Incidents and Disasters Resilience

Homepage: https://disaster.ndri.ac.ir/

Forest fire susceptibility mapping using a hybrid ensemble machine learning model: a performance evaluation of sentinel-2 and landsat 8 data

Mohaddeseh Mesvari[1], Reza Shah Hosseini[2]✉

1. Ph.D. Candidate, Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics Engineering, University of Tehran, Tehran, Iran. Email: m.mesvari@ut.ac.ir
2. Corresponding Author, Associate Professor, Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics Engineering, University of Tehran, Tehran, Iran. Email: rshahosseini@ut.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 26 June 2025
Received in revised form 18 September 2025
Accepted 1 November 2025
Available online 30 December 2025

Keywords:
Forest Fire,
Machine Learning, Hybrid Ensemble Model,
Sentinel-2 & Landsat 8,
Remote Sensing.

ABSTRACT

Objective: Forest fires are among the most destructive natural hazards, causing irreversible ecological and economic consequences. Accurate fire risk mapping is therefore a critical tool for effective prevention and management. This study aimed to develop an integrated framework based on machine learning algorithms and to evaluate the effect of satellite data spatial resolution on the accuracy of forest fire susceptibility modeling.

Method: To achieve this objective, spectral and thermal data derived from Landsat 8 (30 m spatial resolution) and Sentinel-2 (10 m spatial resolution) sensors were combined with topographic and climatic variables. Five well-established machine learning models were initially trained, and their outputs were subsequently integrated through a hybrid ensemble model to predict fire-prone areas. This approach enabled the exploitation of the complementary strengths of individual machine learning algorithms, improved the modeling of complex nonlinear relationships within the input data, and enhanced overall model stability.

Results: Exploratory analyses revealed that the Normalized Difference Vegetation Index (NDVI), the Normalized Burn Ratio (NBR), elevation, and land surface temperature were the most influential variables in distinguishing fire-affected areas. Quantitative evaluation demonstrated that the hybrid ensemble model implemented using Sentinel-2 data outperformed all individual base models, achieving an overall accuracy of 94.52%, an F1-score of 87.76%, and an Intersection over Union (IoU) value of 80.12%. These findings indicate that integrating the capabilities of multiple machine learning algorithms can substantially improve predictive accuracy and model robustness compared with relying on a single model. Furthermore, a comparison of the two satellite sensors showed that the higher spatial resolution of Sentinel-2 significantly enhanced fire susceptibility estimation and produced maps with superior spatial detail and boundary delineation by reducing spatial averaging effects.

Conclusions: The results demonstrate that integrating hybrid ensemble models with high-resolution satellite imagery can provide an efficient and reliable framework for proactive natural resource management and forest fire risk mitigation.

Cite this article: Mesvari. M., Shah Hosseini, R. (2025). Forest fire susceptibility mapping using a hybrid ensemble machine learning model: a performance evaluation of sentinel-2 and landsat 8 data. ***Incidents and Disasters Resilience***, *1*(4), 75-94.

© Author(s) retain the copyright.

Publisher: Natural Disasters Research Institute (NDRI).

تاب‌آوری در برابر حوادث و سوانح

Homepage: https://disaster.ndri.ac.ir

نقشه‌برداری پتانسیل خطر آتش‌سوزی جنگل با استفاده از مدل یادگیری ماشین فراترکیبی: ارزیابی عملکردی داده‌های سنتینل ۲ و لندست ۸

محدثه مسواری[۱]، رضا شاه حسینی[۲]✉

۱. دانشجوی دکتری، گروه فتوگرامتری و سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران. رایانامه: m.mesvari@ut.ac.ir

۲. نویسنده مسئول، دانشیار، گروه فتوگرامتری و سنجش از دور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران. رایانامه: rshahosseini@ut.ac.ir

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

تاریخچه مقاله:
تاریخ دریافت: ۱۴۰۴/۰۴/۰۵
تاریخ بازنگری: ۱۴۰۴/۰۶/۲۷
تاریخ پذیرش: ۱۴۰۴/۰۸/۱۰
تاریخ انتشار: ۱۴۰۴/۱۰/۰۹

کلیدواژه‌ها:
آتش‌سوزی جنگل،
یادگیری ماشین،
مدل فراترکیبی،
سنتینل ۲ و لندست ۸،
سنجش‌ازدور.

چکیده

هدف: آتش‌سوزی جنگل‌ها یکی از مخرب‌ترین سوانح طبیعی است که پیامدهای اکولوژیکی و اقتصادی جبران‌ناپذیری به همراه دارد. تهیه نقشه‌های دقیق خطر حریق، ابزاری حیاتی برای مدیریت و پیشگیری از این بحران است. پژوهش حاضر با هدف توسعه یک چهارچوب یکپارچه مبتنی بر الگوریتم‌های یادگیری ماشین و ارزیابی تأثیر قدرت تفکیک مکانی داده‌های ماهواره‌ای بر دقت مدل‌سازی پتانسیل خطر آتش‌سوزی انجام شد.

روش پژوهش: بدین منظور، داده‌های طیفی و حرارتی دو سنجنده لندست ۸ (با وضوح ۳۰ متر) و سنتینل ۲ (با وضوح ۱۰ متر) به همراه متغیرهای توپوگرافیک و اقلیمی مورداستفاده قرار گرفتند. در این مقاله، پنج مدل یادگیری ماشین پایه مختلف که ازجمله مدل‌های شناخته‌شده یادگیری ماشین هستند، آموزش‌دیده و درنهایت نتایج حاصل از این مدل‌ها به کمک یک مدل فراترکیبی جهت پیش‌بینی مناطق مستعد آتش‌سوزی آموزش داده شدند. این امر سبب گردید تا از تمامی قابلیت‌های متفاوت موجود در مدل‌های یادگیری ماشین بهره برده و الگوهای پیچیده غیرخطی موجود در داده‌های ورودی بهتر مدل شوند و تا جای ممکن مدلی با ثبات بیشتر ایجاد شود.

یافته‌ها: تحلیل‌های اکتشافی نشان داد که شاخص‌های NDVI و NBR به همراه متغیرهای ارتفاع و دمای سطح زمین، بیشترین نقش را در تفکیک مناطق دچار حریق دارند. نتایج ارزیابی کمی نشان داد که مدل فراترکیبی پیاده‌سازی‌شده روی داده‌های سنتینل ۲ با دستیابی به صحت معادل ۹۴/۵۲ درصد، F1–Score برابر با ۸۷/۷۶ درصد و شاخص IoU معادل ۸۰/۱۲ درصد عملکرد مناسبی را در میان تمامی مدل‌های پایه داشته است و این امر خود حاکی از آن است که به‌جای اتکا به یک مدل یادگیری ماشین به‌تنهایی، می‌توان از تلفیق قابلیت‌های متفاوت مدل‌های مختلف جهت بهبود دقت و ثبات مدل بهره برد. همچنین، مقایسه نتایج حاصل از دو سنجنده نشان داد که ارتقای رزولوشن مکانی در سنتینل ۲ منجر به بهبود دقت برآورد پتانسیل خطر آتش‌سوزی شده و نقشه‌های خروجی آن به دلیل کاهش اثر میانگین‌گیری مکانی، از وضوح و مرزبندی به‌مراتب دقیق‌تری برخوردارند.

نتیجه‌گیری: نتایج این مطالعه نشان می‌دهد که ادغام مدل‌های فراترکیبی با داده‌های ماهواره‌ای با وضوح بالا، می‌تواند به‌عنوان یک سیستم کارآمد در فرایند مدیریت پیشگیرانه منابع طبیعی مورداستفاده قرار گیرد.

استناد: مسواری؛ محدثه، شاه حسینی؛ رضا. (۱۴۰۴). نقشه‌برداری پتانسیل خطر آتش‌سوزی جنگل با استفاده از مدل یادگیری ماشین فراترکیبی: ارزیابی عملکردی داده‌های سنتینل ۲ و لندست ۸. *تاب‌آوری در برابر حوادث و سوانح*، ۱ (۴)، ۷۵-۹۴.

CC BY NC

© نویسندگان.

ناشر: پژوهشکده سوانح طبیعی.

مقدمه

آتش‌سوزی‌های طبیعی به‌عنوان یکی از مخرب‌ترین پدیده‌های محیطی، سالانه خسارات و آسیب‌های بسیاری به اکوسیستم‌ها، زیرساخت‌ها، محیط‌های جنگلی و جوامع انسانی وارد می‌کنند. در دهه‌های اخیر، تغییرات اقلیمی و دخالت‌های انسانی سبب تشدید فراوانی و شدت این حوادث شده‌اند. دراین‌بین آتش‌سوزی‌های جنگلی ازجمله مهم‌ترین مخاطرات زیست‌محیطی به شمار می‌آیند که می‌توانند پیامدهای جبران‌ناپذیری بر پویایی اکوسیستم‌ها، چرخه کربن، کیفیت هوا و سلامت جامعه بشری بر جای گذارند (Jones et al., 2020). افزایش فراوانی و شدت این آتش‌سوزی‌ها در سال‌های اخیر تحت تأثیر تغییرات اقلیمی، تغییرات کاربری اراضی و بهره‌برداری بیش‌ازحد از منابع طبیعی، ضرورت بازنگری در رویکردهای پایش، پیش‌بینی و مدیریت پتانسیل خطر آتش‌سوزی را دوچندان کرده است (Abatzoglou et al., 2019). بنابراین، توسعه روش‌های علمی برای پیش‌بینی و ارزیابی پتانسیل خطر آتش‌سوزی به یک نیاز حیاتی در مدیریت بحران و برنامه‌ریزی محیطی تبدیل شده است. فقدان سیستم‌های پایش و پیش‌آگاهی دقیق و بهنگام، چالش عمده‌ای در کاهش آسیب‌پذیری مناطق مستعد محسوب می‌شود.

در این میان، داده‌ها و تصاویر سنجش‌ازدوری به‌عنوان یکی از منابع اطلاعاتی جهت پایش و تحلیل فرایندهای سطح زمین، جایگاه ویژه‌ای در مدیریت مخاطرات طبیعی، ازجمله آتش‌سوزی، یافته‌اند. تصاویر ماهواره‌ای با قابلیت پوشش مکانی گسترده، تکرار زمانی منظم و تنوع طیفی، امکان استخراج اطلاعات متنوعی از وضعیت پوشش گیاهی، رطوبت، ویژگی‌های سطح زمین و شرایط زیست‌محیطی را فراهم می‌سازند. در دهه‌های اخیر، کاربرد تصاویر ماهواره‌ای در مدیریت سوانح مختلف از قبیل سیلاب، خشکسالی، رانش زمین و آتش‌سوزی، به‌طور چشمگیری گسترش یافته است و مزیت آن‌ها در پشتیبانی از تصمیم‌گیری‌های مدیریتی و برنامه‌ریزی راهبردی به‌خوبی به اثبات رسیده است. باوجود پیشرفت‌های قابل‌توجه در حوزه سنجش‌ازدور و یادگیری ماشین، هنوز شکاف دانشی در زمینه به‌کارگیری هم‌زمان داده‌های ماهواره‌ای با رزولوشن‌های مختلف و الگوریتم‌های پیشرفته جهت برآورد پتانسیل خطر آتش‌سوزی وجود دارد. بسیاری از مطالعات پیشین تنها بر یک منبع داده ماهواره‌ای متمرکز بوده‌اند یا از روش‌های سنتی آماری استفاده کرده‌اند که قادر به مدل‌سازی پیچیدگی‌های غیرخطی و تعاملات چندمتغیره در پدیده آتش‌سوزی نیستند. بنابراین، طراحی یک چهارچوب یکپارچه که بتواند از قابلیت‌های مکمل تصاویر چندسنجنده و الگوریتم‌های یادگیری ماشین بهره‌برداری نماید، ضرورتی انکارناپذیر است. سنجش‌ازدور به‌عنوان یک فناوری تحول‌آفرین، امکان پایش پیوسته و در مقیاس‌های گسترده محیطی را فراهم می‌سازد. این داده‌ها با ارائه اطلاعات مکانی–زمانی از شاخص‌های کلیدی مانند دمای سطح زمین، پوشش گیاهی، رطوبت و توپوگرافی، بستر مناسبی برای شناسایی الگوهای خطر و پیش‌بینی وقایع فراهم می‌نمایند.

پیشرفت و تنوع مأموریت‌های سنجش‌ازدور، به‌ویژه پرتاب ماهواره‌هایی با رزولوشن مکانی، طیفی و زمانی بالاتر، فرصت‌های جدیدی برای تحلیل دقیق‌تر الگوهای خطر و پتانسیل خطر فراهم کرده است. در این میان، داده‌های لندست ۸ و سنتینل ۲ به‌عنوان دو منبع کلیدی با پوشش جهانی، رزولوشن مکانی مناسب و دسترسی آزاد، نقش مهمی در مطالعات مرتبط با پوشش گیاهی و مخاطرات مرتبط با آن ایفا می‌کنند (Gitas et al., 2009). تفاوت در ویژگی‌های مکانی و طیفی این دو سنجنده، این امکان را فراهم می‌آورد که اثر رزولوشن مکانی بر عملکرد مدل‌های برآورد پتانسیل خطر موردبررسی قرار گیرد و درک روشن‌تری از محدودیت‌ها و قابلیت‌های هر یک در سنجش پتانسیل خطر آتش‌سوزی حاصل شود.

انقلاب داده‌محور در دو دهه اخیر، الگوریتم‌های یادگیری ماشین و یادگیری عمیق را به ابزارهای استاندارد در تحلیل‌های محیطی و علوم زمین تبدیل کرده است. این روش‌ها با توانایی کشف الگوهای پیچیده از داده‌های با ابعاد بالا، مدل‌سازی روابط غیرخطی و تعمیم‌پذیری به داده‌های مشاهده‌نشده، مزیت‌های قابل‌توجهی نسبت به روش‌های کلاسیک آماری ارائه می‌دهند. در حوزه خطرپذیری محیطی، الگوریتم‌هایی مانند جنگل تصادفی، ماشین بردار پشتیبان و شبکه‌های عصبی عمیق، عملکرد برتر خود را در پیش‌بینی مخاطرات طبیعی به اثبات رسانده‌اند. بااین‌حال، انتخاب بهترین معماری مدل برای داده‌های خاص

سنجش‌ازدور و شرایط محیطی مختلف همچنان یک چالش تحقیقاتی باقی مانده است. به‌طورکلی، توسعه روش‌های یادگیری ماشین و یادگیری عمیق، چشم‌اندازهای نوینی را در تحلیل داده‌های حجیم و پیچیده سنجش‌ازدور گشوده است. در سال‌های اخیر، کاربرد الگوریتم‌های یادگیری ماشین و یادگیری عمیق در حوزه‌های مختلف علوم زمین، ازجمله طبقه‌بندی کاربری و پوشش زمین، پایش تغییرات، برآورد پارامترهای زیست‌محیطی و نیز مدل‌سازی خطر و پتانسیل خطر سوانح، روندی رو به رشد داشته است. بهره‌گیری از این روش‌ها در کنار داده‌های ماهواره‌ای توسعه‌یافته، می‌تواند به ارتقای دقت و کارایی مدل‌های پتانسیل خطر آتش‌سوزی کمک نماید.

بر این اساس، هدف اصلی این پژوهش توسعه یک چهارچوب هوشمند، یکپارچه و مبتنی بر داده‌های سنجش‌ازدور و الگوریتم‌های یادگیری ماشین برای تهیه نقشه پتانسیل خطر آتش‌سوزی جنگل است. بنابراین، اهداف تفصیلی پژوهش شامل موارد زیر است:

نخست، استخراج و آماده‌سازی مجموعه‌ای از متغیرهای مؤثر بر وقوع آتش‌سوزی شامل شاخص‌های طیفی، ویژگی‌های پوشش گیاهی، متغیرهای حرارتی، داده‌های توپوگرافی و پارامترهای اقلیمی از داده‌های ماهواره‌ای لندست ۸، سنتینل ۲، SRTM، ERA5 و نقاط آتش‌سوزی FIRMS.

دوم، بررسی و مقایسه اثر توان تفکیک مکانی داده‌های ماهواره‌ای بر دقت مدل‌سازی پتانسیل خطر آتش‌سوزی ازطریق مقایسه داده‌های سنتینل ۲ با تفکیک مکانی ۱۰ متر و لندست ۸ با تفکیک مکانی ۳۰ متر.

سوم، ارزیابی عملکرد چندین الگوریتم یادگیری ماشین شامل جنگل تصادفی، درخت تصمیم، ماشین بردار پشتیبان، رگرسیون لجستیک و گرادیان بوستینگ هیستوگرام در برآورد مناطق مستعد آتش‌سوزی.

چهارم، طراحی و پیاده‌سازی یک مدل فراترکیبی مبتنی بر ترکیب خروجی مدل‌های پایه با هدف افزایش دقت، کاهش واریانس خطا و بهبود پایداری پیش‌بینی‌ها؛ پنجم، شناسایی نقش نسبی متغیرهای کلیدی مانند NDVI، NBR، دمای سطح زمین و ارتفاع در تفکیک مناطق پرخطر و کم‌خطر.

و درنهایت، تولید و تحلیل نقشه‌های پتانسیل خطر آتش‌سوزی به‌منظور ارائه ابزاری کاربردی برای مدیریت پیشگیرانه منابع طبیعی، پشتیبانی از تصمیم‌گیری مکانی و ارتقای سامانه‌های هشدار زودهنگام آتش‌سوزی. تحقق این اهداف می‌تواند ضمن روشن‌سازی نقش داده‌های ماهواره‌ای با تفکیک مکانی متفاوت در مدل‌سازی خطر آتش‌سوزی، چهارچوبی قابل‌اتکا برای استفاده عملیاتی از مدل‌های هوشمند در مدیریت بحران‌های محیطی فراهم سازد.

پیشینه پژوهش

در سال‌های اخیر، مدل‌سازی و پهنه‌بندی پتانسیل خطر آتش‌سوزی جنگل از رویکردهای سنتی و شاخص‌محور به سمت روش‌های داده‌محور و الگوریتم‌های یادگیری ماشین حرکت کرده است. علت اصلی این تغییر، ماهیت پیچیده و غیرخطی پدیده آتش‌سوزی است؛ زیرا وقوع آتش‌سوزی معمولاً حاصل برهم‌کنش هم‌زمان عوامل گوناگونی مانند وضعیت پوشش گیاهی، دمای سطح زمین، رطوبت، بارش، شیب، ارتفاع، جهت شیب، فاصله از راه‌ها و فعالیت‌های انسانی است. الگوریتم‌های یادگیری ماشین به دلیل توانایی در کشف الگوهای پنهان، مدل‌سازی روابط غیرخطی و استفاده از داده‌های چندمنبعی سنجش‌ازدور و مکانی، در سال‌های اخیر به یکی از ابزارهای اصلی در پیش‌بینی و نقشه‌برداری حساسیت آتش‌سوزی تبدیل شده‌اند. مرورهای جدید نیز نشان می‌دهند که مطالعات نوین آتش‌سوزی به‌طور فزاینده‌ای از داده‌های سنجش‌ازدور، لایه‌های مکانی، داده‌های اقلیمی، موارد ثبت‌شده آتش و مدل‌های یادگیری ماشین، یادگیری عمیق و مدل‌های ترکیبی استفاده می‌کنند؛ هرچند چالش‌هایی مانند کمبود داده آموزشی، تفسیرپذیری مدل‌ها، ناهمگنی داده‌ها و تعمیم‌پذیری مکانی همچنان پابرجا است (Ejaz & Choudhury, 2025; Xu et al., 2025).

یکی از مطالعات مهم در این زمینه توسط تونینی[1] و همکاران (۲۰۲۰) انجام شد که با استفاده از الگوریتم جنگل تصادفی، نقشه حساسیت آتش‌سوزی را برای منطقه لیگوریا در ایتالیا تهیه کردند. در این پژوهش از داده‌های پیرامون آتش‌سوزی در بازه زمانی بلندمدت، به همراه عوامل محیطی مؤثر مانند نوع پوشش زمین، پوشش گیاهی، شبکه راه‌ها، ارتفاع و سایر مشتقات توپوگرافی استفاده شد. مزیت اصلی این مطالعه، استفاده از سری زمانی نسبتاً طولانی و اعتبارسنجی مکانی بود که موجب شد قابلیت تعمیم مدل در شرایط فصلی مختلف بهتر بررسی شود. همچنین، تفکیک تحلیل‌ها برای فصل‌های مختلف نشان داد که الگوهای حساسیت آتش‌سوزی می‌توانند در تابستان و زمستان متفاوت باشند. بااین‌حال، محدودیت اصلی این پژوهش اتکا به یک الگوریتم منفرد، یعنی جنگل تصادفی، و تمرکز بیشتر بر عوامل ایستا و مکانی بود؛ درنتیجه نقش مدل‌های ترکیبی و مقایسه داده‌های ماهواره‌ای با تفکیک مکانی متفاوت به‌طور مستقیم بررسی نشد (Tonini et al., 2020).

گیگوویچ[2] و همکاران (۲۰۱۹) در پارک ملی تارا در صربستان، الگوریتم‌های ماشین بردار پشتیبان، جنگل تصادفی و یک مدل ترکیبی مبتنی بر میانگین‌گیری بیزی را برای تهیه نقشه حساسیت آتش‌سوزی جنگل به کار گرفتند. در این مطالعه، پایگاه داده آتش‌سوزی با استفاده از داده‌های تاریخی، تصاویر MODIS، لندست ۸، WorldView-2، بازدید میدانی و تفسیر تصاویر هوایی تهیه شد و ۱۲۶ نقطه آتش‌سوزی به دو بخش آموزش و اعتبارسنجی تقسیم گردید. نتایج نشان داد که مدل ترکیبی با مقدار AUC برابر با ۰/۸۴۸ عملکرد بهتری نسبت به SVM با AUC برابر با ۰/۸۴۴ و RF با AUC برابر با ۰/۸۳۴ داشته است. مزیت این مطالعه، استفاده از داده‌های چندمنبعی و نشان دادن برتری رویکردهای Ensemble نسبت به مدل‌های منفرد بود. بااین‌حال، تعداد نمونه‌های آتش‌سوزی نسبتاً محدود، استفاده از تقسیم‌بندی تصادفی داده‌ها و عدم بررسی اثر تفکیک مکانی سنجنده‌ها از محدودیت‌های آن به شمار می‌آید.

در پژوهشی دیگر، قربان‌زاده و همکاران (۲۰۱۹a) برای شهرستان آمل در استان مازندران، الگوریتم‌های شبکه عصبی مصنوعی، ماشین بردار پشتیبان و جنگل تصادفی را برای پیش‌بینی مکانی حساسیت آتش‌سوزی جنگل ارزیابی کردند. در این مطالعه، نقاط آتش‌سوزی با استفاده از داده‌های MODIS و داده‌های میدانی GPS تکمیل شد و مجموعه‌ای از عوامل توپوگرافی، هواشناسی، انسانی، پوشش گیاهی و هیدرولوژیکی به‌عنوان متغیرهای ورودی مورداستفاده قرار گرفت. مزیت مهم این پژوهش، ترکیب داده‌های سنجش‌ازدور با داده‌های میدانی و استفاده از چند گروه متغیر اثرگذار بود. این موضوع باعث شد مدل‌ها بتوانند ابعاد مختلف وقوع آتش‌سوزی را بهتر بازنمایی کنند. بااین‌حال، محدودیت آن در استفاده از مدل‌های پایه و نبود چهارچوب فراترکیبی پیشرفته بود؛ همچنین اثر تفکیک مکانی داده‌های ماهواره‌ای بر دقت خروجی مدل‌ها به‌صورت مستقل بررسی نشد (Ghorbanzadeh, Valizadeh Kamran, et al., 2019).

در مطالعه‌ای دیگر، قربان‌زاده و همکاران (۲۰۱۹*b*) نقشه حساسیت و ریسک آتش‌سوزی جنگل را با ترکیب روش یادگیری ماشین و شاخص آسیب‌پذیری اجتماعی–زیرساختی تولید کردند. در این پژوهش، ابتدا نقشه حساسیت آتش‌سوزی با استفاده از روش یادگیری ماشین و ۱۶ متغیر محیطی تهیه شد و سپس با شاخص‌های آسیب‌پذیری اجتماعی و زیرساختی ترکیب گردید. مزیت این رویکرد آن بود که خطر آتش‌سوزی تنها به احتمال وقوع محدود نشد، بلکه پیامدهای احتمالی آن برای جمعیت، زیرساخت‌ها و کاربری‌های انسانی نیز موردتوجه قرار گرفت. بااین‌وجود، ازآنجاکه هدف مطالعه ترکیب حساسیت و آسیب‌پذیری بود، تمرکز اصلی آن بر مقایسه الگوریتم‌های متعدد یادگیری ماشین یا بررسی اثر رزولوشن مکانی داده‌های ماهواره‌ای قرار نداشت (Ghorbanzadeh, Blaschke, et al., 2019).

کلانتر و همکاران (۲۰۲۰) در حوضه چالوس‌رود ایران، سه مدل یادگیری ماشین شامل MARS، SVM و BRT را برای پیش‌بینی حساسیت آتش‌سوزی جنگل بر پایه داده‌های سنجش‌ازدور و متغیرهای محیطی به کار گرفتند. در این مطالعه از ۱۴ متغیر اثرگذار شامل شاخص‌های پوشش گیاهی، عوامل اقلیمی، محیطی و توپوگرافی استفاده شد و برای کاهش آثار ناشی از

1. Tonini
2. Gigović

نحوه تقسیم داده‌ها، روش‌های بازنمونه‌گیری مانند اعتبارسنجی متقاطع، Bootstrap و Optimism Bootstrap مورداستفاده قرار گرفت. مزیت اصلی این پژوهش توجه به مسئله عدم قطعیت و سوگیری ناشی از نمونه‌برداری بود؛ زیرا بسیاری از مطالعات تنها به تقسیم ساده آموزش و آزمون اکتفا می‌کنند. محدودیت این مطالعه آن بود که تمرکز اصلی بر مقایسه روش‌های بازنمونه‌گیری و چند الگوریتم مشخص بود و به موضوع ترکیب چند مدل پایه در یک مدل فراترکیبی یا مقایسه مستقیم داده‌های سنتینل و لندست پرداخته نشد.

در پژوهشی نزدیک به موضوع مقاله حاضر، توکلی و همکاران (۲۰۲۲)، اثر تفکیک مکانی داده‌های سنجش‌ازدور را بر پیش‌بینی حساسیت آتش‌سوزی در استان گیلان بررسی کردند. در این مطالعه، داده‌های با تفکیک مکانی پایین‌تر شامل لندست ۸ و SRTM و داده‌های با تفکیک مکانی متوسط شامل سنتینل ۲ و ALOS با استفاده از مدل‌های جنگل تصادفی و ماشین بردار پشتیبان ارزیابی شدند. نتایج نشان داد که داده‌های با تفکیک مکانی بالاتر عملکرد بهتری داشتند و ترکیب خروجی‌ها توانست بهترین مقدار AUC، یعنی ۷۱/۹۴ درصد، را ایجاد کند. مزیت این مطالعه، پرداختن مستقیم به اثر تفکیک مکانی داده‌های ماهواره‌ای و استفاده از محیط Google Earth Engine بود. بااین‌حال، محدودیت آن در استفاده از تنها دو الگوریتم پایه و اتکا به روش ادغام داده‌ها بود؛ درحالی‌که استفاده از مدل فراترکیبی یادگیرنده می‌تواند روابط میان خروجی مدل‌های پایه را به‌صورت داده‌محور یاد بگیرد (Tavakkoli Piralilou et al., 2022).

پویان و همکاران (۲۰۲۱) نیز در استان کهگیلویه و بویراحمد، از الگوریتم‌های RF،BRT و SVM برای تهیه نقشه چندمخاطره شامل آتش‌سوزی جنگل، سیلاب، فرسایش خندقی و زلزله استفاده کردند. در بخش آتش‌سوزی جنگل، جنگل تصادفی با مقدار AUC برابر با ۰/۸۸۵ عملکرد بهتری نسبت به SVM و BRT نشان داد. مزیت اصلی این مطالعه، بررسی هم‌زمان چند مخاطره و استفاده از معیار اهمیت متغیرها برای تعیین نقش عوامل مؤثر بود. بااین‌حال، آتش‌سوزی در این پژوهش تنها یکی از مؤلفه‌های نقشه چندمخاطره بود و تمرکز تخصصی بر بهینه‌سازی مدل‌های آتش‌سوزی، مقایسه سنجنده‌های ماهواره‌ای و تولید نقشه‌های پرجزئیات پتانسیل خطر حریق وجود نداشت (Pouyan et al., 2021).

در مطالعات بین‌المللی جدیدتر، توجه بیشتری به مدل‌های تقویتی، یادگیری عمیق و تفسیرپذیری مدل‌ها شده است. برای نمونه، در مطالعه‌ای در جنوب اروپا، از جنگل تصادفی به همراه چهارچوب هوش مصنوعی تبیین‌پذیر و مقادیر SHAP برای شناسایی عوامل مؤثر بر وقوع آتش‌سوزی استفاده شد. نتایج این پژوهش نشان داد که مدل توانسته است مناطق دارای احتمال بالای وقوع آتش‌سوزی را با AUC حدود ۸۱/۳ درصد شناسایی کند و عواملی مانند شاخص آب‌وهوای آتش و NDVI را به‌عنوان متغیرهای مهم معرفی نماید. مزیت این مطالعه، تمرکز بر تفسیرپذیری و قابل‌فهم کردن خروجی مدل برای مدیران و تصمیم‌گیران بود؛ اما محدودیت آن این بود که بیشتر بر تبیین وقوع آتش در مقیاس گسترده تمرکز داشت و مقایسه تفصیلی الگوریتم‌های متعدد یا اثر وضوح مکانی داده‌های ماهواره‌ای در آن محور اصلی نبود (Cilli et al., 2022).

در حوزه یادگیری عمیق نیز بیانس و همکاران یک مدل ترکیبی مبتنی بر شبکه‌های عمیق را برای تهیه نقشه حساسیت آتش‌سوزی در دو منطقه در شیلی ارائه کردند. در این پژوهش از ۱۵ عامل مؤثر شامل ارتفاع، شیب، دما، سرعت باد، بارش، کمبود آب اقلیمی، NDVI، نوع پوشش زمین و فاصله از راه‌ها و مناطق شهری استفاده شد. نتایج نشان داد که مدل ترکیبی عمیق با AUC برابر با ۹۵/۳ عملکرد بهتری نسبت به مدل‌های مقایسه‌ای مانند XGBoost و SVM داشته است. مزیت این مطالعه توانایی مدل در استخراج الگوهای پیچیده مکانی از داده‌ها بود؛ اما محدودیت‌هایی مانند نیاز به داده آموزشی زیاد، هزینه محاسباتی بالا، وابستگی به اندازه پنجره نمونه‌برداری و کاهش تفسیرپذیری نسبت به مدل‌های کلاسیک‌تر یادگیری ماشین همچنان مطرح است (Bjånes et al., 2021).

مطالعات اخیر همچنین نشان داده‌اند که تعمیم‌پذیری مکانی و زمانی یکی از چالش‌های مهم در مدل‌سازی آتش‌سوزی است. برای مثال، در مطالعه‌ای در دو منطقه با شرایط فیزیکی متفاوت در آمریکا و کانادا، عملکرد رگرسیون لجستیک و جنگل تصادفی برای تولید نقشه‌های حساسیت آتش‌سوزی ارزیابی شد. نتایج نشان داد که جنگل تصادفی در بیشتر رویکردها عملکرد بهتری از

رگرسیون لجستیک دارد و متغیرهایی مانند پوشش زمین، دما، باد، ارتفاع، بارش و NDVI از عوامل مهم در پیش‌بینی آتش‌سوزی هستند. بااین‌حال، هنگامی‌که مدل در یک منطقه آموزش داده و در منطقه‌ای دیگر اعتبارسنجی شد، دقت کاهش یافت. این یافته نشان می‌دهد که مدل‌های یادگیری ماشین ممکن است در همان منطقه آموزشی عملکرد مطلوبی داشته باشند، اما برای کاربرد عملیاتی در مناطق جدید، نیازمند اعتبارسنجی مکانی و زمانی دقیق هستند (Moghim & Mehrabi, 2024).

در منابع داخلی فارسی‌زبان نیز روند استفاده از الگوریتم‌های یادگیری ماشین در سال‌های اخیر قابل‌مشاهده است؛ هرچند تعداد این مطالعات نسبت به روش‌های سنتی هنوز محدودتر است. رودسرابی و همکاران (۲۰۲۲) در یک مرور داخلی بر مطالعات سنجش‌ازدور آتش‌سوزی جنگل نشان دادند که در میان پژوهش‌های مربوط به شناسایی و پایش آتش‌سوزی در کشور، سهم روش‌های یادگیری نسبتاً اندک بوده و بخش زیادی از مطالعات از روش‌های سنتی استفاده کرده‌اند. این موضوع نشان می‌دهد که استفاده از الگوریتم‌های یادگیری ماشین در مطالعات آتش‌سوزی ایران هنوز ظرفیت توسعه زیادی دارد. همچنین، برخی مطالعات داخلی، مانند پژوهش‌های مبتنی بر جنگل تصادفی و ماشین بردار پشتیبان در تهیه نقشه حساسیت آتش‌سوزی، نشان داده‌اند که RF معمولاً نسبت به SVM عملکرد بهتری دارد و متغیرهایی مانند جهت شیب، فاصله از آبراهه، فاصله از جاده، شیب و شاخص‌های توپوگرافی از عوامل مؤثر در پهنه‌بندی خطر محسوب می‌شوند. در مطالعه‌ای دیگر رودسرابی و همکاران (۲۰۲۲) در مراتع شهرستان اراک نیز از الگوریتم‌های RF و SVM و مجموعه‌ای از عوامل محیطی، توپوگرافی، کاربری اراضی و فاصله از شبکه ارتباطی و آبراهه برای تهیه نقشه خطر آتش‌سوزی استفاده نمودند.

جمع‌بندی مطالعات پیشین نشان می‌دهد که الگوریتم‌های یادگیری ماشین، به‌ویژه جنگل تصادفی، ماشین بردار پشتیبان، شبکه عصبی مصنوعی، مدل‌های بوستینگ، XGBoost و مدل‌های ترکیبی، توانسته‌اند عملکرد قابل قبولی در پهنه‌بندی حساسیت و پتانسیل خطر آتش‌سوزی ارائه دهند. بااین‌حال، چند شکاف پژوهشی همچنان قابل‌شناسایی است. نخست، بسیاری از مطالعات تنها از یک سنجنده یا یک سطح تفکیک مکانی استفاده کرده‌اند و اثر تفاوت رزولوشن مکانی داده‌های ماهواره‌ای بر دقت مدل به‌طور مستقیم و هم‌زمان بررسی نشده است. دوم، در بخش قابل‌توجهی از پژوهش‌ها، مدل‌های منفرد به کار رفته‌اند و استفاده از چهارچوب‌های فراترکیبی که خروجی چند الگوریتم پایه را به‌صورت یادگیرنده ترکیب کند، کمتر موردتوجه قرار گرفته است. سوم، در بسیاری از مطالعات، تقسیم داده‌ها به‌صورت تصادفی انجام شده و اعتبارسنجی زمانی یا شبیه‌سازی شرایط واقعی پیش‌بینی کمتر بررسی شده است. چهارم، بخش مهمی از مطالعات داخلی هنوز به روش‌های سنتی یا تعداد محدودی از الگوریتم‌های یادگیری ماشین وابسته‌اند و نیاز به پژوهش‌هایی وجود دارد که هم‌زمان از داده‌های چندسنجنده، شاخص‌های طیفی، متغیرهای توپوگرافی و اقلیمی و مدل‌های پیشرفته یادگیری ماشین استفاده کنند. بر همین اساس، پژوهش حاضر با مقایسه داده‌های سنتینل ۲ و لندست ۸، ارزیابی چند مدل پایه یادگیری ماشین و توسعه یک مدل فراترکیبی، تلاش می‌کند بخشی از این خلأها را پوشش دهد و چهارچوبی دقیق‌تر و قابل‌اتکاتر برای نقشه‌برداری پتانسیل خطر آتش‌سوزی جنگل ارائه نماید.

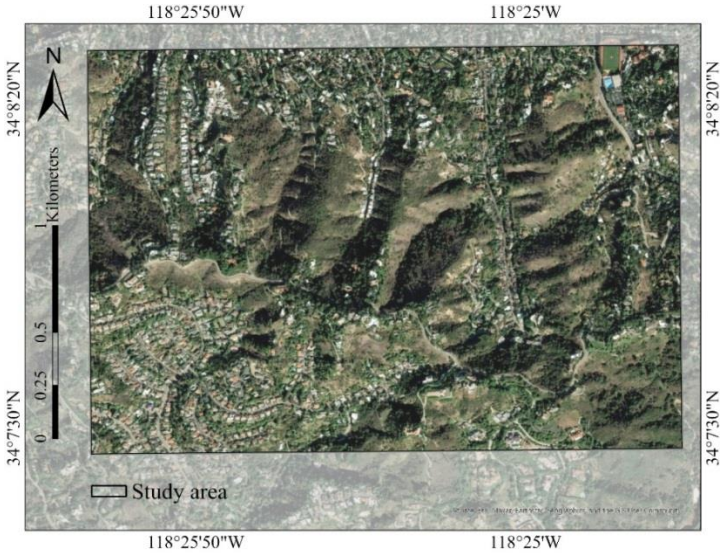
روش‌شناسی پژوهش

این مطالعه با هدف توسعه یک چهارچوب پیشرفته و خودکار جهت برآورد نقشه پتانسیل خطر آتش‌سوزی با استفاده از داده‌های سنجش‌ازدور و الگوریتم‌های یادگیری ماشین انجام شده است. در این تحقیق یک معماری فراترکیبی مورداستفاده قرار گرفته است که ازطریق ترکیب پیش‌بینی‌های چندین مدل پایه یادگیری ماشین قادر است نتایجی با دقت و قابلیت اطمینان بالاتر حاصل نماید. انتخاب این رویکرد به دلیل ماهیت پیچیده، غیرخطی و چندبعدی روابط بین عوامل مؤثر بر وقوع آتش‌سوزی است که تحلیل آن با روش‌های کلاسیک آماری با چالش‌های اساسی مواجه است. الگوریتم‌های یادگیری ماشین با توانایی ذاتی در کشف الگوهای پنهان و تعاملات پیچیده در داده‌های حجیم و ناهمگن، به‌عنوان هسته محاسباتی این پژوهش انتخاب شده‌اند.

جمع‌آوری و پیش‌پردازش داده‌ها

در این مطالعه، وقوع آتش‌سوزی در منطقه‌ای در شهر لس‌آنجلس در ایالت کالیفرنیا در ایالات‌متحده آمریکا موردبررسی قرار

گرفته است. این منطقه در دامنه‌های جنوبی محدوده کوهستانی Santa Monica بوده در نزدیکی محدوده هالیوود واقع است. این منطقه دارای اقلیم مدیترانه‌ای گرم، تابستان‌های گرم و خشک و زمستان‌های ملایم و نسبتاً مرطوب، بوده و پوشش آن از نوع جنگل و بوته‌زار است. این منطقه جزء مناطق پرپتانسیل خطر برای آتش‌سوزی‌های جنگلی محسوب می‌گردد و ازاین‌رو مطالعه پتانسیل خطر آتش‌سوزی در این منطقه از اهمیت بالایی برخوردار است. شکل زیر محدوده مربوط به منطقه موردمطالعه را نشان می‌دهد.



شکل ۱. محدوده موردمطالعه

در این تحقیق به‌منظور برآورد نقشه پتانسیل خطر آتش‌سوزی، داده‌ها و تصاویر ماهواره‌ای معتبر و در دسترس مورداستفاده قرار گرفته‌اند. به‌صورت کلی مجموعه داده موردنظر در این پژوهش در پلتفرم ابری گوگل ارث انجین[3] که امکان دسترسی به آرشیو تصاویر و داده‌های ماهواره‌ای در آن فراهم است، در بازه زمانی ۲۰۱۵ تا پایان سال ۲۰۲۳ اخذ شده‌اند.

مجموعه تصاویر ماهواره‌ای که در این مطالعه مورداستفاده قرار گرفته است از ماهواره لندست-۸ (توان تفکیک مکانی ۳۰ متر برای باندهای مرئی) و سنتینل-۲ (توان تفکیک مکانی ۱۰ متر برای باندهای مرئی)، با توان تفکیک مکانی متفاوت اخذ شده‌اند؛ چراکه یکی از اهداف مدنظر این تحقیق بررسی اثر توان تفکیک مکانی بر برآورد پتانسیل خطر آتش است. در این پژوهش علاوه بر باندهای طیفی مختلف، شاخص‌های طیفی همچون شاخص پوشش گیاهی تفاضلی نرمال‌شده (NDVI) و شاخص نرمال‌شده سوختگی (NBR) مستخرج از این باندهای طیفی نیز در روند برآورد نقشه پتانسیل خطر مورداستفاده قرار گرفته‌اند. تصحیحات رادیومتریک و اتمسفری نیز ابتدا بر روی باندهای طیفی اعمال شده و به ترتیب تصاویر رادیانس و رفلکتنس برای هر باند به دست آمد.

علاوه بر تصاویر ماهواره‌ای، داده‌های ارتفاعی و توپوگرافی SRTM با توان تفکیک مکانی ۳۰ متری مورداستفاده قرار گرفته و از این داده ارتفاعی ویژگی‌های مختلف شامل شیب و جهت شیب نیز به دست آمدند. در کنار این داده‌های از جنس توپوگرافی منطقه، داده‌های محیطی و هواشناسی نیز شامل دمای سطح زمین، میزان بارش، سرعت و جهت باد نیز از مجموعه داده ERA5 ازطریق وبگاه گوگل ارث انجین اخذ گردید.

در مقابل نیز نقاط آتش‌سوزی ازطریق سامانه FIRMS[4] برای بازه زمانی ۲۰۱۵ تا ۲۰۲۵ با توان تفکیک مکانی ۳۰ متر استخراج شد. مختصات این نقاط به‌عنوان مرکز آتش‌سوزی در نظر گرفته شده و نقاط اطراف آن با احتمال بیشتری درگیر وقوع

3. Google Earth Engine
4. Fire Information for Resource Management System

آتش‌سوزی شده است. بدیهی است که با افزایش فاصله از نقاط آتش‌سوزی احتمال وقوع برای سایر نقاط کم و کمتر می‌گردد.

در این پژوهش، باندهای طیفی و شاخص‌های طیفی مورداستفاده و سایر پارامترها به‌صورت میانگین‌گیری‌شده در یک بازه زمانی بلندمدت وارد مدل نشدند؛ بلکه برای هر رخداد آتش‌سوزی، تصاویر و داده‌های ماهواره‌ای مربوط به دوره پیش از وقوع آتش‌سوزی اخذ شدند و نزدیک‌ترین تصویر معتبر پیش از تاریخ رخداد، ازنظر زمانی، به‌عنوان نماینده شرایط پیشاحریق انتخاب شد. هدف از این کار آن بود که وضعیت واقعی پوشش گیاهی، رطوبت سطحی، خشکی زیست‌توده و شرایط طیفی منطقه در نزدیک‌ترین زمان ممکن قبل از وقوع آتش‌سوزی در مدل لحاظ شود. بنابراین، هیچ‌یک از باندها یا شاخص‌های طیفی بر اساس تصاویر پس از آتش‌سوزی محاسبه نشدند تا از ورود اطلاعات مربوط به پیامدهای آتش‌سوزی به مدل و ایجاد نشتی داده جلوگیری شود.

ازآنجاکه داده‌های مورداستفاده در این پژوهش از منابع مختلف با تفکیک مکانی متفاوت گردآوری شده‌اند، پیش از ورود داده‌ها به مدل، تمام لایه‌های ورودی ازنظر سیستم مختصات، محدوده مکانی و اندازه پیکسل یکسان‌سازی شدند. این فرایند به‌صورت جداگانه برای دو سناریوی سنجنده‌ای انجام شد. در سناریوی لندست ۸، باندهای طیفی، شاخص‌های طیفی و سایر متغیرهای ورودی به شبکه مکانی ۳۰ متر بازنمونه‌گیری شدند. در سناریوی سنتینل ۲ نیز تمام لایه‌های ورودی به شبکه مکانی ۱۰ متر بازنمونه‌گیری شدند. روش مورداستفاده برای بازنمونه‌گیری در هر دو سناریو، روش نزدیک‌ترین همسایگی بود. استفاده از این روش باعث شد مقادیر اصلی پیکسل‌ها حفظ شده و از تغییرات ناشی از درون‌یابی یا میانگین‌گیری مقادیر طیفی جلوگیری شود. پس‌ازاین مرحله، مقادیر لایه‌های هم‌مکان‌شده برای نمونه‌های آتش‌سوزی و عدم آتش‌سوزی استخراج و به‌عنوان ورودی مدل‌های یادگیری ماشین استفاده شد.

پس از اخذ داده‌های موردنیاز جهت برآورد نقشه پتانسیل خطر آتش‌سوزی در بازه زمانی ۲۰۱۵ تا ۲۰۲۵ این مجموعه داده به دو مجموعه داده آموزشی و آزمایشی، ۷۰ درصد داده آموزشی و ۳۰ درصد داده آزمایشی، تقسیم شد. درنهایت نیز نقشه پتانسیل خطر آتش‌سوزی برای آتش‌سوزی مربوط به ۱۵ ام ماه فوریه سال ۲۰۲۵ برآورد گردید.

معماری مدل‌سازی

در این پژوهش با ادغام داده‌های ماهواره‌ای و سنجش‌ازدوری مختلف و به‌کارگیری مجموعه‌ای از مدل‌های یادگیری ماشین با قابلیت‌های مکمل و درنهایت ترکیب بهینه پیش‌بینی‌های هریک از مدل‌های یادگیری ماشین با یک مدل فراترکیبی، چهارچوبی جامع، دقیق و قابل‌اتکا جهت برآورد پتانسیل خطر آتش‌سوزی ارائه شده است. این رویکرد نه‌تنها قادر به بهبود دقت پیش‌بینی خواهد بود بلکه با کاهش واریانس و افزایش قابلیت اطمینان ابزار مناسبی جهت تصمیم‌گیری‌های مدیریت پتانسیل خطر آتش‌سوزی در مقیاس‌های مختلف مکانی و زمانی فراهم می‌نماید.

در این پژوهش، پنج الگوریتم یادگیری ماشین با قابلیت‌های مکمل به‌عنوان مدل‌های پایه انتخاب شدند که شامل جنگل تصادفی (Leo., 2001)، درخت تصمیم، ماشین بردار پشتیبان، رگرسیون لجستیک و گرادیان بوستینگ هیستوگرام بودند. انتخاب این مدل‌ها با هدف پوشش طیفی از روش‌های خطی، غیرخطی، مبتنی بر درخت و مبتنی بر یادگیری جمعی انجام شد. مطالعات جدید در حوزه پیش‌بینی و پهنه‌بندی خطر آتش‌سوزی نشان داده‌اند که ترکیب داده‌های سنجش‌ازدور، متغیرهای مکانی و اقلیمی با الگوریتم‌های یادگیری ماشین می‌تواند روابط پیچیده میان پوشش گیاهی، شرایط حرارتی، توپوگرافی و وقوع آتش‌سوزی را با دقت بالایی مدل‌سازی نماید (Ejaz & Choudhury, 2025). در این میان، مدل‌های مبتنی بر درخت مانند جنگل تصادفی و مدل‌های بوستینگ به دلیل توانایی در مدل‌سازی روابط غیرخطی، تعاملات پیچیده بین متغیرها و مقاومت نسبی در برابر نویز، در مطالعات اخیر آتش‌سوزی جنگل عملکرد مطلوبی نشان داده‌اند (Shahzad et al., 2025). همچنین استفاده از رگرسیون لجستیک و ماشین بردار پشتیبان در کنار مدل‌های درختی، امکان مقایسه عملکرد مدل‌های خطی، فاصله‌محور و غیرخطی را فراهم می‌سازد و موجب می‌شود ارزیابی نهایی تنها به یک خانواده خاص از الگوریتم‌ها محدود نشود.

عملکرد مدل‌های یادگیری ماشین به‌شدت به انتخاب مقادیر مناسب برای هایپرپارامترهایشان وابسته است. هایپرپارامترها،

پارامترهایی هستند که قبل از شروع فرایند آموزش مدل توسط کاربر تنظیم می‌شوند و ساختار و فرایند یادگیری مدل را کنترل می‌کنند (مانند تعداد درختان در مدل جنگل تصادفی یا پارامتر تنظیم C در ماشین بردار پشتیبان). یافتن ترکیب بهینه از این هایپرپارامترها یک چالش کلیدی در ساخت مدل‌های با کارایی بالا است. در این پژوهش، برای یافتن بهترین ترکیب هایپرپارامترها برای هر یک از مدل‌های پایه شامل جنگل تصادفی، درخت تصمیم، رگرسیون لجستیک و گرادیان بوستینگ هیستوگرام، از روش جستجوی تصادفی[5] همراه با اعتبارسنجی متقاطع[6] استفاده شد. برخلاف روش جستجوی شبکه‌ای[7] که تمام ترکیبات ممکن در یک شبکه از پیش تعریف‌شده را به‌صورت جامع و زمان‌بر ارزیابی می‌کند، جستجوی تصادفی تعداد معینی از ترکیبات هایپرپارامتر را از توزیع‌های آماری مشخصی به‌صورت تصادفی نمونه‌برداری می‌کند. مطالعات نشان داده‌اند که جستجوی تصادفی در بسیاری از موارد ازنظر محاسباتی کارآمدتر است، زیرا تمام هایپرپارامترها از اهمیت یکسانی برخوردار نیستند و این روش شانس بیشتری برای یافتن مقادیر بهینه برای پارامترهای مهم‌تر فراهم می‌آورد. فرایند بهینه‌سازی برای هر مدل به شرح زیر انجام می‌شود:

– تعریف فضای جستجو: برای هر هایپرپارامتر کلیدی مدل، یک محدوده یا توزیع مقادیر ممکن تعریف گردید. به‌عنوان‌مثال، برای مدل جنگل تصادفی، فضای جستجو شامل هایپرپارامترهایی نظیر تعداد درختان[8]، حداکثر عمق هر درخت[9]، حداقل تعداد نمونه برای تقسیم یک گره[10] و حداکثر تعداد ویژگی‌های مورداستفاده در هر تقسیم[11] بود.

– نمونه‌برداری و ارزیابی: الگوریتم جستجوی تصادفی، تعداد n ترکیب مختلف از هایپرپارامترها را از فضای جستجوی تعریف‌شده نمونه‌برداری کرد. برای هر ترکیب نمونه‌برداری شده، یک فرایند اعتبارسنجی متقاطع بر روی مجموعه داده آموزشی اجرا شد. استفاده از اعتبارسنجی متقاطع تضمین می‌کند که ارزیابی عملکرد هر ترکیب قوی بوده و از بیش‌برازش[12] به یک بخش خاص از داده‌های آموزشی جلوگیری می‌شود.

– انتخاب بهترین مدل: معیاری که برای ارزیابی عملکرد در هر لایه از اعتبارسنجی متقاطع استفاده شد، امتیاز F1 بود. این معیار به دلیل توانایی آن در ایجاد توازن بین دقت[13] و بازیابی[14]، برای مسائل دسته‌بندی نامتوازن مانند پیش‌بینی آتش‌سوزی، انتخاب مناسبی است. پس از اتمام تمام تکرارها، ترکیبی از هایپرپارامترها که منجر به بالاترین میانگین امتیاز F1 در طول فرایند اعتبارسنجی متقاطع شد، به‌عنوان ترکیب بهینه انتخاب شد.

در این تحقیق، ساختار اعتبارسنجی متقاطع به‌صورت K-Fold پنج‌بخشی تعریف گردید و در هر مرحله، چهار بخش از داده‌ها برای آموزش و یک بخش برای اعتبارسنجی مورداستفاده قرار گرفت. این فرایند پنج بار تکرار شد تا هر بخش از داده‌ها یک‌بار به‌عنوان مجموعه اعتبارسنجی به کار رود. همچنین، برای کاهش اثر ترتیب نمونه‌ها، داده‌ها پیش از تقسیم‌بندی Fold ها به‌صورت تصادفی جابه‌جا شدند. در فرایند بهینه‌سازی، برای هر الگوریتم، ۱۰۰ ترکیب تصادفی از فضای هایپرپارامترها ارزیابی شد و عملکرد هر ترکیب با میانگین F1-Score حاصل از اعتبارسنجی متقاطع پنج‌بخشی سنجیده شد. درنهایت، ترکیبی از هایپرپارامترها که بیشترین مقدار میانگین F1-Score را به دست آورد، به‌عنوان تنظیمات بهینه مدل انتخاب شد.

درنهایت، ترکیبی از هایپرپارامترها که بیشترین مقدار میانگین F1-Score را تولید کرد، به‌عنوان تنظیمات بهینه مدل انتخاب شد. دامنه مقادیر بررسی‌شده برای هر هایپرپارامتر و مقدار بهینه انتخاب‌شده برای هر الگوریتم در جدول ۱ ارائه شده است.

5. Random Search
6. Cross-Validation
7. Grid Search
8. n_estimators
9. max_depth
10. min_samples_split
11. max_features
12. Overfitting
13. Precision
14. Recall

جدول ۱. دامنه هایپرپارامترهای مورداستفاده در جستجوی تصادفی و مقادیر بهینه انتخاب‌شده برای مدل‌های یادگیری ماشین

مدل	هایپرپارامترها	دامنه تغییرات	مقدار بهینه برای سنتینل ۲	مقدار بهینه برای لندست ۸
جنگل تصادفی	n_estimators	100، 200، 400، 600	200	400
	max_depth	None، 20، 40، 80، 120	40	40
	min_samples_split	2، 5، 10، 20	2	10
	min_samples_leaf	1، 2، 4، 8	1	1
	max_features	sqrt، log2، 0.3، 0.5، 0.8	0.5	0.3
	bootstrap	True، False	True	False
	class_weight	balanced، balanced_subsample، None	balanced_subsample	balanced
درخت تصمیم	max_depth	None، 10، 20، 40، 80	20	10
	min_samples_split	2، 5، 10، 20، 50	2	2
	min_samples_leaf	1، 2، 4، 8، 10، 20	1	4
	max_features	None، sqrt، log2، 0.5	None	0.5
	class_weight	balanced، None	balanced	None
ماشین بردار پشتیبان	C	0.1-1000	1	1000
	gamma	scale، auto، 0.0001 -0.1	0.01	0.1
	kernel	rbf	rbf	rbf
	class_weight	balanced، None	None	balanced
رگرسیون لجستیک	C	0.1-100	0.33	54.55
	penalty	l2	l2	l2
	solver	lbfgs، saga	saga	lbfgs
	class_weight	balanced، None	None	None
گرادیان بوستینگ هیستوگرام	max_iter	100، 200، 300	300	300
	learning_rate	0.01، 0.03، 0.05، 0.1	0.03	0.1
	max_depth	None، 5، 10، 20	10	None
	l2_regularization	0.0، 0.1، 1.0	1	0
	max_leaf_nodes	15، 31، 63	63	15
	min_samples_leaf	20، 50، 100	20	20

در مرحله نهایی، برای استفاده همزمان از قابلیت‌های مدل‌های پایه، از راهبرد فراترکیبی یا Meta-Blending استفاده شد. در این رویکرد، خروجی احتمالاتی مدل‌های پایه به‌عنوان ویژگی‌های جدید یا Meta-features در نظر گرفته شده و سپس یک مدل متا بر اساس این ویژگی‌ها آموزش داده می‌شود. هدف اصلی از این کار، بهره‌گیری از نقاط قوت مدل‌های مختلف، کاهش واریانس خطا و افزایش پایداری پیش‌بینی نهایی است. مطالعات اخیر در زمینه آتش‌سوزی نشان داده‌اند که مدل‌های Stacking Ensemble با ترکیب چند مدل پایه در یک پیش‌بین متا، می‌توانند دقت و قابلیت اعتماد خروجی را در نقشه‌برداری شدت یا حساسیت آتش‌سوزی بهبود دهند. در پژوهش حاضر، برای جلوگیری از نشتی داده، ابتدا پیش‌بینی‌های خارج از نمونه یا Out-of-Fold Predictions ازطریق اعتبارسنجی متقاطع K-fold تولید شد و سپس این پیش‌بینی‌ها برای آموزش مدل متا مورداستفاده قرار گرفتند. این ساختار مشابه منطق StackingClassifier در Scikit-learn است که در آن مدل نهایی با استفاده از پیش‌بینی‌های حاصل از cross_val_predict آموزش داده می‌شود و در صورت استفاده از مدل‌های از پیش آموزش‌دیده بر همان داده‌ها، خطر بیش‌برازش افزایش می‌یابد. در این تحقیق، رگرسیون لجستیک به‌عنوان مدل متا انتخاب شد، زیرا علاوه بر سادگی و تفسیرپذیری، امکان ترکیب وزن‌دار خروجی احتمالاتی مدل‌های پایه را فراهم می‌کند و در مطالعات جدید نیز به‌عنوان یکی از گزینه‌های مؤثر برای لایه متا در چهارچوب‌های فراترکیبی معرفی شده است (Nguyen Van & Lee, 2025).

در مرحله بعدی پیش‌بینی‌های خارج از نمونه هر پنج مدل پایه به‌عنوان ویژگی‌های ورودی برای مدل متا در نظر گرفته می‌شوند. این ویژگی‌ها یا اصطلاحاً meta features دانش جمعی مدل‌های پایه را در قالب احتمالات طبقه‌بندی کدگذاری

می‌کنند. در این تحقیق مدل رگرسیون لجستیک به‌عنوان مدل متا انتخاب شده است که به‌نوعی ارتباط خطی بین پیش‌بینی‌های مدل‌های پایه را مدل می‌نماید. تابع اتلاف این مدل به‌صورت زیر است (معادله ۱):

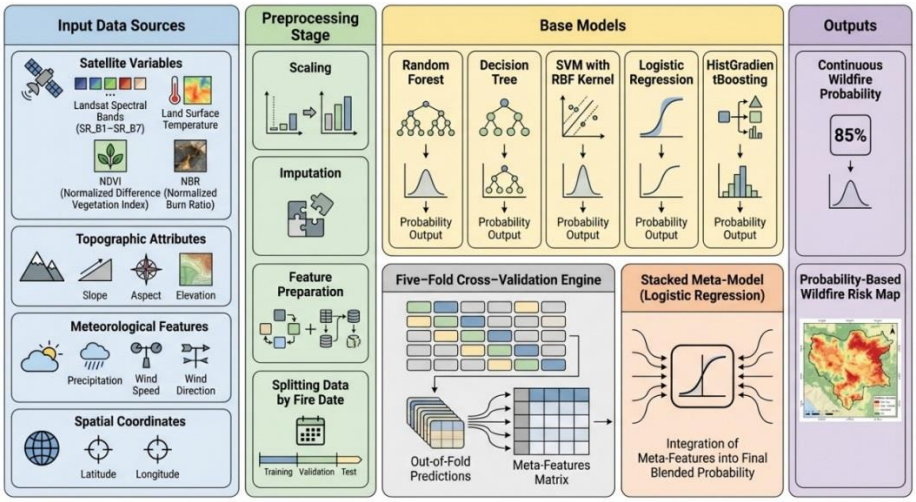
$$J(\omega) = \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n} log\left(1 + e^{-y_i(\omega^T x_i + b)}\right) \quad (1)$$

که در آن $\boldsymbol{\omega}$ بردار وزن، b بایاس، C پارامتر Regularization جهت برقراری تعادل و کنترل بیش‌برازش و کم‌برازش، $\boldsymbol{x_i}$ بردار ویژگی و $\boldsymbol{y_i}$ مقدار نقشه واقعیت برای نمونه i ام است. ازجمله مزایای این روش، ترکیب چندین مدل مستقل و به‌تبع کاهش واریانس خطای پیش‌بینی، بهبود عملکرد مدل و تعمیم‌پذیری آن، استفاده از قابلیت‌های متنوع مدل‌های پایه و کاهش احتمال وقوع بیش‌برازش مدل است.

پس از آموزش مدل متا، هر مدل پایه بر روی داده‌های آزمایشی پیش‌بینی انجام می‌دهد. این پیش‌بینی‌ها به‌عنوان ورودی به مدل متا داده می‌شود و مدل با ترکیب خطی این پیش‌بینی‌ها احتمال نهایی پتانسیل خطر آتش‌سوزی را محاسبه می‌نماید (معادله ۲).

$$P_{final} = \sigma\left(\sum_{j=1}^{M} \omega_j . P_j + b\right) \quad (2)$$

که در آن M تعداد مدل‌های پایه، ω_j وزن تخصیص‌یافته به هر مدل، P_j پیش‌بینی مدل و σ تابع سیگموید[15] است. معماری کلی مدل در شکل ۲ نشان داده شده است.



شکل ۲. فرایند انجام تحقیق و معماری کلی مدل فراترکیبی

در این تحقیق عملکرد مدل‌های پایه و مدل متا به کمک معیارهای کمی شامل صحت[16]، دقت، F1-Score، IoU و نتایج بصری شامل نقشه پتانسیل خطر آتش‌سوزی برای تاریخ مدنظر موردبررسی قرار گرفته‌اند که به‌تفصیل در بخش یافته‌های پژوهش موردبررسی قرار خواهد گرفت.

به‌منظور ارزیابی جامع‌تر عملکرد مدل‌ها، علاوه بر معیارهای Accuracy، Precision، Recall و F1-Score، از معیار سطح زیر منحنی[17] نیز استفاده شد. در این راستا، منحنی Precision-Recall ترسیم گردید. منحنی Precision-Recall عملکرد مدل را در تشخیص کلاس مثبت، به‌ویژه در داده‌های نامتوازن، بهتر نمایان می‌سازد. مقدار AUC بیانگر توان کلی مدل در تفکیک کلاس‌ها بوده و هرچه این مقدار به عدد ۱ نزدیک‌تر باشد، عملکرد مدل بهتر است.

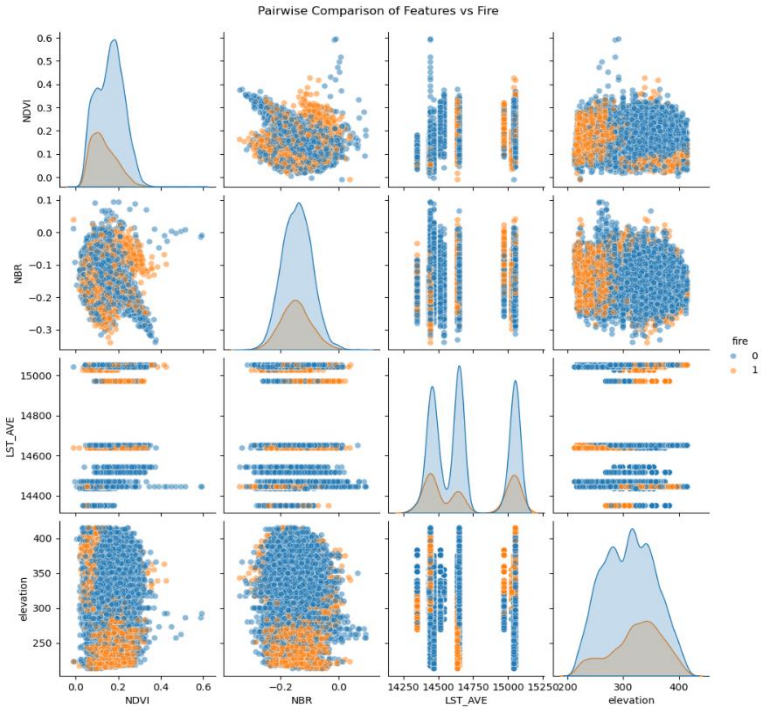
یافته‌های پژوهش

در این بخش، نتایج حاصل از پیاده‌سازی مدل‌های یادگیری ماشین ارائه و مورد تجزیه‌وتحلیل قرار می‌گیرد. ابتدا، یک تحلیل

15. Sigmoid
16. Accuracy
17. Area Under Curve

اکتشافی بر روی داده‌ها به‌منظور درک روابط بین متغیرهای کلیدی و پدیده آتش‌سوزی انجام شده و سپس عملکرد مدل‌های پایه و مدل ترکیبی نهایی به‌تفصیل ارزیابی می‌شود.

به‌منظور درک اولیه از توزیع داده‌ها و شناسایی الگوهای بالقوه، یک نمودار ماتریسی جفتی[18] برای چهار متغیر ورودی کلیدی شامل شاخص تفاضلی نرمال‌شده پوشش گیاهی، شاخص نرمال‌شده سوختگی، میانگین دمای سطح زمین و ارتفاع ترسیم گردید (شکل ۳). در این نمودار، نقاط داده بر اساس کلاس هدف (۰ برای عدم وقوع آتش و ۱ برای وقوع آتش) رنگ‌آمیزی شده‌اند. نمودارهای روی قطر اصلی، توزیع چگالی کرنل[19] هر متغیر را برای دو کلاس به تفکیک نشان می‌دهند و نمودارهای خارج از قطر، نمودارهای پراکندگی[20] بین جفت متغیرها هستند.



شکل ۳. نمودار ماتریسی مقایسه جفتی بین ویژگی‌های منتخب و کلاس هدف

نمودارهای توزیع روی قطر اصلی به‌وضوح نشان می‌دهند که مقادیر NDVI و NBR برای نقاطی که دچار آتش‌سوزی شده‌اند (کلاس ۱، نارنجی) به‌طور معناداری پایین‌تر از نقاط بدون آتش‌سوزی (کلاس ۰، آبی) است. منحنی توزیع برای کلاس آتش‌سوزی در هر دو شاخص به سمت مقادیر کمتر متمایل است. این امر ازنظر فیزیکی قابل‌انتظار است، زیرا مقادیر پایین‌تر این شاخص‌ها نشان‌دهنده پوشش گیاهی کمتر، خشک‌تر یا تحت تنش است که به‌عنوان سوخت مستعد برای آتش‌سوزی عمل می‌کند. نمودار پراکندگی بین NDVI و NBR (ردیف دوم، ستون اول) نیز این رابطه را تأیید می‌کند؛ نقاط نارنجی (آتش‌سوزی) به‌وضوح در گوشه پایین سمت چپ (مقادیر پایین هر دو شاخص) متمرکز شده‌اند، درحالی‌که نقاط آبی (عدم آتش‌سوزی) عمدتاً در ناحیه با مقادیر بالاتر قرار دارند. این جدایی‌پذیری بالا نشان می‌دهد که این دو شاخص، پیش‌بینی‌کننده‌های بسیار قدرتمندی برای پتانسیل خطر آتش‌سوزی هستند.

توزیع داده‌ها بر اساس ارتفاع نشان می‌دهد که رخدادهای آتش‌سوزی (نقاط نارنجی) به‌صورت کلی در ارتفاعات پایین‌تر (تقریباً بین ۲۵۰ تا ۳۵۰ متر) متمرکز شده‌اند. در مقابل، مناطق بدون آتش‌سوزی در گستره وسیع‌تری از ارتفاعات، به‌ویژه ارتفاعات بالاتر، پراکنده هستند. این الگو ممکن است به نوع پوشش گیاهی غالب در ارتفاعات مختلف یا افزایش فعالیت‌های

18. Pair Plot
19. KDE
20. Scatter Plots

انسانی در مناطق کم‌ارتفاع‌تر مرتبط باشد. همچنین مشاهدات توزیع LST با این فرض که دمای بالاتر سطح زمین منجر به خشک شدن سریع‌تر سوخت و افزایش احتمال اشتعال می‌شود، کاملاً سازگار است. نمودارهای پراکندگی متقابل، جداسازی بهتری را نشان می‌دهند. به‌عنوان‌مثال، در نمودار پراکندگی بین ارتفاع و NDVI (ردیف آخر، ستون اول)، یک خوشه متراکم از نقاط آتش‌سوزی در ناحیه‌ای با ارتفاع کم و NDVI پایین قابل‌مشاهده است. این نشان می‌دهد که ترکیب این دو ویژگی قدرت تفکیک‌پذیری بیشتری نسبت به هر یک به‌تنهایی دارد.

درمجموع، تحلیل اکتشافی داده‌ها نشان داد که بین متغیرهای ورودی و متغیر هدف (وقوع آتش) الگوهای مشخص و قابل‌تفکیکی وجود دارد. این جدایی‌پذیری مشهود در فضای ویژگی‌ها، مبنای محکمی برای موفقیت الگوریتم‌های یادگیری ماشین فراهم می‌آورد، زیرا این الگوریتم‌ها دقیقاً برای یادگیری چنین مرزهای تصمیم پیچیده و غیرخطی طراحی شده‌اند.

به‌منظور ارزیابی کمی و کیفی مدل‌های یادگیری ماشین توسعه داده‌شده، از پنج معیار استاندارد استفاده شد: صحت، دقت، Recall، F1-Score و IoU. این ارزیابی به‌صورت مجزا برای داده‌های حاصل از سنجنده سنتینل ۲ (با تفکیک مکانی بالاتر) و سنجنده لندست ۸ (با تفکیک مکانی پایین‌تر) انجام شد تا تأثیر وضوح مکانی بر عملکرد مدل‌ها نیز موردبررسی قرار گیرد. نتایج کامل این ارزیابی در جدول ۲ ارائه شده است.

همان‌طور که در جدول ۲ مشاهده می‌شود، مدل فراترکیبی به‌طور قاطع بهترین عملکرد را در میان تمام مدل‌ها بر روی داده‌های سنتینل ۲ از خود نشان داده است. این مدل با دستیابی به صحت ۹۴/۵۲ درصد، F1-Score برابر با ۸۷/۷۶ درصد و IoU معادل ۸۰/۱۲ درصد، برتری خود را در تمام معیارهای ارزیابی به اثبات رساند. مقدار بالای Recall در این مدل نشان‌دهنده توانایی استثنایی آن در شناسایی صحیح پیکسل‌های دارای آتش‌سوزی واقعی است که برای کاربردهای مدیریت بحران از اهمیت ویژه‌ای برخوردار است.

در میان مدل‌های پایه، جنگل تصادفی با امتیاز F1 برابر با ۸۴/۱۲ درصد و گرادیان بوستینگ هیستوگرام با امتیاز ۸۲/۳۴ درصد، بهترین عملکرد را داشتند. این نتایج نشان می‌دهد که مدل‌های مبتنی بر Ensemble به دلیل توانایی در ترکیب پیش‌بینی‌کننده‌های ضعیف‌تر و کاهش واریانس، برای این مسئله بسیار مناسب هستند. در مقابل، مدل‌هایی مانند رگرسیون لجستیک و ماشین بردار پشتیبان عملکرد ضعیف‌تری داشتند که این امر به ماهیت روابط پیچیده و غیرخطی حاکم بر داده‌های پتانسیل خطر حریق نسبت داده می‌شود. موفقیت چشمگیر مدل فراترکیبی نشان می‌دهد که ترکیب هوشمندانه پیش‌بینی‌های مدل‌های متنوع (ازجمله جنگل تصادفی و گرادیان بوستینگ هیستوگرام) ازطریق یک فرایادگیرنده[21]، قادر است نقاط ضعف هر مدل پایه را پوشش داده و به یک پیش‌بینی نهایی قوی‌تر و پایدارتر دست یابد.

الگوی مشابهی در نتایج حاصل از داده‌های لندست ۸ نیز مشاهده می‌شود. برای این سنجنده نیز، مدل فراترکیبی با کسب F1-Score برابر با ۸۶/۱۵ درصد و IoU معادل ۷۹/۰۵ درصد، بهترین عملکرد را در این گروه به ثبت رساند و برتری خود را نسبت به تمامی مدل‌های پایه حفظ کرد. این مدل توانست به صحت ۹۳/۵۱ درصد دست یابد که نشان‌دهنده پایداری و کارایی بالای رویکرد فراترکیبی حتی با داده‌هایی با وضوح مکانی کمتر است.

جدول۲. بررسی عملکرد مدل‌های پایه و مدل فراترکیبی به کمک معیارهای ارزیابی کمی

سنجنده	مدل	صحت	دقت	F1-Score	Recall	IoU
سنتینل ۲	جنگل تصادفی	۹۱/۴۸	۷۸/۰۵	۸۴/۱۲	۹۰/۰۳	۷۴/۲۲
	درخت تصمیم	۸۹/۶۳	۷۵/۱۲	۸۱/۰۵	۸۷/۱۵	۷۰/۱۸
	ماشین بردار پشتیبان	۸۹/۰۷	۷۴/۵۵	۸۰/۵۱	۸۶/۱۲	۶۹/۰۳
	رگرسیون لجستیک	۸۸/۵۴	۷۳/۲۹	۷۹/۰۸	۸۵/۲۲	۶۸/۱۰
	گرادیان بوستینگ هیستوگرام	۹۰/۵۷	۷۶/۴۷	۸۲/۳۴	۸۸/۰۲	۷۲/۱۵
	مدل فراترکیبی	۹۴/۵۲	۸۲/۱۳	۸۷/۷۶	۹۳/۰۲	۸۰/۱۲

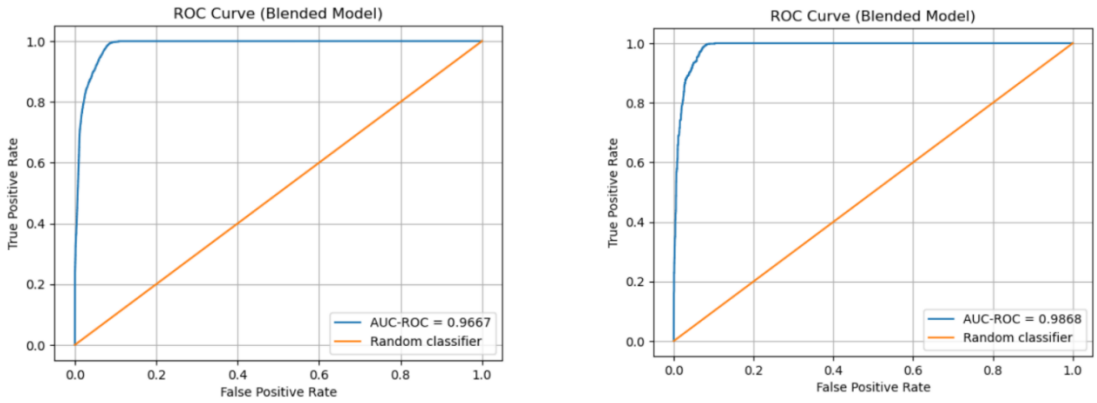
21. Meta-learner

سنجنده	مدل	صحت	دقت	F1-Score	Recall	IoU
لندست ۸	جنگل تصادفی	۹۱/۰۳	۷۷/۱۸	۸۳/۰۱	۸۵/۱۲	۷۴/۱۰
	درخت تصمیم	۸۸/۵۴	۷۴/۰۸	۸۰/۰۲	۸۲/۱۵	۷۰/۲۲
	ماشین بردار پشتیبان	۸۸/۰۲	۷۳/۵۵	۷۹/۵۱	۸۱/۱۲	۶۹/۰۸
	رگرسیون لجستیک	۸۷/۵۴	۷۲/۱۲	۷۸/۰۳	۸۰/۲۲	۶۸/۰۵
	گرادیان بوستینگ هیستوگرام	۹۰/۱۲	۷۶/۰۲	۸۲/۰۵	۸۴/۲۱	۷۲/۱۲
	مدل فراترکیبی	۹۳/۵۱	۸۰/۲۴	۸۶/۱۵	۸۷/۳۳	۷۹/۰۵

مقایسه مستقیم نتایج بین دو سنجنده، بینش ارزشمندی را در مورد تاثیر وضوح مکانی داده‌های ماهواره‌ای بر دقت مدل‌سازی پتانسیل خطر آتش‌سوزی فراهم می‌کند. مدل فراترکیبی آموزش‌دیده بر روی داده‌های سنتینل ۲ در تمام پنج معیار ارزیابی، عملکرد بهتری نسبت به مدل مشابه آموزش‌دیده بر روی داده‌های لندست ۸ داشت.

به‌طور مشخص F1-Score از ۸۶/۱۵ درصد در لندست ۸ به ۸۷/۷۶ درصد در سنتینل ۲ (بهبود ۱/۶۱ درصدی) و معیار IoU از ۷۹/۰۵ درصد به ۸۰/۱۲ درصد (بهبود ۱/۰۷ درصدی) افزایش یافت. این برتری به‌احتمال‌زیاد ناشی از تفکیک مکانی بالاتر باندهای سنجنده سنتینل ۲ (۱۰ متر) نسبت به لندست ۸ (۳۰ متر) است. وضوح بالاتر به مدل اجازه می‌دهد تا الگوهای مکانی ظریف‌تر و تغییرات محلی در پوشش گیاهی، رطوبت و ساختار زمین را که در مقیاس‌های کوچک‌تر رخ می‌دهند و عوامل مهمی در شروع و گسترش آتش هستند، بهتر شناسایی کند. این نتایج به‌صورت تجربی تأیید می‌کند که سرمایه‌گذاری بر روی داده‌های ماهواره‌ای با وضوح مکانی بالاتر می‌تواند منجر به تولید نقشه‌های پتانسیل خطر آتش‌سوزی با دقت و قابلیت اطمینان بیشتری شود.

به‌منظور ارزیابی جامع‌تر عملکرد مدل فراترکیبی، علاوه بر معیارهای صحت، دقت، Recall، F1-Score وIoU، از منحنی ROC نیز استفاده شد (شکل ۴). منحنی ROC توانایی مدل را در تفکیک کلاس‌های مثبت و منفی در آستانه‌های تصمیم‌گیری مختلف نشان می‌دهد و مقدار AUC-ROC به‌عنوان شاخصی برای سنجش قدرت تفکیک کلی مدل مورداستفاده قرار می‌گیرد.



لندست ۸ سنتینل ۲

شکل ٤. منحنی ROC مدل فراترکیبی برای داده‌های لندست ۸ و سنتینل ۲

نتایج نشان داد که مدل فراترکیبی در هر دو رویکرد داده‌ای عملکرد بسیار مطلوبی دارد. در رویکرد مبتنی بر داده‌های لندست ۸، مقدار AUC-ROC برابر با ۰/۹۶۶۷ به دست آمد. در مقابل، در رویکرد مبتنی بر داده‌های سنتینل ۲، مقدار AUC-ROC برابر با ۰/۹۸۶۸ حاصل شد. مقایسه این نتایج نشان می‌دهد که مدل فراترکیبی در هر دو سناریو توان تفکیک بالایی میان مناطق مستعد آتش‌سوزی دارد؛ بااین‌حال، استفاده از داده‌های سنتینل ۲ به دلیل تفکیک مکانی بالاتر، موجب بهبود عملکرد مدل در معیار AUC-ROC شده است.

جهت شفاف‌سازی عملکرد الگوریتم‌ها در مرحله بهینه‌سازی و مرحله ارزیابی نهایی، نتایج مدل‌ها در دو سطح گزارش شد. نخست، مقدار Best CV Score برای هر الگوریتم پس از انتخاب بهترین ترکیب هایپرپارامترها و بر اساس اعتبارسنجی متقاطع

پنج‌بخشی محاسبه گردید. این مقدار نشان‌دهنده عملکرد مدل در مرحله تنظیم پارامترها است. سپس، مدل‌های بهینه‌شده روی مجموعه داده ارزیابی مستقل آزمون شدند و مقدار Accuracy حاصل از داده‌های ارزیابی گزارش شد (جدول ۳). مقایسه این دو مقدار نشان می‌دهد که اگرچه الگوریتم‌ها در مرحله اعتبارسنجی متقاطع عملکرد بسیار بالایی داشته‌اند، عملکرد آن‌ها روی داده ارزیابی مستقل واقع‌بینانه‌تر بوده و بیانگر توان تعمیم مدل به داده‌های مشاهده‌نشده است.

جدول ۳. مقایسه Best CV Score مدل‌های بهینه‌شده و دقت حاصل از داده‌های ارزیابی

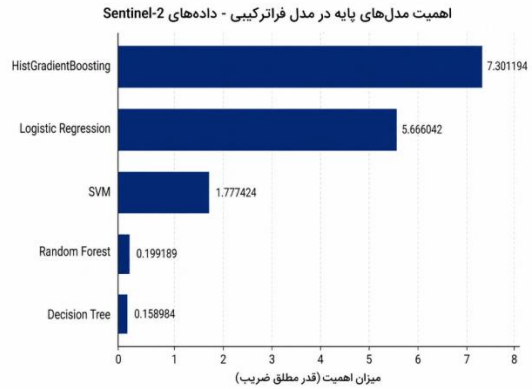
سنجنده	الگوریتم	Best CV (%)	Accuracy (%) داده آموزشی
لندست ۸	جنگل تصادفی	۹۹,۱۵	۹۱,۳
	درخت تصمیم	۹۵,۷۰	۸۸,۵۴
	SVM	۹۸,۲۶	۸۸,۰۲
	رگرسیون لجستیک	۹۶,۴۲	۸۷,۵۴
	گرادیان بوستینگ هیستوگرام	۹۹,۱۹	۹۰,۱۲
سنتینل ۲	جنگل تصادفی	۹۹,۶۳	۹۱,۴۸
	درخت تصمیم	۹۹,۴۴	۸۹,۶۳
	SVM	۹۸,۴۴	۸۹,۰۷
	رگرسیون لجستیک	۹۶,۴۶	۸۸,۵۴
	گرادیان بوستینگ هیستوگرام	۹۹,۵۶	۹۰,۵۷

همان‌گونه که در جدول ۳ مشاهده می‌شود، در هر دو رویکرد لندست ۸ و سنتینل ۲، مدل‌های جنگل تصادفی و گرادیان بوستینگ هیستوگرام بالاترین مقادیر Best CV Score را به خود اختصاص داده‌اند. در داده‌های لندست ۸، گرادیان بوستینگ هیستوگرام با مقدار ۹۹/۱۹ درصد و جنگل تصادفی با مقدار ۹۹/۱۵ درصد بهترین عملکرد را در مرحله اعتبارسنجی متقاطع داشته‌اند. در داده‌های سنتینل ۲ نیز جنگل تصادفی با مقدار ۹۹/۶۳ درصد و گرادیان بوستینگ هیستوگرام با مقدار ۹۹/۵۶ درصد عملکرد برتری نشان داده‌اند. بااین‌حال، در مرحله ارزیابی مستقل، مقادیر Accuracy کاهش یافته است که امری طبیعی بوده و نشان می‌دهد داده‌های آزمون، معیار واقع‌بینانه‌تری برای سنجش قابلیت تعمیم مدل‌ها ارائه می‌دهند.

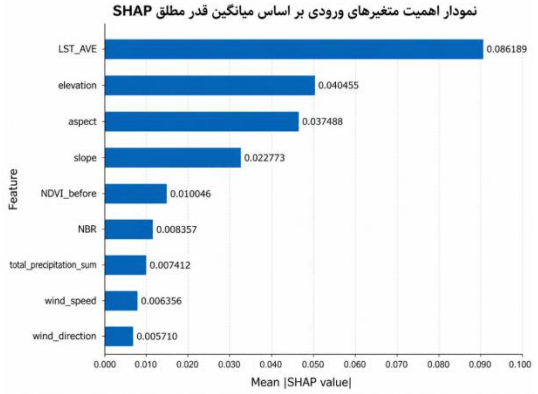
تحلیل اهمیت شاخه‌های مدل فراترکیبی نشان داد که مشارکت مدل‌های پایه در تصمیم‌گیری نهایی یکسان نبوده است (شکل ۵). در مدل دارای بالاترین دقت، یعنی رویکرد مبتنی بر داده سنتینل ۲، شاخه گرادیان بوستینگ هیستوگرام با بیشترین ضریب، مهم‌ترین نقش را در پیش‌بینی پتانسیل خطر آتش‌سوزی ایفا کرده است. پس‌ازآن، شاخه رگرسیون لجستیک نیز سهم قابل‌توجهی در خروجی نهایی مدل داشته است. سهم SVM در سطح متوسط ارزیابی شد، درحالی‌که شاخه‌های جنگل تصادفی و درخت تصمیم اثر نسبتاً محدودی در ساختار نهایی مدل نشان دادند. این موضوع بیانگر آن است که مدل فراترکیبی به‌صورت هوشمندانه‌تر بر مدل‌هایی تکیه کرده است که اطلاعات مؤثرتری برای تفکیک مناطق مستعد و غیرمستعد آتش‌سوزی فراهم کرده‌اند.

این نتایج نشان می‌دهد که در ساختار مدل فراترکیبی مبتنی بر داده‌های سنتینل ۲، بیشترین قدرت پیش‌بینی نهایی از سوی شاخه‌های گرادیان بوستینگ هیستوگرام و رگرسیون لجستیک تأمین شده است و سایر شاخه‌ها نقش تکمیلی داشته‌اند.

علاوه بر بررسی اهمیت شاخه‌های مدل فراترکیبی، به‌منظور افزایش تفسیرپذیری مدل، اهمیت متغیرهای ورودی نیز با استفاده از روش SHAP (Ponce-Bobadilla et al., 2024) محاسبه شد که نتایج حاصل در شکل۶ نشان داده شده است. نتایج نشان داد که در میان متغیرهای ورودی، LST بیشترین اهمیت را داشته و پس‌ازآن متغیرهای ارتفاع، جهت شیب و شیب در رتبه‌های بعدی قرار گرفته‌اند. در میان شاخص‌های طیفی نیز، NDVI و NBR نقش موثری در پیش‌بینی خطر آتش‌سوزی داشته‌اند، درحالی‌که متغیرهای جوی شامل مجموع بارش، سرعت و جهت باد اهمیت کمتری نسبت به متغیرهای توپوگرافی و حرارتی نشان داده‌اند. این نتایج بیانگر آن است که شرایط حرارتی سطح زمین و ویژگی‌های توپوگرافی مهم‌ترین عوامل کنترل‌کننده پتانسیل خطر آتش‌سوزی در محدوده مطالعه بوده‌اند.

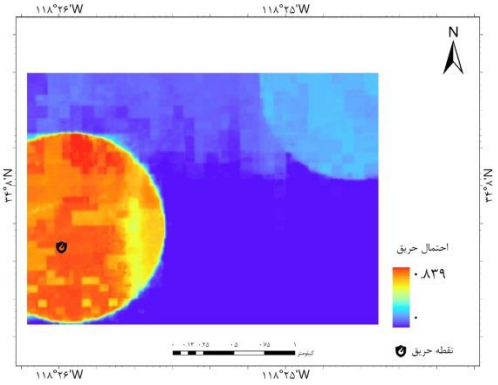


شکل ۵. اهمیت نسبی شاخه‌های مدل فراترکیبی در رویکرد مبتنی بر داده سنتینل ۲

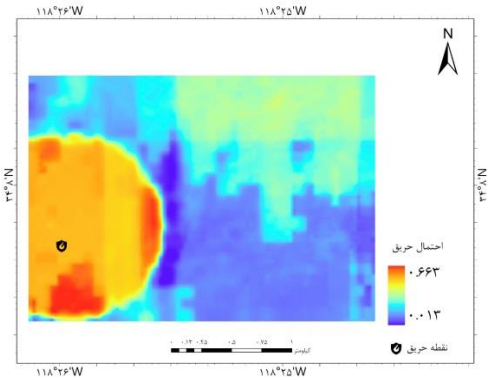


شکل ٦. میزان اهمیت پارامترهای ورودی شامل شاخص‌های طیفی و پارامترهای جوی در رویکرد مبتنی بر داده سنتینل ۲

فراتر از ارزیابی‌های کمی، هدف نهایی این پژوهش، تولید نقشه‌های پتانسیل خطر آتش‌سوزی با قابلیت تفسیر مکانی وقوع آتش‌سوزی است. برای این منظور، مدل فراترکیبی که بهترین عملکرد را در میان مدل‌های موجود داشت، جهت پیش‌بینی احتمال وقوع آتش‌سوزی برای هر پیکسل در منطقه موردمطالعه به کار گرفته شد. این احتمالات، که مقادیری پیوسته بین ۰ (پتانسیل خطر بسیار پایین) و ۱ (پتانسیل خطر بسیار بالا) هستند، به‌عنوان مبنای تولید نقشه‌های نهایی پتانسیل خطر مورداستفاده قرار گرفتند. نتایج حاصل در شکل‌های ۷ و ۸ ارائه شده است.



شکل ۷. نقشه تخمین پتانسیل خطر آتش‌سوزی با استفاده از مدل فراترکیبی و داده‌های سنجنده سنتینل ۲



شکل ۸. نقشه تخمین پتانسیل خطر آتش‌سوزی با استفاده از مدل فراترکیبی و داده‌های سنجنده لندست ۸

نقشه حاصل از داده‌های سنتینل ۲ (شکل ۷)، نمایشی با جزئیات بیشتر و انسجام مکانی بالاتر از پتانسیل خطر آتش‌سوزی ارائه می‌دهد. الگوهای پتانسیل خطر در این نقشه دارای مرزهای مشخص‌تری هستند و می‌توان مناطق کوچک و مجزا با پتانسیل خطر بالا را به‌خوبی تشخیص داد. این جزئیات دقیق، جهت برنامه‌ریزی‌های عملیاتی مدیریت بحران، مانند تخصیص منابع اطفای آتش، بسیار ارزشمند است. برتری بصری نقشه سنتینل ۲ مستقیماً به توان تفکیک مکانی بالاتر داده‌های آن نسبت داده می‌شود. وضوح مکانی ۱۰ متری سنتینل ۲ به مدل اجازه می‌دهد تا تغییرات محلی در پارامترهای کلیدی مانند *NDVI* و *NBR* را با دقت بیشتری ثبت کند. در مقابل، نقشه حاصل از داده‌های لندست ۸ (شکل ۸) به دلیل وضوح مکانی پایین‌تر (۳۰ متر)، ظاهری هموارتر و تا حدی بلوکی دارد. در این نقشه، مناطق پتانسیل خطر به‌صورت لکه‌های بزرگ‌تر و با مرزهای نامشخص‌تر نمایش داده می‌شوند. این اثر در پیکسل‌های بزرگ‌تر لندست ۸ باعث می‌شود که برخی از الگوهای مکانی که ممکن است برای شناسایی نقاط بحرانی کلیدی باشند، از بین بروند.

برتری بصری نقشه سنتینل ۲، انطباق کاملی با برتری آن در معیارهای کمی مانند IoU و F1-Score دارد. IoU بالاتر به این معناست که مدل مبتنی بر سنتینل ۲ هم‌پوشانی بسیار بهتری بین مناطق پیش‌بینی‌شده به‌عنوان پرخطر و مناطق واقع در اطراف نقطه آتش‌سوزی دارد. این دقت بالاتر در مرزبندی، خود را به‌صورت یک نقشه خروجی واضح‌تر و کاربردی‌تر نشان می‌دهد. درنتیجه، تحلیل مکانی نشان می‌دهد که درحالی‌که هر دو سنجنده قادر به تولید نقشه‌های پتانسیل خطر معنادار هستند، داده‌های با تفکیک مکانی بالاتر مانند سنتینل ۲، منجر به تولید محصولاتی با دقت عملیاتی و قابلیت اطمینان بسیار بیشتری می‌شوند که برای مدیریت مؤثر سوانح طبیعی مانند آتش‌سوزی‌های گسترده، امری حیاتی است.

بحث

هدف از این پژوهش، توسعه یک چهارچوب یکپارچه مبتنی بر الگوریتم‌های یادگیری ماشین جهت نقشه‌برداری دقیق پتانسیل خطر آتش‌سوزی جنگل و ارزیابی تأثیر قدرت تفکیک مکانی داده‌های ماهواره‌ای (سنتینل ۲ در برابر لندست ۸) بر عملکرد مدل‌های پیش‌بینی بود. نتایج حاصل از تحلیل‌های اکتشافی و ارزیابی‌های کمی نشان داد که ادغام داده‌های چندطیفی با مدل‌های یادگیری ماشین فراترکیبی، رویکردی بسیار کارآمد برای غلبه بر پیچیدگی‌های پدیده حریق جنگل است.

در این پژوهش تحلیل اکتشافی داده‌ها تأیید کرد که متغیرهای طیفی و توپوگرافیک دارای الگوهای متمایز و قابل‌تفکیکی در مناطق دچار حریق نسبت به مناطق ایمن هستند. پایین بودن معنادار مقادیر شاخص‌های NDVI و NBR در کلاس هدف (وقوع آتش‌سوزی) ازنظر اکولوژیک کاملا قابل توجیه است؛ چراکه مقادیر پایین این شاخص‌ها نشان‌گر کاهش شادابی پوشش گیاهی، خشکی و وجود تنش رطوبتی در گیاه است که بستر ایدئالی برای وقوع آتش‌سوزی فراهم می‌کند. علاوه بر این، تمرکز نقاط آتش‌سوزی در بازه ارتفاعی خاص (۲۵۰ تا ۳۵۰ متر) و همبستگی آن با مقادیر پایین NDVI نشان داد که آتش‌سوزی‌ها صرفاً تابع یک عامل نیستند، بلکه برآیند تعاملات پیچیده میان توپوگرافی و وضعیت پوشش گیاهی‌اند. این جدایی‌پذیری واضح در فضای ویژگی‌ها دلیل اصلی موفقیت الگوریتم‌های یادگیری ماشین در این مطالعه است، زیرا این مدل‌ها ظرفیت بالایی در کشف مرزهای تصمیم غیرخطی و چندبعدی دارند.

در ارزیابی مدل‌های پایه، الگوریتم‌های مبتنی بر درخت و یادگیری جمعی مانند جنگل تصادفی و گرادیان بوستینگ هیستوگرام نسبت به مدل‌های خطی و فاصله محور (مانند رگرسیون لجستیک و ماشین بردار پشتیبان) عملکرد بسیار بهتری نشان دادند. این امر ناشی از ماهیت غیرخطی روابط میان متغیرهای محیطی و پتانسیل خطر حریق است که مدل‌های مبتنی بر درخت بهتر از عهده مدل‌سازی آن برمی‌آیند. بااین‌حال، مهم‌ترین یافته تحلیلی این بخش، برتری قاطع مدل فراترکیبی بود. این مدل با ترکیب نقاط قوت پیش‌بینی‌کننده‌های پایه و کاهش واریانس خطای آن‌ها، روی داده‌های سنتینل ۲ به صحت معادل ۹۴/۵۲ درصد، امتیاز و IoU معادل ۸۰/۱۲ درصد دست یافت. مقدار بسیار بالای معیار Recall در این مدل از منظر مدیریت بحران اهمیتی حیاتی دارد؛ چراکه نشان‌دهنده کمترین میزان خطای نوع دوم[22] است و تضمین می‌کند که اکثر پیکسل‌های مستعد حریق در طبیعت، توسط مدل به‌درستی و با موفقیت شناسایی شده‌اند. پایداری این مدل فراترکیبی روی داده‌های لندست ۸ نیز اثبات می‌کند که معماری پیشنهادی، وابستگی مطلقی به یک سنجنده خاص ندارد و از استحکام[23] بالایی برخوردار است.

مقایسه مستقیم خروجی دو سنجنده، فرضیه تأثیر مثبت توان تفکیک مکانی بر مدل‌سازی پتانسیل خطر حریق را به‌صورت کمی و بصری اثبات کرد. ارتقای تفکیک مکانی از ۳۰ متر (لندست ۸) به ۱۰ متر (سنتینل ۲)، منجر به بهبود ۱/۶۱ درصدی در F1-Score و ۱/۰۷ درصدی در شاخص IoU در مدل فراترکیبی شد. از منظر تحلیل فضایی، پیکسل‌های ۱۰ متری سنتینل ۲ به مدل اجازه دادند تا تغییرات محلی بسیار ظریف در رطوبت، ساختار زمین و پوشش گیاهی را که به‌عنوان کانون‌های اولیه[24]

22. False Negative
23. Robustness
24. Hotspots

شروع حریق عمل می‌کنند، با دقت بالایی شناسایی کند. در مقابل، خروجی لندست ۸ به دلیل پدیده میانگین‌گیری مکانی[۲۵] در پیکسل‌های ۳۰ متری، لکه‌های حریق را به شکل بلوکی‌تر، با مرزهای محوتر و یکپارچه‌تر نمایش داد. تطابق برتری بصری نقشه‌های سنتینل ۲ با شاخص IoU بالاتر، نشان می‌دهد که هم‌پوشانی هندسیِ پیش‌بینی‌های مدل با واقعیت زمینی در داده‌های با رزولوشن بالا به‌مراتب دقیق‌تر است. این یافته‌ها در تضاد با برخی رویکردهای کلاسیک است که از داده‌های مقیاس منطقه‌ای (مانند MODIS) استفاده می‌کنند و بر لزوم استفاده از سنجنده‌های وضوح بالا در مدیریت عملیاتی حریق تأکید دارد.

از منظر اثر تفکیک مکانی، یافته‌های این پژوهش با مطالعه توکلی پیرعلی‌لو و همکاران (۲۰۲۲) نیز همسو است. آن‌ها اثر داده‌های سنجش‌ازدور با تفکیک‌های مکانی متفاوت را در پیش‌بینی حساسیت آتش‌سوزی بررسی کردند و داده‌های لندست ۸ و SRTM را به‌عنوان داده‌های با توان تفکیک مکانی پایین‌تر و داده‌های سنتینل ۲ و ALOS را به‌عنوان داده‌های با تفکیک بالاتر ارزیابی نمودند. نتایج آن مطالعه نشان داد که استفاده از داده‌های با تفکیک مکانی بهتر می‌تواند به بهبود نقشه‌های حساسیت آتش‌سوزی کمک کند و ترکیب پیش‌بینی‌ها نیز عملکرد نهایی را ارتقا دهد. نتایج پژوهش حاضر نیز این موضوع را تأیید می‌کند؛ زیرا سنتینل ۲ نه‌تنها در معیارهای کمی عملکرد بهتری نسبت به لندست ۸ داشته، بلکه نقشه خروجی آن نیز مرزبندی دقیق‌تر و جزئیات مکانی بیشتری در شناسایی کانون‌های پرخطر ارائه کرده است.

مقایسه نتایج این پژوهش با مطالعاتی که از مدل‌های منفرد استفاده کرده‌اند نیز نشان می‌دهد که رویکرد فراترکیبی می‌تواند بر محدودیت مدل‌های تکی غلبه کند. برای مثال، در مطالعه‌ای در سال ۲۰۲۴، عملکرد رگرسیون لجستیک و جنگل تصادفی برای تهیه نقشه حساسیت آتش‌سوزی در دو منطقه با ویژگی‌های فیزیکی متفاوت در آمریکا و کانادا مقایسه شد. نتایج آن مطالعه نشان داد که جنگل تصادفی در بسیاری از رویکردها عملکرد بهتری از رگرسیون لجستیک دارد، اما هنگام انتقال مدل از یک منطقه به منطقه دیگر، عملکرد مدل کاهش پیدا می‌کند. این موضوع نشان می‌دهد که تعمیم‌پذیری مکانی همچنان یکی از چالش‌های اصلی مدل‌های یادگیری ماشین در پیش‌بینی آتش‌سوزی است. در پژوهش حاضر، اگرچه اعتبارسنجی در محدوده مطالعه انجام شده است، استفاده از مدل فراترکیبی می‌تواند تا حدی واریانس پیش‌بینی را کاهش دهد و نسبت به مدل‌های منفرد پایداری بیشتری فراهم کند (Moghim & Mehrabi, 2024).

در مقایسه با مطالعاتی که از مدل‌های پایه مانند RF، SVM، شبکه عصبی و XGBoost استفاده کرده‌اند، نتایج پژوهش حاضر نشان می‌دهد که ترکیب چند الگوریتم می‌تواند در شرایطی که متغیرهای ورودی از جنس‌های مختلف هستند، عملکرد مناسب‌تری ایجاد کند. مطالعه بوزرا و همکاران در سال ۲۰۲۵ (Bouzeraa et al., 2025) نیز یک چهارچوب یادگیری ماشین برای نقشه‌برداری حساسیت آتش‌سوزی ارائه کرد و با استفاده از عوامل اقلیمی، توپوگرافی، پوشش گیاهی و فعالیت انسانی، مدل‌هایی مانند RF، SVM، NN و XGBoost را مقایسه نمود. اهمیت این دسته از مطالعات در تأکید بر چندعاملی بودن پدیده آتش‌سوزی است؛ بااین‌حال، بسیاری از آن‌ها همچنان بر مقایسه مدل‌های منفرد متمرکز هستند. پژوهش حاضر با استفاده از مدل فراترکیبی، یک گام فراتر رفته و تلاش کرده است خروجی چند مدل پایه را در قالب یک چهارچوب واحد و قابل‌تعمیم‌تر ادغام کند.

تولید نقشه‌های پتانسیل خطر با تفکیک‌پذیری مکانی و انسجام بالا، یک ابزار تصمیم‌یار قدرتمند برای مدیران منابع طبیعی است. نقشه‌های مستخرج از داده‌های سنتینل ۲ و مدل فراترکیبی در این مطالعه، به دلیل توانایی در مرزبندی دقیق مناطق پرخطر و مجزا کردن لکه‌های کوچک، قابلیت استفاده مستقیم در برنامه‌ریزی‌های پیشگیرانه نظیر تخصیص بهینه تجهیزات اطفای حریق و پایش مناطق حساس را دارا هستند. علی‌رغم نتایج امیدوارکننده، پتانسیل خطر آتش‌سوزی پدیده‌ای پویاست. پیشنهاد می‌شود در مطالعات آتی، به‌منظور ارتقای این چهارچوب یکپارچه، داده‌های هواشناسی با توان تفکیک زمانی بالا (نظیر تغییرات روزانه دما، رطوبت و باد) با متغیرهای مکانی فعلی ادغام شوند تا نقشه‌های پتانسیل خطر از حالت استاتیک به حالت

25. Spatial Averaging

دینامیک (پویا) تبدیل گردند (Xu et al., 2025). همچنین ارزیابی این مدل فراترکیبی در سایر اکوسیستم‌های جنگلی با رویکردهای آتش‌سوزی متفاوت، می‌تواند گام مؤثری در سنجش میزان تعمیم‌پذیری[۲۶] آن باشد.

نتیجه‌گیری

این پژوهش با هدف توسعه یک سیستم هوشمند جهت پیش‌بینی مناطق مستعد حریق و ارزیابی نقش قدرت تفکیک مکانی در ارتقای دقت نقشه‌های پتانسیل خطر، گام‌های موثری در جهت تلفیق الگوریتم‌های پیشرفته یادگیری ماشین با داده‌های سنجش‌ازدور برداشت. یافته‌های این مطالعه به روشنی نشان داد که پیچیدگی‌های حاکم بر وقوع آتش‌سوزی در جنگل‌ها، که از تعاملات غیرخطی میان پوشش گیاهی، وضعیت توپوگرافیک و شرایط حرارتی سطح زمین ناشی می‌شود، با استفاده از مدل‌های یادگیری ماشین فراترکیبی به بهترین شکل قابل مدل‌سازی است. برتری قاطع مدل فراترکیبی نسبت به الگوریتم‌های پایه نظیر ماشین بردار پشتیبان یا رگرسیون لجستیک، ثابت کرد که ترکیب استراتژیک چندین پیش‌بینی‌کننده می‌تواند واریانس خطا را کاهش داده و سوگیری‌های مدل‌های منفرد را به حداقل برساند.

یکی از دستاوردهای محوری این تحقیق، اثبات برتری معنادار سنجنده سنتینل ۲ نسبت به لندست ۸ در تهیه نقشه‌های پتانسیل خطر حریق بود. اگرچه هر دو سنجنده عملکردی فراتر از ۹۳ درصد در صحت کلی داشتند، اما ارتقای رزولوشن مکانی از ۳۰ متر به ۱۰ متر در سنتینل ۲ منجر به بهبود نتایج حاصل شد. این بهبود هم در معیارهای کمی و هم در تحلیل‌های بصری و مکانی نیز به‌وضوح مشاهده گردید؛ جایی که نقشه‌های مستخرج از سنتینل ۲ توانستند مرزبندی‌های دقیق‌تر و کانون‌های پتانسیل خطر جزئی‌تری را شناسایی کنند که در داده‌های لندست ۸ به دلیل پدیده میانگین‌گیری مکانی و بزرگ بودن ابعاد پیکسل، به‌صورت لکه‌های بلوکی و هموار نمایش داده می‌شدند. این موضوع تأکید می‌کند که برای مدیریت عملیاتی در سطوح محلی، استفاده از سنجنده‌های با قدرت تفکیک بالا نه یک انتخاب، بلکه یک ضرورت فنی برای شناسایی کانون‌های دقیق آتش‌سوزی است.

از سوی دیگر، نتایج این مطالعه بر اهمیت معیار Recall در مدیریت بحران تاکید ورزید. دستیابی مدل فراترکیبی به مقدار Recall برابر ۹۳/۰۲ درصد به این معناست که سیستم پیشنهادی توانسته است اکثریت قریب به وقوع مناطق مستعد آتش‌سوزی را به‌درستی شناسایی کند و خطای نوع دوم (عدم‌تشخیص منطقه پرخطر) را به حداقل برساند. این ویژگی برای سازمان‌های متولی منابع طبیعی بسیار حیاتی است؛ زیرا هزینه نادیده گرفتن یک منطقه مستعد آتش‌سوزی به‌مراتب سنگین‌تر از هشدارهای اشتباه در مناطق ایمن است. تحلیل‌های اکتشافی نیز نشان داد که وابستگی شدید وقوع حریق به مقادیر پایین شاخص‌های NDVI و NBR در بازه‌های ارتفاعی خاص، می‌تواند به‌عنوان یک شاخص هشداردهنده زودهنگام در پایش‌های ماهواره‌ای مستمر مورداستفاده قرار گیرد.

چهارچوب ارائه‌شده در این پژوهش با استفاده از داده‌های رایگان و در دسترس ماهواره‌ای، ابزاری قدرتمند، دقیق و اقتصادی را برای پیش‌بینی خطر آتش‌سوزی فراهم کرده است. باوجود موفقیت مدل در شرایط استاتیک، پیشنهاد می‌شود در مطالعات آتی، مؤلفه‌های پویای هواشناسی نظیر سرعت باد و رطوبت لحظه‌ای اتمسفر به سیستم اضافه شوند تا قدرت پیش‌بینی مدل در مقیاس‌های زمانی کوتاه‌مدت (روزانه) نیز ارزیابی گردد. همچنین، استفاده از تکنیک‌های یادگیری عمیق و شبکه‌های عصبی کانولوشنی برای استخراج الگوهای بافتی از تصاویر ماهواره‌ای می‌تواند افق‌های جدیدی را در ارتقای دقت نقشه‌برداری پتانسیل خطر آتش‌سوزی در اکوسیستم‌های متنوع باز نماید. این مطالعه ثابت کرد که گذار از روش‌های کلاسیک به سمت مدل‌های هوشمند و داده‌محور، مسیری اجتناب‌ناپذیر برای حفاظت پایدار از عرصه‌های جنگلی در عصر تغییرات اقلیمی است.

26. Generalizability

References

Abatzoglou, J. T., Williams, A. P., & Barbero, R. (2019). Global emergence of anthropogenic climate change in fire weather indices. *Geophysical Research Letters*, *46*(1), 326–336.

Bjånes, A., De La Fuente, R., & Mena, P. (2021). A deep learning ensemble model for wildfire susceptibility mapping. *Ecological Informatics*, *65*, 101397.

Bouzeraa, Y., Bouchemal, N., Djaaboub, S., Hristov, G., & Zahariev, P. (2025). Machine Learning-Based Wildfire Susceptibility Mapping: A GIS-Integrated Predictive Framework. *Applied Sciences*, *15*(22), 12188.

Cilli, R., Elia, M., D'Este, M., Giannico, V., Amoroso, N., Lombardi, A., Pantaleo, E., Monaco, A., Sanesi, G., & Tangaro, S. (2022). Explainable artificial intelligence (XAI) detects wildfire occurrence in the Mediterranean countries of Southern Europe. *Scientific reports*, *12*(1), 16349.

Ejaz, N., & Choudhury, S. (2025). A comprehensive survey of the machine learning pipeline for wildfire risk prediction and assessment. *Ecological Informatics*, 103325.

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., & Aryal, J. (2019). Forest fire susceptibility and risk mapping using social/infrastructural vulnerability and environmental variables. *Fire*, *2*(3), 50.

Ghorbanzadeh, O., Valizadeh Kamran, K., Blaschke, T., Aryal, J., Naboureh, A., Einali, J., & Bian, J. (2019). Spatial prediction of wildfire susceptibility using field survey GPS data and machine learning approaches. *Fire*, *2*(3), 43.

Gitas, I. Z., de Santis, A., & Mitri, G. H. (2009). Remote sensing of burn severity. In *Earth observation of wildland fires in mediterranean ecosystems* (pp. 129–148). Springer.

Roodsarabi, Z. , Sam Khaniani, A. and Kiani, A. (2022). A Review of Remote Sensing Methods in Identifying and Monitoring Forest Fires. *Iranian Journal of Remote Sensing and GIS*, *14*(4), 19-52. doi: 10.52547/gisj.14.4.19. (in Persian)

Jones, M. W., Smith, A., Betts, R., Canadell, J. G., Prentice, I. C., & Le Quéré, C. (2020). Climate change increases the risk of wildfires.

Leo., B. (2001). Random forests.

Moghim, S., & Mehrabi, M. (2024). Wildfire assessment using machine learning algorithms in different regions. *Fire Ecology*, *20*(1), 104.

Nguyen Van, L., & Lee, G. (2025). Optimizing stacked ensemble machine learning models for accurate wildfire severity mapping. *Remote Sensing*, *17*(5), 854.

Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. *Clinical and translational science*, *17*(11), e70056.

Pouyan, S., Pourghasemi, H. R., Bordbar, M., Rahmanian, S., & Clague, J. J. (2021). A multi-hazard map-based flooding, gully erosion, forest fires, and earthquakes in Iran. *Scientific reports*, *11*(1), 14889.

Shahzad, F., Mehmood, K., Anees, S. A., Adnan, M., Muhammad, S., Haidar, I., Ali, J., Hussain, K., Feng, Z., & Khan, W. R. (2025). Advancing forest fire prediction: A multi-layer stacking ensemble model approach. *Earth Science Informatics*, *18*(3), 270.

Tavakkoli Piralilou, S., Einali, G., Ghorbanzadeh, O., Nachappa, T. G., Gholamnia, K., Blaschke, T., & Ghamisi, P. (2022). A Google Earth Engine approach for wildfire susceptibility prediction fusion with remote sensing data of different spatial resolutions. *Remote sensing*, *14*(3), 672.

Tonini, M., D'Andrea, M., Biondi, G., Degli Esposti, S., Trucchia, A., & Fiorucci, P. (2020). A machine learning-based approach for wildfire susceptibility mapping. The case study of the Liguria region in Italy. *Geosciences*, *10*(3), 105.

Xu, Z., Li, J., Cheng, S., Rui, X., Zhao, Y., He, H., Guan, H., Sharma, A., Erxleben, M., & Chang, R. (2025). Deep learning for wildfire risk prediction: Integrating remote sensing and environmental data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *227*, 632–677.