاسبه می‌شود. در مرحله آموزش ۱۰۰۰ بار فرآیند آموزش تکرار شد و شبکه بهینه با کم‌ترین خطا آزمون و بیشترین کارایی در بین سایر شبکه‌ها به‌عنوان بهترین ساختار شبکه برای مدل‌سازی خطر احتمال وقوع آتش‌سوزی انتخاب و مدل‌سازی انجام گردید.

پیش‌بینی خطر احتمال وقوع آتش‌سوزی جنگل با الگوریتم ماشین بردار پشتیبان: شکلی از ماشین بردار پشتیبان، که برای مسائل رگرسیونی مورد استفاده قرار می‌گیرد، ماشین بردار پشتیبان رگرسیون (SVR) نامیده می‌شود. تاکنون کرنل‌های مختلفی از جمله پایه‌های خطی، شعاعی، چند جمله‌ای و سیگموئید مورد استفاده قرار گرفته است. کرنل تابع پایه شعاعی^۱ (RBF) به‌عنوان بهترین انتخاب از

1- Radial base function

بین دیگر کرنل‌ها در پیش‌بینی احتمال خطر وقوع آتش‌سوزی گزارش شده است (۴۱، ۴۹). کرنل تابع پایه شعاعی که توسط γ در رابطه (۳) تعریف شده است محاسبه می‌شود.

$$\gamma > 0, k(x_i, x) = \exp(-\gamma|x_i - x|^2) \quad \text{رابطه (۳)}$$

برای مدل‌سازی و برآورد مشخصه‌های کمی توسط مدل SVR و کرنل تابع پایه شعاعی لازم است سه پارامتر γ, ϵ, C بهینه شود. برای بهینه‌سازی سه پارامتر مذکور از روش جستجوی شبکه استفاده می‌شود (۲۶).

برای پیدا کردن مقادیر پارامترهای ϵ, C از روش جستجو شبکه مشخص شده و اعتبارسنجی ۱۰ قسمتی با ۱۰۰۰ تکرار و γ ثابت، استفاده گردید (۲۶، ۵۲). پارامتر γ برابر ۰/۰۱۸ از طریق ۰/۱ تعداد متغیرهای مستقل محاسبه شد (۲۶). در روش جستجو شبکه مشخص شده برای C از مقادیر ۱ تا ۲۰ که برابر با دامنه متغیرهای ورودی می‌باشد (۳۹)، و برای مقادیر ϵ از ۰/۱ تا ۱ با فواصل ۰/۱ استفاده گردید. در این مطالعه هم‌چنین از کرنل پایه شعاعی به‌عنوان بهترین کرنل معرفی شده در اکثر تحقیقات (۱۲، ۵۲)، استفاده شد.

پیش‌بینی خطر احتمال وقوع آتش‌سوزی جنگل با الگوریتم جنگل تصادفی: برای بهینه‌سازی و اجرای الگوریتم جنگل تصادفی سه مولفه خیلی مهم هستند (۳۸): ۱. تعداد درختان رگرسیون (K tree)؛ ۲. تعداد متغیر برآورد کننده در هر گره (K predictor) و ۳. حداقل اندازه گره نهایی درختان که بیشتر از آن درخت تقسیم نمی‌شود. برای دستیابی به دقت کلی بیشتر از اعضا باید کل درختان دارای دقت فردی بهتری نسبت به حدس و تنوع بالایی باشند. مجموعه یادگیری برای هر درخت از طریق دسته‌بندی انجام می‌شود. هر درخت برآورد کننده $T_{L\theta}$ به یک وکتور یا خط تصادفی θ وابسته است که نشان می‌دهد نمونه از مجموعه آموزشی خود L دسته‌بندی شده‌اند. برآوردکننده نهایی f اکثریت رای یا میانگین برای همه درختان می‌باشد (رابطه (۱ و ۲)) (۵).

$$\hat{Y}_n = f(X_n) = \frac{1}{K} \sum_{k=1}^K T_{L(\theta_k)} \cdot (X_n)_1^K \quad \text{رابطه (۱)}$$

$$\hat{Y}_n = f(X_n) = \text{Majority vote}[T_{L(\theta_k)}]_1^K \quad \text{رابطه (۲)}$$

برای افزایش دقت مدل لازم است اریب کم (درختان تا حداکثر اندازه خود افزایش یابند) و همبستگی کم (برای دستیابی به همبستگی کم باید داده‌ها تصادفی شوند) باشد (۵۸).

به منظور افزایش عملکرد پیش‌بینی الگوریتم رندوم فارست نیاز به تنظیم بهینه پارامترهای ورودی است. بنابراین در این مطالعه ابتدا تعداد درختان بهینه تعیین شد. الگوریتم جنگل تصادفی با تعداد ۵۰۰ درخت اجرا گردید. سپس با تفسیر نمودار میانگین مربعات خطا بر حسب تعداد درخت داده‌های آموزش و آزمون حاصل شده، جایی که با افزایش تعداد درختان تغییری در کاهش میزان میانگین مربعات خطای آموزش و آزمون مشاهده نمی‌شد به‌عنوان تعداد درخت بهینه انتخاب شد که به‌عنوان ابزاری قدرتمند به منظور کشف و تأیید تعداد درختان بهینه در رندوم فارست است (۵۲)، سپس مدل با استفاده از تعداد درختان بهینه تعیین شده و سایر پارامترها ثابت اجرا شد. یکی دیگر از پارامترهای مهم در اجرای الگوریتم جنگل تصادفی تعداد برآورد کننده‌ها k (متغیرهای مستقل) در هر گره برای برآورد متغیر وابسته می‌باشد. یکی از ساده‌ترین راه‌های ممکن، استفاده از جذر کل تعداد متغیرهای مستقل مورد استفاده در مدل می‌باشد که این مقدار به‌عنوان برآورد کننده‌ها در هر گره مورد استفاده قرار می‌گیرد. بنابراین k برابر ۷ و حداقل گره در درختان تصمیم ۵ و حداکثر آن ۱۰۰ در نظر گرفته شد.

ارزیابی صحت نتایج طبقه‌بندی: هدف از بررسی صحت ارزیابی میزان همخوانی نقشه حاصل از طبقه‌بندی با واقعیت زمینی می‌باشد به این ترتیب، میزان درستی نتایج به‌دست آمده، ارزیابی و سطح اعتماد مربوط به آن‌ها مشخص می‌گردد. در این مطالعه به منظور ارزیابی صحت طبقه‌بندی، مقادیر پیش‌بینی شده خطر احتمال وقوع حاصل از داده‌های آزمون به دو طبقه ۰ (عدم وقوع آتش‌سوزی) و ۱ (وقوع آتش‌سوزی) تقسیم‌بندی شدند. مقادیر پیش‌بینی شده خطر احتمال وقوع آتش بین ۰ تا ۰/۵ به طبقه ۰ (عدم وقوع) و مقادیر بین ۰/۵ تا ۱ به طبقه ۱ (وقوع) نسبت داده شدند. سپس ارزیابی صحت طبقه‌بندی از طریق محاسبه معیار صحت کلی^۱ صورت گرفت.

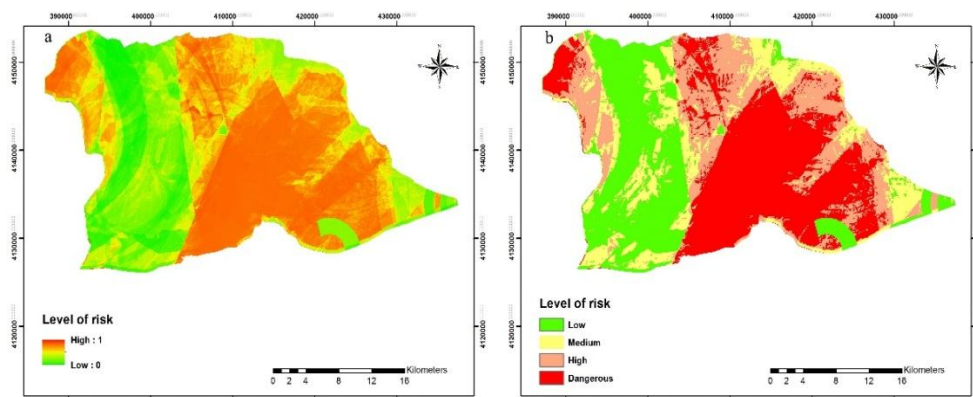
$$\text{رابطه (۷)} = \frac{\text{مجموع پیکسل‌های صحیح طبقه بندی شده}}{\text{مجموع پیکسل‌های مورد طبقه بندی}} = \text{صحت کلی}$$

در این معیار فقط از پیکسل‌های صحیح طبقه‌بندی شده استفاده شده است و از شرکت دادن پیکسل‌هایی که به غلط به طبقات تعلق گرفتند صرف‌نظر می‌شود.

1- Overall accuracy

نتایج و بحث

الگوریتم شبکه عصبی مصنوعی پرسپترون چند لایه: در مرحله آموزش ۱۰۰۰ بار فرآیند آموزش تکرار شد و نهایتاً شبکه با ۵۵ نرون در لایه ورودی، ۸ لایه مخفی و یک لایه خروجی با تابع محرک لجستیک سیگموئیدی در لایه مخفی و تابع محرک نمایی در لایه خروجی با کم‌ترین خطا آزمون و بیشترین کارایی در بین سایر شبکه‌ها به‌عنوان بهترین ساختار شبکه جهت مدل‌سازی انتخاب شد. سپس پیش‌بینی احتمال نسبی وقوع آتش‌سوزی برای سایر نقاط مجهول یک هکتاری با استفاده از شبکه منتخب انجام گرفت. ابتدا نقشه پیوسته احتمال خطر وقوع آتش‌سوزی ایجاد شد و سپس نقشه پیوسته ایجاد شده به چهار کلاس کم خطر، متوسط خطر، پرخطر و خطرناک طبقه‌بندی شد (شکل ۷).



شکل ۷- (a) نقشه پیوسته احتمال خطر وقوع آتش

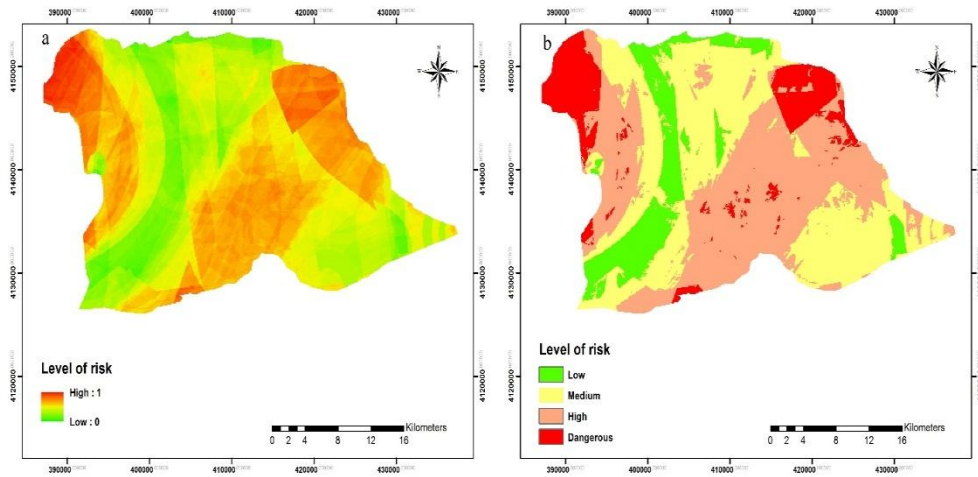
تهیه شده با الگوریتم شبکه عصبی مصنوعی

(b) نقشه طبقات احتمال خطر وقوع آتش

تهیه شده با الگوریتم شبکه عصبی مصنوعی

Figure 7. a) Fire risk assessment map; b) Fire risk classes map prepared by ANN algorithm.

الگوریتم ماشین بردار پشتیبان: الگوریتم ماشین بردار پشتیبان یکی از روش‌های یادگیری ماشین است که نتایج آن کاملاً وابسته به تنظیم صحیح پارامترهای ورودی مدل می‌باشد. بهترین پارامترها با استفاده از روش اعتبارسنجی متقابل و جستجوی خودکار برای پیدا کردن مقادیر پارامترهای C ، ϵ برای C مقدار ۱۰ و برای ϵ مقدار 0.3 و γ ثابت 0.18 با استفاده از کرنل پایه شعاعی جهت پیش‌بینی احتمال خطر وقوع آتش‌سوزی به‌دست آمد. پس از مدل‌سازی و مشخص شدن مقادیر نقاط مجهول ابتدا نقشه پیوسته احتمال خطر وقوع آتش‌سوزی ایجاد شد و سپس نقشه پیوسته ایجاد شده به چهار کلاس کم خطر، متوسط خطر، پرخطر و خطرناک طبقه‌بندی گردید (شکل ۸).



شکل ۸. (a) نقشه پیوسته احتمال خطر وقوع آتش سوزی

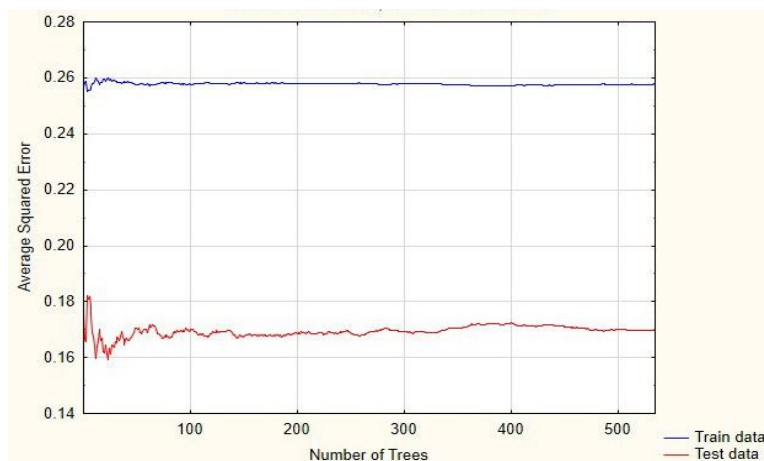
(b) نقشه طبقات احتمال خطر وقوع آتش سوزی

تهیه شده با الگوریتم ماشین بردار پشتیبان

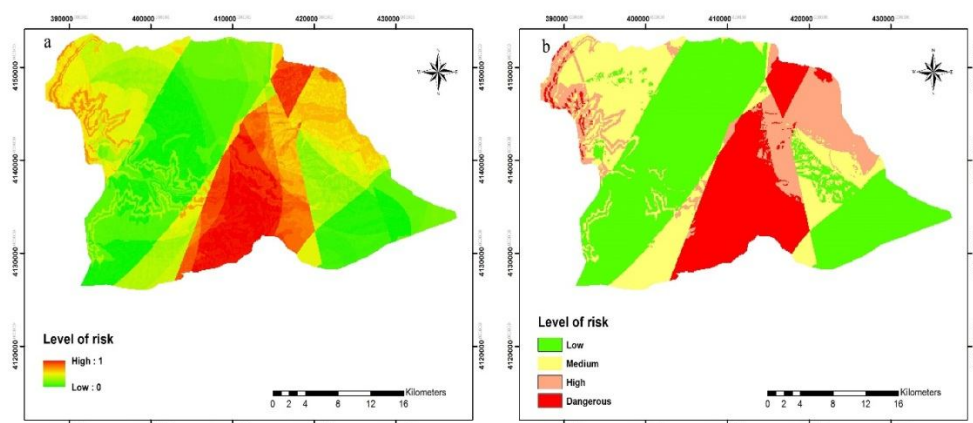
تهیه شده با الگوریتم ماشین بردار پشتیبان

Figure 8. a) Fire risk assessment map, b) Fire risk classes map prepared by SVM algorithm.

الگوریتم جنگل تصادفی: همان‌گونه که بیان شد به منظور اجرای بهینه مدل جنگل تصادفی تعیین صحیح تعداد درختان تصمیم و تعداد پیش‌بینی کننده‌ها در هر گره از ضروریات است (۵۲). شکل (۹) نمودار میانگین مربعات خطا بر حسب تعداد درخت به منظور پیش‌بینی خطرات احتمال وقوع آتش سوزی با استفاده از الگوریتم جنگل تصادفی می‌باشد. صحت مدل از حدود ۳۰ درصد به کندی افزایش می‌یابد و در ۲۰۰ درخت خطا ثابت و کم‌ترین میزان خود می‌رسد بنابراین تعداد درختان بهینه ۲۰۰ درخت و تعداد برآورد کننده‌ها (k) ۷ در نظر گرفته شد. حداقل گره در درختان تصمیم ۵ و حداکثر آن ۱۰۰ در نظر گرفته شد. پس از مدل‌سازی و مشخص شدن مقادیر نقاط مجهول ابتدا نقشه پیوسته احتمال خطر وقوع آتش سوزی ایجاد شد و سپس نقشه پیوسته ایجاد شده به چهار کلاس کم خطر، متوسط خطر، پرخطر و خطرناک طبقه‌بندی گردید (شکل ۱۰).



شکل ۹- نمودار میانگین مربعات خطا بر حسب تعداد درخت در داده‌های آموزش و آزمون احتمال خطر وقوع آتش‌سوزی.
Figure 9. Graph of average square error against number of trees in training and test data.

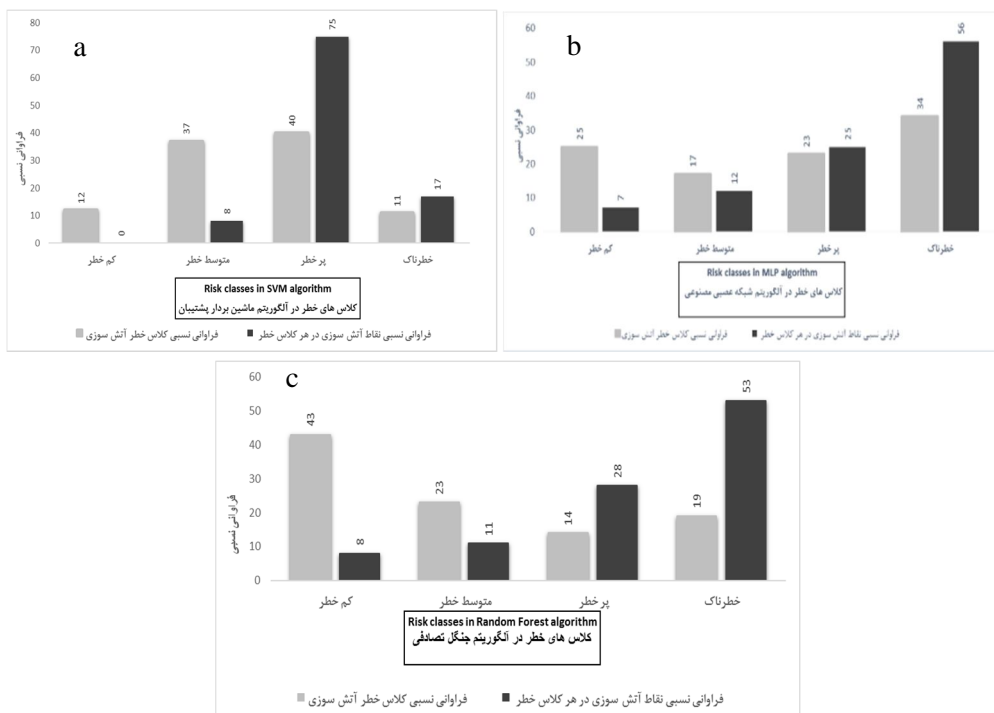


شکل ۱۰- (a) نقشه پیوسته احتمال خطر وقوع آتش‌سوزی تهیه شده با الگوریتم جنگل تصادفی
(b) نقشه طبقات احتمال خطر وقوع آتش‌سوزی تهیه شده با الگوریتم جنگل تصادفی

Figure 10. a) fire risk assessment map b) fire risk classes map prepared by RF algorithm

آتش‌سوزی‌های وقوع یافته در هر کلاس خطر با استفاده از الگوریتم‌های داده کاوی: با توجه به شکل‌های (۱۱) که درصد فراوانی نسبی کلاس‌های خطر در نقشه احتمال خطر وقوع آتش‌سوزی تهیه شده در کلاس‌های کم خطر، متوسط خطر، پرخطر و خطرناک با الگوریتم شبکه عصبی مصنوعی به ترتیب برابر با ۲۵ درصد، ۱۷ درصد، ۲۳ درصد و ۳۴ درصد، با الگوریتم ماشین بردار پشتیبان به

ترتیب برابر با ۱۲ درصد، ۳۷ درصد، ۴۰ درصد و ۱۱ درصد، با الگوریتم جنگل تصادفی به ترتیب برابر با ۴۳ درصد، ۲۳ درصد، ۱۴ درصد و ۱۹ درصد بوده است. درصد فراوانی نسبی نقاط آتش سوزی به وقوع پیوسته در سنوات قبل در هر یک از کلاس‌های کم خطر، متوسط خطر، پر خطر و خطرناک با الگوریتم شبکه عصبی برابر (۷ درصد، ۱۲ درصد، ۲۵ درصد، ۵۶ درصد) با الگوریتم ماشین بردار پشتیبان برابر (۰ درصد، ۸ درصد، ۷۵ درصد، ۱۷ درصد) و در الگوریتم جنگل تصادفی برابر (۸ درصد، ۱۱ درصد، ۲۸ درصد، ۵۳ درصد) بوده است. از نظر میزان مطابقت وقوع آتش با نتایج الگوریتم‌ها نشان می‌دهد که همه الگوریتم‌ها توانستند منطقه را از نظر احتمال وقوع آتش طبقه‌بندی نمایند به طوری که بیش از ۸۰ درصد نقاط آتش سوزی در مناطق با کلاس‌های خطر، پرخطر و خطرناک واقع شده است. نتایج حاصل شده با مطالعات (۱۷، ۵۵)، هم‌خوانی دارد.



شکل ۱۱- فراوانی کلاس‌های خطر و نقاط آتش سوزی در هر کلاس تهیه شده با الگوریتم‌های ناپارامتریک (a) SVM، (b) MLP، (c) RF.

Figure 11. Frequency of risk classes and fire points in any class prepared by non-parametric algorithms, a) SVM; b) ANN; and c) RF.

نتایج حاصل از الگوریتم‌ها نشان دادند که صحت کلی طبقه‌بندی الگوریتم شبکه عصبی برابر (۷۲/۵ درصد)، الگوریتم ماشین بردار پشتیبان (۶۸/۳ درصد) و با الگوریتم جنگل تصادفی (۷۵ درصد) بوده است. بنابراین هر سه الگوریتم با صحت کلی مناسبی توانستند احتمال خطر وقوع آتش‌سوزی را پیش‌بینی کنند. الگوریتم جنگل تصادفی با صحت کلی بالاتر نسبت به دو الگوریتم دیگر توانست احتمال خطر وقوع آتش‌سوزی را پیش‌بینی کند. نتایج صحت کلی طبقه‌بندی حاصل شده در این تحقیق در مقایسه با مطالعات (۴۳) (صحت کلی = ۶۵ درصد)، بالاتر و با (۵۷)، (صحت کلی = ۷۵ درصد)، (۲۸)، (صحت کلی = ۷۹ درصد)، تفاوت زیادی ندارد. اما در مقایسه با مطالعات (۳)، (صحت کلی = ۹۳ درصد)، (۵۴)، (صحت کلی = ۸۶ درصد)، پایین‌تر می‌باشد.

نتیجه‌گیری کلی

در این مطالعه به بررسی احتمال خطر وقوع آتش‌سوزی با استفاده از سه الگوریتم ناپارامتریک ماشین بردار پشتیبان، شبکه عصبی مصنوعی و جنگل تصادفی در پارک ملی گلستان پرداخته شد و اقدام به مدل‌سازی و تهیه نقشه احتمال خطر وقوع آتش‌سوزی با استفاده از الگوریتم‌های فوق‌گرددید و در نهایت الگوریتم جنگل تصادفی با صحت کلی ۷۵ درصد به‌عنوان مدل برتر معرفی شد. ویژگی‌های منطقه مورد مطالعه شامل توپوگرافی، عوامل مؤثر بر وقوع آتش‌سوزی، دقت و نوع لایه‌های متغیرهای مستقل جهت مدل‌سازی احتمال خطر وقوع آتش‌سوزی، دقت نقاط و محدوده‌های ثبت شده آتش‌سوزی‌های به وقوع پیوسته در گذشته و نوع الگوریتم پیش‌بینی مورد استفاده از جمله عوامل تأثیرگذار بر صحت کلی طبقه‌بندی در مطالعات می‌باشد. با توجه به این نکته که عامل وقوع اکثر آتش‌سوزی‌های رخ داده در پارک ملی گلستان انسانی بوده (۱، ۱۸)، می‌توان با افزایش متغیرهای انسانی تأثیرگذار مانند فاصله از روستاها، فاصله از محل تجمع توریست‌ها و فاصله از خطوط ریلی باعث افزایش صحت و دقت نقشه‌های خطر در این منطقه شد. با توجه به بالا بودن صحت کلی حاصل از بررسی احتمال وقوع آتش‌سوزی در این منطقه می‌توان از این مدل‌ها جهت پیش‌بینی و تهیه نقشه احتمال وقوع آتش‌سوزی در جهت مدیریت اصولی، جلوگیری از بروز و کاهش خسارات جبران‌ناپذیر آتش‌سوزی در سایر مناطق شمال کشور هم مورد در مقیاس وسیع استفاده کرد. بنابراین تهیه این نوع نقشه‌ها می‌تواند به مدیریت ریسک آتش‌سوزی در جنگل، پیشگیری از وقوع آتش‌سوزی و در نهایت اطفاء سریع آن کمک شایانی نمایند. براساس این نوع نقشه‌ها می‌توان امکانات موردنیاز را

برای مقابله با آتش‌سوزی در مناطق پر خطر و خطرناک از جمله احداث برج‌های دیده‌بانی، احداث مخازن آب در مناطق صعب‌العبور، پوشش منطقه با شبکه‌های حسگر بی‌سیم جهت کشف هر چه سریع‌تر آتش‌سوزی، ایجاد پایگاه‌های امداد و نجات قبل از شروع فصل آتش‌سوزی متمرکز نمود. هم‌چنین با آموزش روستاییان و ایجاد آتش‌برها و آتش‌بندها باعث پیشگیری و کاهش خسارات ناشی از آن شد.

منابع

1. Adab, H., Kasturi, D., and Karim, S. 2013. Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques, *Natural Hazards* 65(3): 1723–1743
2. Alexandridis, A., Vakalis, D., Siettos, C.I., and Bafas, G.V. 2008. A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through Spetses Island in 1990, *Applied Mathematics and Computation*, 204: 191–201.
3. Alonso-Betanzos, A., Fontenla-Romero, O., Guijarro-Berdin, B., Hernandez-Pereira, E., Canda, J., Jimenez, E., Legido, J., Muñoz, S., Paz-Andrade, C., and Paz-Andrade, M.I. 2002. A neural network approach for forestal fire risk estimation. In: van Harmelen F (ed) *The 15th European conference on artificial intelligence, ECAI'2002*, Lyon, France. IOS Press, 643–647.
4. Arpaci, A., Malowerschnig, B., Sass, O., and Vacik, H. 2014. Using multivariate data mining techniques for estimating fire susceptibility of Tyrolean forests, *Applied Geography* 53: 258-270.
5. Auret, L., and Aldrich, C. 2012. Interpretation of nonlinear relationships between process variables by use of random forests, *Minerals Engineering*. 35: 27-42.
6. Baier, D., and Wernecke, K.D. 2004. *Innovations in classification, data science and information systems*. Springer press, 613p.
7. BanjShafiei, A., Akbarinia, M., Jalali, S., and Azizi, P. 2006. The effects of fire on forest structure: Case study in Chelir, Kheyroudkenar (Watershed number 45 Golband, Nowshahr) Pajouhesh and Sazandegi. 76: 105-112. (In persian)
8. Breiman, L. 2002. *A Manual on Setting Up, Understanding and Using Random Forests v.3.1*. http://oz.berkeley.edu/users/breiman/using_random_forests.pdf Burke A, 2004, *Keepers of the Flame*. High Country News.
9. Cambardella, C.A., Moorman, T.B., Parkin, T.B., Karlen, D.L., Turco, R.F., and Konopka, A.E. 1994. Field scale variability of soil properties in Central Iowa soil, *Soil Science Society of America Journal*, 58: 1501-1511.
10. Campbell, C., and Ying, Y. 2011. *Learning with Support Vector Machines*,

- Morgan and Calypool Publishers series, 95p.
11. Chaei, A. 2000. Fire effects on vegetation changes in the Golestan National Park. MSc. Thesis, University of Mazandaran, 85p.
 12. Cortez, P.G., and Morais, A. 2007. A data mining approach to predict forest fires using meteorological data. In: Proceeding of the 13th Portuguese conference an artificial intelligence, 512-523.
 13. Dibike, Y.B., Velikov, S., Solomatin, D., and Abbot, M.B. 2001. Model induction with support vector machines-introduction and applications. *Journal of computing in Civil Engineering*, 15(3): 208-2016.
 14. Dietterich, T.G. 2000. Ensemble methods in machine learning In: Proceedings of the First International Workshop on Multiper Classifier Systems. Lecture Notes in Computer Science. MCS. Springer-Verlag, Pp: 1-15.
 15. Dayhoff, J.E. 1990. Neural network architectures, VNR, USA.
 16. Dlamini, W.M. 2006. Application of Bayesian networks for fire risk mapping using GIS and remote sensing data. *GeoJournal*, 76: 283–296.
 17. Eshaghi, M.A., and Shataee, Sh. 2014. Golestan National Park Fire Risk Zoning Using Tree-Based Random Forest Algorithm, International Forest Fire Conference in Black Sea Region, Kastamonu, Turkey.
 18. Eshaghi, M.A., and Shataee, Sh. 2014. Spatial distribution of occurred forest fires in the Golestan national park, northern east of Iran, International Forest Fire Conference in Black Sea Region, Kastamonu, Turkey.
 19. Fayad, U., Piatetsky-Shapiro, G., Symth, P., and Uthurusamy, R. 1996. *Advances in Knowledge Discovery and Data Mining*, MIT press 601p.
 20. Florinsky, I.V., and Kuryakova, G.A. 1966. Influence of topography on some vegetation cover properties Elsevier Science, 27: 123-141.
 21. Ghadami, H. 2011. Khorasan newspaper, viewed 15 April 2012 .<http://www.khorasannews.com/News.aspx?type=1&year=1390&month=2&day=1&id=2219745>.
 22. Ganawa, E.S.M., and Mohammad Sharif, A.R. 2003. Spatial variability of total nitrogen and available phosphorus of large rice field in Sawah Sepadan Malaysia, *Science Asia Journal*, 29: 7-12.
 23. Hagan, M.T. 1995. Neural network design, PWS, USA.
 24. Hagan, M.T., Menhaj, M. 1999. Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*. 5(6): 989-993.
 25. Hastie, T., Tibshirani, R., and Friedman, J. 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, NY, USA.
 26. Hsu, C.W., Chang, C.C., and Lin, C.J. 2003. A practical guide to support vector classification, www.csie.ntu.edu.tw, 4 May 2009.
 27. Hsu, C.W., Chang, C.C., and Lin, C.J. 2010. *A Practical Guide to Support Vector Classification* (Taipei: Department of Computer Science, National Taiwan University), Available online at: <http://www.csie.ntu.edu.tw/~cjlin>.

28. Holden, Z., Morgan, P., and Evans, J. 2009. A predictive model of burn severity based on 20-year satellite-inferred burn severity data in a large southwestern US wilderness area, *For Ecol Manag.* 258(11): 2399–2406.
29. Jain, A.K., Mao, J., and Mohiuddin, K.M. 1996. Artificial Neural Networks: A Tutorial”, *Computer*. Pp: 31-44.
30. Jianhua, Z., Jiancai, G., Kejian, W. 2013. Comparative Analysis of the Frequency of Forest Fires in Inner Mongolia and the Greater Khingan Mountains, *AISS*. 5(10): 1234-1241.
31. James, K., Brown, J.K.S. 2000. Wild land Fire in Ecosystems Effects of Fire on Flora, *USDA Forest Service Gen*, vol2. 266.13.
32. Jaiswal, R.K., Mukherjee, S., and Raju, D.K. 2002. Forest fire risk zone mapping from satellite imagery and GIS, *International Journal of Applied Earth Observation and Geoinformation*, 4: 1-10.
33. Kandy, A.K., Kimothi, M.M., Jadhav, R.N., Agarwal, J.P. 1998. Application of GIS in identification of fire prone areas, a feasibility study in parts of Junagarth, (Gujrat, India). *The indian Forester*, 124(7): 531-535.
34. Lindsay, J.B. 2005. The terrain analysis system: a tool for hydro-geomorphic applications, *Journal Hydrological Processes*, 19: 1123-1130
35. Makela, H., and Pekkarinen, A. 2004. Estimation of forest stands volumes by Landsat TM imagery and stand-level field-inventory data, *Forest Ecology and Management*, 196: 245–255
36. Madjnoonian, H., Zehzad, B., and Kiabi, B. 1999. Golestan National Park (Biosphere Reserve) – Department of the Environment, 130p.
37. Maeda, E.E., Arcoverde, G.F.B., Pellikka, P.K.E., Shimabukuro, E.Y. 2011. Fire risk assessment in the Brazilian Amazon using MODIS imagery and change vector analysis, *Applied Geography*, 31: 76-84
38. Mutanga, O., Adama, E., and Choa, M.A. 2012. High density biomass estimation for wetland vegetation using World View-2 imagery and random forest regression algorithm, *International Journal of Applies Earth Observation and Geoinformation*, 18: 399-406.
39. Mattera, D., and Haykin, S. 1999. Support vector machines for dynamic reconstruction of a chaotic system, In *Advances in Kernel Methods: Support Vector Learning* Schölkopf B, Burges EY, Smola. Cambridge, MA: MIT Press AJ. Pp: 211–242
40. Nyatondo, U.N. 2010. Fire spread modeling in Majella National Park, Italy. M.Sc., Thesis. International Institute for Geo Information Science and Earth Observation, Enscheda, the Netherlands.
41. Noori, R., Abdoli, M., Farokhnia, A., and Abbasi, M. 2009. Results uncertainty of solid waste generation forecasting by hybrid of wavelet transform-ANFIS and wavelet transform-neural network. *Expert Systems with Applications* 36(6): 9991-9999

42. Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., and Pereira, J.M.C. 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecology and Management*, 275: 117–129
43. Ozbayoglu, A., and Recep, B. 2012. Estimation of the burned area in forest fires using computational intelligence techniques, Conference Organized by Missouri University of Science and Technology, 2012- Washington D.C.
44. Peng, C., and Wen, X. 1999. Recent Applications of Artificial Neural Networks in Forest Resource Management: An Overview. In: *Environmental Decision Support Systems and Artificial Intelligence*, Ulises Corte and Miquel Srnchez-Marc, Cochairs (eds.). Pp: 12-22. Tech. Rep. WS-99-07, AAAI press, Menlo Park, California.
45. Rothermel, R.C. 1983. How to predict the spread and intensity of forest and Range fires. Gen. Tech.Rep.INT-143.USDA Forest Service, Intermountain Forest and Range Experiment Station.
46. Samui, P. 2008. Support vector machine applied to settlement of shallow foundations on cohesion less soil, *Journal Computers and Geotechnics*, 35: 419-427.
47. Safi, Y., and Bouroumi, A. 2013. Prediction of Forest Fire Using Artificial Neural Networks. *Mathematical Sciences*. 7(6): 271–286
48. Salmanmahiny, A. 2011. Necessity and Possibility of Predicting Forest Areas Prone to Fire and Modeling Fire Spread Pattern, international conference on wildfires in natural resources lands, Gorgan, iran
49. Sakr, G., Elhadj, E.I.H., and Mitri, G. 2011. Efficient forest fire occurrence prediction for developing countries using two weather parameters, *Engineering Applications of Artificial Intelligence* 24: 888–894
50. Salis, M. 2008. Fire behavior simulation in Mediterranean Maquis using FARSITE. PhD Doctoral Thesis, Universita' Degli Studi Di Sassari.
51. Shafiei, A. 2006. Fire effects on vegetation changes in Chelier's Forest, Iran, PhD Thesis, Tarbiat Modares University
52. Shataee, Sh., Kalbi, S., Falah, A., and Pelz, D.R. 2012. Forest attribute imputation using machine –learning methods and ASTER data: comparison of K-NN, SVR and random forest regression algorithms, *International Journal of Remote Sensing*. 33(19): 6254-6280.
53. Sibanda, C. 2011. Modeling forest fire behavior and carbon emission in the Ludikhola Watershed, Gorkha District, Nepal. M.Sc. Thesis. University of Twente.
54. Stojanova, D.G., Pano, P.G., Kobler, A.G., Dzeroski, S.G., and Taskova, K. 2006. Learning to predict forest fires with different data mining techniques, In: *Conference on Data Mining and Data Warehouses*, Pp: 255-258.
55. Tian, X., Zhao, F., Shu, I., and Wang, M. 2013. Distribution characteristics and the influence factors of forest fires in China, *Forest Ecology and Management*, 310: 460-467.

56. Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., and Matsinos, I. 2009. Identifying wild land fire ignition factors through sensitivity analysis of a neural network, *Nat Hazards*, 50(1): 125–143.
57. Vasconcelos, M.J.P., Silva, S., Tomé, M., Alvim, M., Cardoso Pereira, J.M. 2001. Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks, *Photogrammetric Engineering and Remote Sensing*, 67: 73-81.
58. Verikas, A., Gelzinis, A., and Bacauskiene, M. 2011. Mining data with random forest: a survey and result of new tests, *Pattern Recognition*, 44: 330-349.
59. Wang, Z., and Brenner, A. 2009. An integrated method for forest canopy cover mapping using Landsat ETM+ imagery, In *ASPRS/MAPPS Fall Conference 16–19 November 2009, San Antonio, TX*
60. Webster, R., and Oliver, M.A. 2000. *Geostatistics for environmental scientists*. Wiley press. 271p.
61. Wilson, J.P., and Gallant, J.C. 2000. Digital terrain analysis, In Wilson J.P. and Gallant JC (eds) *Terrain Analysis: Principles and Applications*, New York John Wily and Sons, Pp: 1-27.
62. Yin, H.W., Kong, F.H., and X.Z. 2004. RS and GIS-based forest fire risk zone mapping in Dahinggan Mountains. *Chinese Geographical Science*, 14(3): 251-257.
63. Zahedi Amiri, G.h. 1998. Relation between ground vegetation and soil characteristic in a mixed hardwood stand, Ph.D. thesis, academic press, University of Gent, Belgium. 319p.
64. Zhang, Q.F., and Chen, W.J. 2007. Fire cycle of the Canada's boreal region and its potential response to global change, *Journal of Forestry Research*, 18(1): 55-61.



Gorgan University of Agricultural
Sciences and Natural Resources

J. of Wood & Forest Science and Technology, Vol. 23 (4), 2016
<http://jwfst.gau.ac.ir>

Preparation map of Forest Fire Risk Using SVM, RF and MLP Algorithms (Case Study: Golestan National Park, Northeastern Iran)

***M.A. Eshaghi¹ and Sh. Shataee Jouibary²**

¹M.Sc. Graduated, Dept., of Forestry, Gorgan University of Agricultural Sciences and Natural Resources, ²Associate Prof., Dept., of Forestry, Faculty of Forest Science, Gorgan University of Agricultural Sciences and Natural Resources

Received: 06/08/2015 ; Accepted: 12/22/2015

Abstract

Background and objectives: Spatial prediction of fire risk and preparing the forest fire risk map across the natural areas are among the ways that can be used to prevent and to manage fire. The aim of this research was zonation of forest fire risk in Golestan National Park using non-parametric algorithms, namely Artificial Neural Network (ANN), Support Vector Machine (SVM) and Random Forest (RF).

Materials and methods: About 100 occurred fire points were considered for modeling the fire risk. The effective factors on fire occurring including vegetation types, physiographic, climatic, and human factors were identified and their relevant maps were prepared from different sources. To modeling purposes, initially the zone was divided into 1-ha levels of decision-making and modeling and then the pixel values of the effective factors on classes of fire occurring, across the 1-ha levels, were extracted and standardized. Based on non-parametric algorithms, fire risk was modeled with 70 percent of the fire points, as training samples. The prepared forest fire risk map was zoned in terms of four classes of low-risk, medium-risk, high-risk and dangerous. The classification accuracy of the maps, resulted from this modeling, was assessed through the overall classification accuracy given 30 percent of the remained fire points.

Results and Conclusion: The results indicated that RF algorithm, with the overall accuracy of 75%, was the best algorithm in predicting the fire risk compare to other ones. Likewise, after matching the fire risk occurring with the results gained from algorithms, it turned out that all algorithms were able to classify the area properly in terms of the fire risk, as more than 80 percent of fire points were placed in the high-risk and dangerous classes.

Keywords: Forest fire risk map, Support Vector Machine, Artificial Neural Network, Random Forest, Golestan National Park.

*Corresponding author: amin.eshaghi@gmail.com