

Modeling vegetation changes in Zagros forests based on machine learning algorithms and satellite images

Amin Zoratipour ^{*1} , Ghasem Shahparifar² , Ashkan Yusefi ³ 

¹ Corresponding Author, Associate Professor, Department of Natural Engineering, Faculty of Agriculture, Khuzestan University of Agricultural Sciences and Natural Resources, Khuzestan, Iran.

Email: Zoratipour@asnrukh.ac.ir

² Master's degree student, Department of Engineering Geology, Faculty of Science, University of Isfahan, Isfahan, Iran. Email: ghasemshahparifar@yahoo.com

³ Department of Soil Science and Engineering, Faculty of Agriculture, Khuzestan University of Agricultural Sciences and Natural Resources, Khuzestan, Iran. Email: ashkanyusefi2013@yahoo.com

ABSTRACT

This research investigated and modeled changes in vegetation cover in southwestern Iranian forests using Landsat 8 satellite images and six machine learning algorithms including random forest(RF), Boosted Regression Tree (BRT), support vector machine(SVM), generalized additive model(GAM), Classification and Regression Tree (CART) and generalized linear model(GLM). First, satellite images were processed to determine deforestation areas over a decade from 2013 to 2023, and changes in Normalized Difference Vegetation Index in forest of the study area were calculated. Then, the parameters affecting forest cover changes were examined, and validation and verification of machine learning models were performed using efficiency indices. The results showed that in this study area, several factors affected forest cover and in some areas led to a 20% decrease in Normalized Difference Vegetation Index. The results also indicate that the random forest model has greater accuracy and precision than the other five models. The area under the curve (AUC) of these models is 0.986, 0.0.92, 0.865, 0.761, 0.743 and 0.682 respectively. On the other hand, the impact of different factors on forest cover reduction in different models is not the same; in general, the importance of the parameters of distance from the river, precipitation, evaporation, distance from villages, distance from the dam lake, erosion, distance from the road, slope direction, topographic wetness index, fire and slope are 0.176, 0.170, 0.170, 0.169, 0.167, 0.165, 0.165, 0.163, 0.162, 0.161 and 0.160 respectively.

Keywords: Algorithm, vegetation, deforestation, Khuzestan, machine learning.

EXTENDED ABSTRACT

Introduction

Today, studying the parameters affecting forest regeneration and modeling them is expensive and complex. The phenomenon of forest destruction will lead to the release of millions of tons of carbon dioxide into the atmosphere, creating greenhouse conditions and disrupting the balance of the planet. Deforestation is one of the most important environmental challenges, as it is one of the most important factors in the weakening and destruction of the environment. Recently, by training machine learning models using experimental data and using the R programming language, it is possible to design algorithms and provide accurate predictions of forest growth and changes based on environmental conditions. Machine learning is not limited to environmental issues, but has been successfully used in diverse fields such as facial and voice recognition, medicine, robotics, aquaculture, food security, and climatology. Although this technology faces obstacles such as one-sided conclusions and data limitations, it will be accompanied by many advances and applications in the future. Machine learning is a type of data analysis that automates model analysis. This method allows machines to learn from experience.

Materials and Methods

In this study, eleven different parameters were used as factors affecting vegetation reduction in southwestern Iranian forests based on machine learning algorithms and satellite images. The factors studied include distance from village, distance from river, distance from lake, distance from road, soil erosion, precipitation, evaporation, fire, topographic wetness index, slope percentage and slope direction. In this regard, the digital elevation model (DEM) of the region was extracted from the USGS website. Also, slope percentage, slope direction, topographic wetness index and distance from the river were extracted using a digital elevation model. Distance from roads and villages were also prepared using Landsat 8 satellite data. Changes in Normalized Difference Vegetation Index (NDVI) in forest areas were obtained by processing Landsat 8 data. For this purpose, Landsat 8 satellite data was

used during the years 2013 to 2023. Forest areas with a decrease in Normalized Difference Vegetation Index of more than 20% are considered as deforestation areas. Then, in this stage, in order to train machine learning models, 270,000 points were randomly selected for areas with and without deforestation (with all land uses) in the entire study area. Of these, about 90,000 points that had more than 20 percent changes in the Normalized Difference Vegetation Index in forest areas and 180,000 points that had less than 20 percent changes in the Normalized Difference Vegetation Index in forest areas. All the points were randomly divided into two groups (70% for learning and 30% for testing the model). The learning data was used to train the selected model and the testing data was used for validation. The available data was also used as input to the trained models to analyze the probability of future deforestation.

Results

In order to identify areas experiencing deforestation, Landsat 8 satellite data was used from 2013 to 2023. Areas with a decrease of more than 20% in the Normalized Difference Vegetation Index a ten-year period were considered as areas experiencing deforestation. In this study, different methods were used to validate the accuracy of the models. Six different machine learning models, random forest(RF), support vector machine(SVM), generalized additive model(GAM), generalized linear model(GLM), classification, Boosted Regression Tree (BRT) and Classification and Regression Tree (CART) were validated using different parameters including ROC and AUC plots, efficiency, TPR and FPR . These criteria indicate the predictive ability of machine learning models. The AUC values of RF, BRT, SVM, GAM, CART and GLM models are 0.986, 0.92, 0.865, 0.761, 0.743 and 0.682, respectively. The results showed that the random forest model was more accurate in predicting deforestation than other models in the present study. The sensitivity (TPR) values of RF, BRT, CART, SVM, GAM and GLM models were estimated as 0.953, 0.835, 0.826, 0.768, 0.668 and 0.604, respectively, and the specificity (FPR) values of the models were estimated as 0.047, 0.142, 0.240, 0.174, 0.233 and 0.306, respectively. These figures indicate sufficient accuracy of all models used in this study for investigating forest cover changes . Considering the AUC, TPR and FPR indices, it can be concluded that all models have good accuracy, but overall, the random forest (RF) and boosted tree regression (BRT) models have more accuracy than the other four models. The importance of the parameters in the models used varied, in general, five parameters distance from the river, precipitation, evaporation, distance from the lake, and distance from the village are more important than other parameters. Therefore, climatic and anthropogenic factors simultaneously affect the destruction of forests in Khuzestan province.

Conclusions

Although the six models used have sufficient accuracy, the comparison of the models shows that the random forest and tree-reinforced regression models are more accurate than the other four models. While the generalized additive and generalized linear models showed less accuracy than the other models. Finally, the results showed that among the six models used, the random forest (RF) model has the highest accuracy. Since the calculation method used by the models is different from each other; therefore, the extent of the areas at risk in the Khuzestan province area is not the same in different models. Studies show that the random forest model introduces the risk of deforestation in more limited areas than other models. While the support vector model and generalized ensemble models act conservatively and show more areas of the forests of the study area at risk. The risk of deforestation in the study area was investigated and predicted using six models. The models used show between 2 and 8 percent of the total area of Khuzestan province (11 to 51 percent of the province's forest area) at high risk of deforestation. On average, about 4 percent of the total area of the province has a high probability of deforestation, which includes about 29 percent of the province's forests.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank all participants of the present study.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Conflict of interest

The author declares no conflict of interest.

مدل سازی تغییرات پوشش گیاهی جنگل های استان خوزستان مبتنی بر الگوریتم یادگیری ماشین و تصاویر ماهواره ای

امین ذرتی پور^۱، قاسم شه پری فر^۲، اشکان یوسفی^۳

^۱ نویسنده مسئول، دانشیار، گروه مهندسی طبیعت، دانشکده کشاورزی، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان، خوزستان، ایران.

رایانامه: Zoratipour@asnrukh.ac.ir

^۲ دانش آموخته کارشناسی ارشد، گروه زمین شناسی مهندسی، دانشکده علوم، دانشگاه اصفهان، اصفهان، ایران.

رایانامه: ghasemshahparifar@yahoo.com

^۳ گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه علوم کشاورزی و منابع طبیعی خوزستان، خوزستان، ایران.

رایانامه: ashkanyusefi2013@yahoo.com

چکیده

در سال های اخیر پوشش جنگلی، تحت تأثیر عوامل متعددی همچون تغییرات اقلیم و فعالیت های انسانی در معرض خطر قرار گرفته است، با توجه به اهمیت بالای این منابع عظیم، این پژوهش به بررسی و مدل سازی تغییرات پوشش گیاهی جنگل های استان خوزستان با استفاده از شش الگوریتم یادگیری ماشین شامل جنگل تصادفی، رگرسیون تقویت شده درختی، بردار پشتیبان، مدل جمعی تعمیم یافته، درخت طبقه بندی و رگرسیون و مدل خطی تعمیم یافته می پردازد. به طوری که به منظور تعیین و تشخیص مناطق جنگل زدایی شده در استان، تصاویر ماهواره ای لندست ۸، طی دوره زمانی ده ساله (از سال ۲۰۱۳ تا ۲۰۲۳)، مورد پردازش قرار گرفتند و تغییرات شاخص پوشش گیاهی در مناطق جنگلی در محدوده کل استان خوزستان، به مساحت ۶۴۰۵۷ کیلومتر مربع، محاسبه گردید. پارامترهای متعددی بر تغییرات پوشش جنگلی موثر هستند که از آن میان، عواملی همچون فاصله از روستا، فاصله از رودخانه، فاصله از دریاچه سد، فاصله از جاده، فرسایش خاک، بارش، تبخیر، آتش سوزی، شاخص رطوبت توپوگرافی، درصد شیب و جهت شیب مورد بررسی قرار گرفتند. در نهایت اعتبارسنجی و صحت سنجی کلیه الگوریتم های یادگیری ماشین با استفاده از شاخص های راندمان انجام گرفت. مساحت زیر نمودار (AUC) این مدل ها که بیانگر میزان دقت هر یک از مدل ها می باشد، به ترتیب ۰/۹۸۶، ۰/۹۲۰، ۰/۸۶۵، ۰/۷۶۱، ۰/۷۴۳ و ۰/۶۸۲ برآورد شد. نتایج نشان دادند، مدل جنگل تصادفی دارای صحت و دقت بیشتری نسبت به پنج مدل دیگر است. همچنین، تأثیر عوامل مختلف بر کاهش پوشش جنگلی در مدل های مختلف یکسان نیست؛ به طور کلی اهمیت پارامترهای فاصله از رودخانه، بارش، تبخیر، فاصله از روستاها، فاصله از دریاچه سد، فرسایش، فاصله از جاده، جهت شیب شاخص رطوبت توپوگرافی، آتش سوزی و شیب به ترتیب برابر با ۰/۱۷۶، ۰/۱۷۰، ۰/۱۶۹، ۰/۱۶۷، ۰/۱۶۵، ۰/۱۶۵، ۰/۱۶۳، ۰/۱۶۲، ۰/۱۶۱ و ۰/۱۶۰ می باشند. نتایج این پژوهش به طور کلی دقت بالای الگوریتم های یادگیری ماشین را در بررسی مسئله جنگل زدایی در استان خوزستان به خوبی نشان می دهند. مدل های مورد استفاده بین ۲ تا ۸ درصد از مساحت کل استان را در معرض خطر زیاد جنگل زدایی نشان می دهند. به عبارت دیگر، به طور میانگین یک سوم جنگل های استان در معرض خطر بالای جنگل زدایی قرار دارند.

کلیدواژه ها: الگوریتم، پوشش گیاهی، جنگل زدایی، خوزستان، یادگیری ماشین.

مقدمه

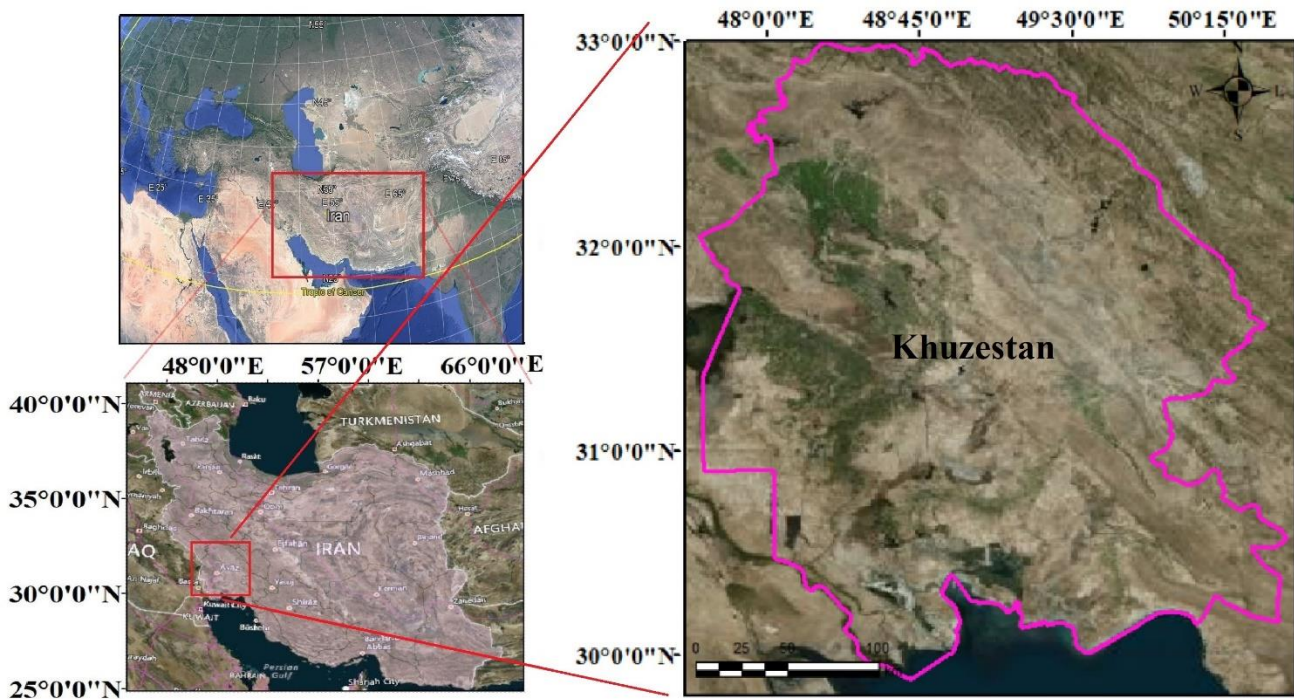
زندگی بیش از ۱/۶ میلیارد انسان در سراسر جهان به جنگل‌ها وابسته است. طبق مطالعات فائو، اراضی جنگلی به زمین‌های با حداقل مساحت نیم هکتار و درختان بیش از ۵ متر ارتفاع با حداقل سطح پوشش ده درصد اطلاق می‌شود (Tovar, 2009). طی این گزارش، جنگل‌ها بیش از ۳۱ درصد سطح خشکی‌های زمین، به مساحت حدود ۴ میلیارد هکتار را پوشانده‌اند، بطوری‌که بیش از پنجاه درصد این جنگل‌ها در پنج کشور روسیه، برزیل، کانادا، چین و ایالات متحده آمریکا قرار دارند (FAO and UNEP, 2020). جنگل‌ها از طریق تأثیرگذاری بر پدیده تبخیر و تعرق اراضی و نیاز آبی گیاهی، مکانسیم اقلیم و وضعیت هوا را تحت تأثیر قرار می‌دهند (CPF, 2012). نتایج مطالعات صورت گرفته در خصوص تغییرات کاربری اراضی در کشور در دو دهه گذشته نشان می‌دهد کاربری جنگل‌بین سال‌های ۱۹۹۳ تا ۲۰۱۵ تخریب گسترده داشته است، اراضی جنگلی به میزان ۱۴/۶ درصد معادل ۳۰۰ هزار هکتار کاهش پیدا کرده و تنها ۲ میلیون هکتار از مساحت آن از کشور باقی مانده است. نتایج به دست آمده از این پژوهش بیان می‌کند که توسعه‌های فیزیکی انسان ساخت، نقش ناچیزی در تخریب جنگل و تغییر سیمای سرزمین دارد. که در این میان مراتع با ۴ میلیون هکتار بیشترین کاهش را داشته است. از طرفی، جنگل‌ها به صورت چشمگیری به میزان ۱۴ درصد کاهش پیدا کرده‌اند. منابع آبی نیز از این قاعده مستثنا نیستند و با کاهش ۳۱ درصدی مواجه شده‌اند. از طرف دیگر، افزایش اراضی انسان ساز با رشد بسیار زیاد ۱۳ درصدی همراه بوده است (کشتکار، ۱۴۰۳). به‌طوری‌که مناطق جنگلی با جذب دی اکسیدکربن هوا بر چرخه تثبیت کربن تأثیرگذار هستند (Senior et al., 2017). پدیده جنگل‌زدایی ۱۲ تا ۱۷ درصد از گاز دی اکسید کربن را به جو می‌افزاید (CPF, 2012). پدیده تخریب جنگل، منجر به آزادسازی میلیون‌ها تن دی اکسیدکربن به جو زمین و ایجاد شرایط گلخانه‌ای و به هم خوردن تعادل کره خاکی خواهد شد. جنگل‌زدایی یکی از مهم‌ترین چالش‌های زیست‌محیطی به شمار می‌آید، چرا که از مهم‌ترین عوامل تضعیف و نابودی زیست محیط است (Kumar et al., 2014). یکی از عوامل تأثیرگذار در تعیین حد جنگل‌زدایی تعیین احتمال جنگل‌زدایی است. تعیین محدوده‌های در معرض جنگل‌زدایی می‌تواند در برنامه‌ریزی‌های دقیق و تعیین روش‌های کارآمد جهت رفع این مشکل بسیار مؤثر واقع شود. امروزه مشکل جنگل‌زدایی و راهکار حل آن یکی از موضوعات مورد بحث توسط محققان در سراسر جهان است. بدین منظور روش‌های آماری با استفاده از تکنیک‌های یادگیری ماشین به منظور پیش‌بینی دقیق‌تر جنگل‌زدایی توسعه یافته‌اند (Bera et al., 2020). با توجه به توانایی الگوریتم‌های یادگیری ماشین در پیش‌بینی روابط پیچیده بین پارامترهای وابسته و پیش‌بینی‌کننده، این مدل‌ها در زمینه حل مسائل متعدد زیست محیطی به کمک مدیران و محققان آمده‌اند (Mayfield et al., 2017; Fenton and Neil, 2018). نتایج پژوهش Tuoku et al., (۲۰۲۴) نشان داد که بارش و تابش خورشید بیشترین تأثیر را بر میزان شاخص پوشش گیاهی دارند، در حالی که رطوبت نسبی و تجمع دی اکسید کربن کم‌ترین تأثیرگذاری را دارند. وانگ و همکاران (۲۰۲۳) تغییرات شاخص پوشش گیاهی NDVI را در کوه‌های کویلین چین بررسی کرده و بیان کردند که در مناطق کم‌ارتفاع این کوه‌ها، بارندگی و دما مهم‌ترین پارامترهای مؤثر هستند در حالی که در مناطق مرتفع‌تر، دمای هوا بیشتر تأثیرگذار است (Wang et al., 2023). از سوی دیگر Kumar et al., (۲۰۲۲)، با استفاده از تصاویر ماهواره‌ای لندست ۸، تغییرات بیابان‌زایی در مناطق نیمه خشک را مورد بررسی قرار دادند. نتایج کارآنان قابلیت استفاده از تصاویر لندست ۸ را در پایش بیابان‌زدایی و دمای محیط نشان داد (Kumar et al., 2022). یکی از مهم‌ترین پارامترهایی که می‌توان با استفاده از تصاویر ماهواره‌ای مورد مطالعه قرار داد، بررسی شدت و تأثیرات جنگل‌زدایی در منطقه است. در همین راستا Matthew et al., (۲۰۱۴) با استفاده از تصاویر لندست ۸، نشان دادند که سرعت جنگل‌زدایی طی سال‌های ۱۹۹۰ تا ۲۰۱۸ در جنگل‌های آمازون به سرعت افزایش یافته است. این روند در منطقه پرو از ۱۹۰۰ هکتار در سال به ۴۲۰۰ هکتار در سال افزایش یافت (Marcus et al., 2024). همچنین مطالعات فلکس دوکسون بوبا و همکاران در نیجریه نشان داد که جنگل‌زدایی طی سال‌های ۱۹۸۶ تا ۲۰۱۴ تبعات زیادی به همراه داشته است. نتایج این مطالعات ارتباط نزدیک جنگل‌زدایی را با عوامل اقلیمی از جمله دما و بارش نشان داد (Buba et al., 2020). بر اساس آخرین تخمین سازمان خواربار و کشاورزی جهانی در سال ۲۰۱۴ وسعت جنگل‌های ایران به ۱۰/۶ میلیون هکتار کاهش پیدا کرده است. همچنین مطابق گزارشات مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی کشور، سالانه ۱۲۵ هزار هکتار جنگل در کشور تخریب می‌شود که سهم جنگل‌های شمال ۴۵ هزار هکتار یعنی ۳۶

درصد آن است (FAO and UNEP, 2020). بطور کلی پایش و ارزیابی پارامترهای مؤثر بر احیای جنگل و مدل سازی آن‌ها، زمان‌بر و پیچیده هستند. امروزه با آموزش مدل‌های یادگیری ماشین با استفاده از داده‌های تجربی به همراه بهره جستن از زبان برنامه نویسی R، می‌توان طراحی الگوریتم‌ها را انجام داده و پیش‌بینی دقیقی از روند رشد و تغییرات جنگل‌ها براساس شرایط محیطی ارائه داد (Jin Zhao et al., 2024). یادگیری ماشین محدود به مسائل زیست محیطی نیست، بلکه به طور موفقیت آمیزی در زمینه‌های متنوع از جمله تشخیص چهره و صدا، پزشکی (Kang et al., 2015, Zhang et al., 2017)، رباتیک (Takahashi et al., 2017)، کشت آبی (Zhou et al., 2018)، امنیت غذایی (Maione et al., 2019)، اقلیم شناسی (Fang et al., 2017) و محیط زیست (شعبانی و همکاران، ۱۴۰۰)، (Moscovini et al., 2024) مورد استفاده قرار گرفته شده است. گرچه این تکنولوژی با موانعی از جمله نتیجه گیری یک طرفه و محدودیت داده‌مواجه هست. با این وجود، در آینده این تکنولوژی با پیشرفت‌ها و کاربردهای فراوان همراه خواهد بود (شریفانی و امینی، ۲۰۲۳). یادگیری ماشین نوعی از آنالیز داده‌است که تحلیل مدل‌ها را به صورت خودکار انجام می‌دهد. این روش امکان یادگیری ماشین را با استفاده از تجربه فراهم می‌نماید (Jabeen et al. 2023). مطالعات زیادی در زمینه به کارگیری مدل‌های مختلف یادگیری ماشین در حل مسائل پیچیده صورت گرفته است؛ Saha et al., (۲۰۲۰)، با استفاده از مدل جنگل تصادفی به بررسی احتمال جنگل زدایی پرداختند و نشان دادند که جنگل‌های متراکم در هند از سال ۱۹۹۰ تا ۲۰۱۷ به میزان ۹ درصد کاهش یافته‌اند (Saha et al., 2020). همچنین، شعبانی و همکاران (۱۴۰۰) با بکارگیری مدل‌های رگرسیون تقویت شده درختی (BRT) و رگرسیون منطقی (Logestic regression) به بررسی آسیب به جنگل‌ها در طول عملیات کشت پرداختند و پارامترهای تأثیرگذار از جمله فاصله از جاده، تراکم جنگل و را مورد بررسی قرار دادند. نتایج کار آنان تأثیر بیشتر پارامترهای شیب و طول شیب را تایید می‌کند (شعبانی و همکاران، ۱۴۰۰). Moscovini et al., (۲۰۲۴) با استفاده از روش یادگیری ماشین، امکان پیش‌بینی نسبتاً دقیق شاخص پوشش گیاهی را با استفاده از تصاویر RGB بررسی کرده و نتیجه را مثبت ارزیابی کردند. نتایج کار آنان نشان داد که صحت پیش‌بینی به روش یادگیری ماشین بیش از ۷۰ درصد می‌باشد. Aksoy et al., (۲۰۲۴) با استفاده از تصاویر لندست ۸ و به روش یادگیری ماشین گسترش شوری دریاچه ارومیه را تخمین زدند. نتایج کار آنان نشان داد که پارامترهای فیزیکی خاک از جمله درخشندگی سطح و پوشش گیاهی متاثر از شوری خاک محدوده دریاچه ارومیه هستند و همچنین نتایج امکان استفاده از یادگیری ماشین به منظور کاهش هزینه‌های نمونه برداری و کاهش صحرایی را نشان داد. ناحیه ریشی زاگرس با مساحتی بیش از ۵/۵ میلیون هکتار، بزرگترین منطقه جنگلی کشور بوده که حدود ۴۰ درصد از جنگل‌های کشور را در بر می‌گیرد و پوشش غالب و گونه اصلی درختی این جنگل‌ها، گونه بلوط است. در همین راستا اداره کل منابع طبیعی و آبخیزداری استان خوزستان سطح جنگل‌های استان خوزستان را حدود ۹۸۰ هزار هکتار اعلام کرده است، که شامل جنگل‌ها و مراتع مشجر و بیشه زارها می‌باشد. در این رابطه عالی محمودی (۲۰۱۸) نشان داد که وسعت پوشش جنگلی زاگرس در منطقه خوزستان در طی سال‌های ۲۰۰۳ تا ۲۰۱۱ به دلیل فعالیت‌های انسانی و تغییرات کاربری و پدیده آتش سوزی طبیعی، سیر قهقراپی داشته است. افزایش روند آتش سوزی در این جنگل‌ها با افزایش متوسط بارش منطقه به همراه توسعه گیاهان یکساله معنی دار بوده است (عالی محمودی، ۱۳۹۷). شیرمردی و همکاران (۱۳۹۹) نشان دادند که احیاء گونه‌های بوته‌ای و درختچه‌ای کوه‌های زاگرس باتوجه به تغییرات اقلیم به بیش از ۲۵ سال زمان نیاز دارد. بنابراین ارزیابی و پایش تغییرات زمانی و مکانی جنگل‌ها در جهت افزایش تولید ناخالص ملی و جلوگیری از تخریب آن جز اقدامات ضروری هر کشوری محسوب می‌گردد. با توجه به اهمیت مساله حفاظت، احیاء و پایش جنگل‌های جنوب غرب کشور و تغییرات ناشی از تغییر کاربری اراضی و تغییر اقلیم در آن‌ها در دهه‌های اخیر به‌خصوص جنگل‌های بلوط زاگرس، این پژوهش با استفاده از مدل‌های یادگیری ماشین و تصاویر ماهواره‌ای در ناحیه جنگلی زاگرس در استان خوزستان انجام گرفت (شیرمردی و همکاران، ۱۳۹۹).

منطقه مورد مطالعه

استان خوزستان با مساحت ۶/۴ میلیون هکتار بین مدارهای ۴۷ درجه و ۴۱ دقیقه تا ۵۰ درجه و ۳۹ دقیقه طول شرقی از نصف‌النهار گرینویچ و ۲۹ درجه و ۵۸ دقیقه تا ۳۳ درجه و ۴ دقیقه عرض شمالی از خط استوا، در قسمت جنوب غربی ایران واقع شده است (شکل ۱). این پهنه شامل دشت خوزستان و مناطق کوهستانی شمالی استان است که به علت تنوع آب و هوای این استان از تنوع

گونه‌ای بالایی برخوردار است. حدود ۱۵ درصد از مساحت استان از درختان و بوته‌های جنگلی پوشیده شده است. بخش اعظم پوشش گیاهی مناطق کوهستانی را درختان و درختچه‌های بلوط، گون، بنه وحشی و زالزالک تشکیل می‌دهد (عالی‌محمودی، ۱۳۹۷). متوسط بارندگی در استان خوزستان ۲۱۸ میلی‌متر و متوسط درجه حرارت ۲۶/۵ درجه سانتی‌گراد می‌باشد. از منظر زمین‌شناسی منطقه شامل دشت خوزستان، زاگرس چین‌خورده و زاگرس گسل‌خورده در مناطق کوهستانی است، که بخش اعظم دشت خوزستان از نهشته‌های کواترنری تشکیل شده در حالی که در مناطق چین‌خورده و کوهستانی سنگ‌های رسوبی سازندهای گچساران، اغاجاری، بختیاری، میشان، ایلام، سروک، پابده، گورپی و آسماری بیشترین فراوانی را دارند (ذرتی پور، ۱۴۰۱).



شکل ۱. موقعیت منطقه مورد مطالعه

روش اجرای مطالعه

گام اول - انتخاب عوامل مؤثر بر پدیده جنگل زدایی و رقومی سازی:

در این مطالعه به منظور مدل‌سازی تغییرات پوشش گیاهی جنگل‌های استان خوزستان مبتنی بر الگوریتم یادگیری ماشین و تصاویر ماهواره‌ای در اولین گام تعداد یازده پارامتر مختلف، به‌عنوان عوامل مؤثر بر کاهش پوشش گیاهی مورد استفاده قرار گرفتند و با استفاده از آزمون هم‌خطی بین پارامترها از عدم وجود رابطه قوی بین هر کدام با سایر پارامترها اطمینان حاصل شد. عوامل مورد بررسی عبارتند از فاصله از روستا، فاصله از رودخانه، فاصله از دریاچه سد، فاصله از جاده، فرسایش خاک، بارش، تبخیر، آتش‌سوزی، شاخص رطوبت توپوگرافی، درصد شیب و جهت شیب. در انتخاب عوامل مؤثر بر جنگل‌زدایی سعی شده است که حتی‌الامکان همه عوامل انسانی و طبیعی در قالب لایه‌های رستری تهیه شوند. در این رابطه مدل رقومی ارتفاعی (DEM) با دقت ۳۰ متر منطقه از وبسایت USGS استخراج گردید (<https://earthexplorer.usgs.gov>). همچنین درصد شیب، جهت شیب، شاخص رطوبت توپوگرافی، فاصله از رودخانه با استفاده از مدل ارتفاعی رقومی استخراج شدند. فاصله از جاده‌ها و روستاها نیز با استفاده از داده‌های ماهواره‌ای لندست ۸ آماده شدند. ابتدا نقشه کاربری اراضی با استفاده از داده‌های لندست ۸ استخراج و سپس فاصله اقلیدسی از عوارض مختلف در قالب لایه‌های رستری تهیه گردید. علاوه بر این با پردازش تصاویر ماهواره‌مادیس و لندست ۷ در گوگل ارث انجین، آتش‌سوزی‌های صورت گرفته (مساحت و موقعیت) در بازه زمانی ده ساله استخراج (اداره کل منابع طبیعی و آبخیزداری

استان خوزستان) و مورد استفاده قرار گرفت. نقشه های فرسایش (از مدل USLE) و زمین شناسی و سازندها، تبخیر و تعرق نیز به ترتیب از سازمان زمین شناسی و هواشناسی کشور تهیه شدند (اشکال ۲ و ۳).

– آزمون هم خطی

جهت جلوگیری از مشکل هم خطی بین پارامترهای مستقل، آزمون هم خطی انجام گردید. دو پارامتر خطای مجاز و عامل تورم واریانس جهت تشخیص هم خطی بودن پارامترها مورد استفاده قرار گرفتند. عامل تورم واریانس بیش از ۱۰ و خطای مجاز کم تر از ۰/۱ مشکل هم خطی را نشان می دهند (رابطه ۱ و ۲)، (Johnston et al., 2018). که در این معادله R^2_T ضریب همبستگی رگرسیون رابطه بین پارامترها را نشان می دهد.

گام دوم – پایش تغییرات شاخص پوشش گیاهی (NDVI) در مناطق جنگلی:

در گام دوم، تغییرات شاخص پوشش گیاهی (NDVI) در مناطق جنگلی با پردازش داده های لندست ۸ بدست آمد. بدین منظور از داده های ماهواره ای لندست ۸ طی سال های ۲۰۱۳ تا ۲۰۲۳ استفاده گردید. مناطق جنگلی با کاهش بیش از ۲۰ درصد شاخص پوشش گیاهی در سال ۲۰۱۳ نسبت به این شاخص در سال ۲۰۲۳ به عنوان نقاط دچار جنگل زدایی در نظر گرفته شده اند (Saha et al., 2022). در واقع نسبت تغییرات شاخص پوشش گیاهی طی دوره ده ساله نسبت به همین شاخص در سال ۲۰۱۳ به عنوان نقاط دچار جنگل زدایی در نظر گرفته شدند. حد آستانه جنگل زدایی برابر درصد تغییرات شاخص پوشش گیاهی در مناطق جنگلی در نظر گرفته شد.

گام سوم – آموزش و امتحان (تست) الگوریتم های یادگیری ماشین و کلاسه بندی:

سپس در مرحله سوم، به منظور آموزش مدل های یادگیری ماشین، تعداد ۲۷۰۰۰۰ نقطه به صورت تصادفی برای مناطق با و بدون جنگل زدایی (با کلیه کاربری های اراضی) در کل منطقه مطالعاتی انتخاب گردید که از این تعداد حدود ۹۰۰۰۰ نقطه که بیش از ۲۰ درصد تغییرات در شاخص پوشش گیاهی در مناطق جنگلی داشتند، با توجه به پیچیده و وسیع محدوده مورد مطالعه، جهت آموزش مدل های یادگیری ماشین سعی شده است تا شبکه نقاط آموزش از تعداد کافی برخوردار باشد. در محدود بودن تعداد نقاط آموزش، امکان بررسی روابط بین پارامترهای مستقل با جنگل زدایی وجود ندارد و آموزش مدل های به کار برده شده به نحو مناسب صورت نمی پذیرد. در نهایت به منظور شناسایی مناطق دچار جنگل زدایی، از داده های ماهواره ای لندست ۸ طی سال های ۲۰۱۳ تا ۲۰۲۳ استفاده گردید. مناطق با کاهش بیش از ۲۰ درصد شاخص پوشش گیاهی طی دوره ده ساله، به عنوان نقاط دچار جنگل زدایی در نظر گرفته شده اند (شکل ۴). در این گام، آموزش الگوریتم های یادگیری ماشین بود. طی پروسه یادگیری ماشین مناطق دچار تغییرات پوشش جنگلی و مناطق فاقد تغییرات این شاخص، به عنوان لایه باینری با دو کلاس کلی به مدل معرفی گردید. مناطق دچار جنگل زدایی با کد یک و سایر نقاط با کد صفر معرفی می شوند. همه نقاط به صورت کاملاً تصادفی به دو دسته تقسیم شدند (۷۰ درصد برای یادگیری و ۳۰ درصد برای تست کردن مدل). داده های یادگیری در این مرحله، جهت آموزش مدل انتخابی و داده های تست جهت اعتبارسنجی مورد استفاده قرار می گیرند. برای آنالیز احتمال جنگل زدایی در آینده نیز از داده های موجود جهت ورودی مدل های آموزش دیده شده استفاده شد. در پایان این مرحله شامل تهیه نقشه های مخاطره جنگل زدایی است. در ادامه مطالعه به منظور بررسی و مقایسه نتایج مدل ها نقشه های احتمال جنگل زدایی در همه الگوریتم ها بدست آمد و طبقه بندی گردید سپس بررسی آماری شدند. نتایج این پیش بینی به صورت سه کلاس خطر پایین (کم تر از ۴۰ درصد)، خطر متوسط (۴۰ تا ۶۰ درصد) و خطر زیاد (بیش از ۶۰ درصد) طبقه بندی و ارائه شده است. در گام پایانی تحلیل اهمیت پارامترهای موثر، جهت مقایسه نسبی تاثیر این عوامل بر جنگل زدایی یا استفاده از شش مدل مختلف یادگیری ماشین، صورت گرفت.

گام چهارم _ ارزیابی و اعتبارسنجی مدل‌ها و ارزیابی پارامترها:

میزان دقت هر مدل و توانایی آن در برآورد دقیق پارامتر اصلی با اعتبارسنجی مدل مشخص می‌شود. در این زمینه روش‌های متعدد مورد استفاده قرار می‌گیرد. یکی از بهترین روشها استفاده از نمودار مشخصه نسبی عملکرد (ROC) می‌باشد که میزان دقت یک مدل را نشان می‌دهد. مقدار نزدیک به یک دقت بالا و توانایی مدل در پیش‌بینی درست پارامتر را نشان می‌دهد. در حالی که هرچه قدر این عدد به صفر نزدیک شود مدل دقت کم تری دارد. در کنار استفاده از ROC پارامترهای FNR ، TNR ، FPR ، TPR ، $Efficiency$ و ... را می‌توان جهت بررسی دقت و توانایی یک مدل به کار گرفت (جدول ۱)، (رابطه ۳ و ۴).

که TP جواب‌های درست مثبت، FN جواب‌های اشتباه منفی، FP جواب‌های اشتباه مثبت و TN جواب‌های درست منفی را نشان می‌دهد (رابطه ۵). مطابق رابطه زیر کارایی یک مدل محاسبه می‌شود (جدول ۱).

ارزیابی اهمیت پارامترهای تأثیرگذار

در این پژوهش همچنین آنالیز فاکتورهای تأثیرگذار بر جنگل‌زدایی به طور واضح مورد بررسی قرار گرفت. بدین منظور در ادامه مطالعه، تغییرات صحت مدل با حذف هر پارامتر بررسی گردید. تغییر صحت هر مدل با حذف هر پارامتر به عنوان اهمیت آن برآورد شد و سپس میانگین وزنی هر مدل با لحاظ کردن صحت آن مدل، محاسبه و مورد مقایسه قرار گرفت.

کاربرد مدل‌های مختلف یادگیری ماشین

مدل جنگل تصادفی

مدل جنگل تصادفی یکی از کاربردی‌ترین مدل‌های یادگیری خردجمعی است که توسط بریمان (Breiman, 2001) ارائه گردید. این مدل با ایجاد چندین درخت طبقه‌بندی و برآورد خرد جمعی همه درخت‌ها مدل‌نهایی را ایجاد می‌کند. مهم‌ترین پارامترهای این مدل تعداد درخت و ریشه مربع تعداد پارامترها هستند. همچنین این مدل از شاخص جینی به عنوان شاخص تقسیم استفاده می‌کند (جدول ۱)، (رابطه ۶). در اینجا، T بیانگر دیتاست و z تعداد داده‌هایی است که در هیچ یک از کلاس‌ها قرار نمی‌گیرند.

مدل بردار پشتیبان

مدل بردار پشتیبان یکی از کاربردی‌ترین مدل‌های یادگیری ماشین بر اساس کاهش ریسک خطا است که توسط وپنیک (Vapnik, 1995) ارائه شد. این روش داده‌های ورودی را در یک سطح تقسیم کرده و مرز بین هر یک از طبقات را به وضوح مشخص می‌کند. از داده‌های یادگیری ماشین جهت ایجاد تمایز بین کلاس‌های مختلف و ارتباط آن‌ها با پارامتر وابسته استفاده می‌شود (Pradhan, 2013) (رابطه ۷ تا ۱۰).

برای اجرای SVM، داده‌های آموزشی مورد نیاز است. این داده‌ها فرایند جداسازی طبقات را بهینه‌سازی می‌کنند با استفاده از یک تابع پایه شعاعی (RBF)، پراکنش طبقات با محدوده‌های غیرخطی را می‌توان در یک محدوده جدا شده خطی ترسیم نمود (هوانگ و همکاران، ۲۰۰۲). آموزش SVM با هسته تابع پایه شعاعی مستلزم تنظیم دو پارامتر است. یکی از این پارامترها، پارامتر منظم سازی است که تعادل بین به حداکثر رسانی حساسیت حواشی و به حداقل رساندن خطای آموزشی را برقرار کرده و کنترل می‌کند. طبقه بندی SVM شامل چهار نوع خطی، چند جمله‌ای، RBF و سیگموئید می‌باشد. نوع RBF در بیشتر موارد به خوبی عمل می‌کند. ساختار ریاضی این چهار نوع در معادلات زیر نشان داده شده است. در این معادلات، xi بردار پشتیبانی i ، xj نقاط آموزش داده شده برای j ها، t پارامتر هموارسازی، k تابع مرکزی، II شاخص اقلیدسی، γ پهنای هسته در توابع هسته مدل‌ها به جز مدل خطی، d نسبت درجه چند جمله‌ای در توابع هسته، r نسبت بایاس در توابع هسته‌ای مدل‌های چند جمله‌ای و سیگموئید بوده و γ ، d و r پارامترهای کنترل شده توسط کاربر می‌باشد زیرا تعریف صحیح آن‌ها بطور قابل توجهی باعث افزایش دقت SVM می‌شود.

مدل جمعی تعمیم یافته

مدل جمعی تعمیم یافته یک مدل ناپارامتری بوده و بسط مدل‌های خطی تعمیم یافته است که برخلاف مدل رگرسیون خطی، داده‌ها

شکل منحنی پاسخ را مشخص می‌کنند. در این مدل فرض بر این است که متغیر پاسخ دارای توزیعی از خانواده نمایی با میانگین است که از طریق تابع پیوند به متغیرهای تخمین گر متصل می‌شود. در مدل جمعی تعمیم یافته رابطه ۸ فرض بر این است که متغیر پاسخ Y دارای توزیع از خانواده نمایی با میانگین می‌باشد که از طریق تابع پیوند (g) به متغیرهای تخمین گر (X_j) متصل می‌شود (رابطه ۱۱)، (Hastie and Tibshirani, 1987)، (جدول ۱).

در اینجا فرض می‌شود که S_j ها توابعی نامعلوم و هموار و X_j متغیرهای تخمین گر هستند. به طور خاص S_j با استفاده از داده‌ها و با پراکنش استفاده از تکنیک های پیشرفته هموارساز نمودار پراکنش برآورد می‌شود. تفاوت اصلی مدل های جمعی تعمیم یافته با مدل های پارامتری در این است که توابع به وسیله توابع هموار نامعلوم جانشین می‌شوند.

مدل خطی تعمیم یافته

مدل خطی تعمیم یافته یک مدل خطی وابسته جهت ارتباط بین پارامتر وابسته و پارامترهای مستقل فاقد توزیع نرمال است. این مدل طیف وسیعی از داده‌ها با توزیع نمایی را پوشش می‌دهد. با استفاده از تابع پیوند، متغیر پیشگو و اندازه گیری شده باهم ترکیب می‌شوند. این مدل ارتقاء یافته مدل خطی است که از حالت طبیعی داده‌ها استفاده می‌کند در نتیجه داده‌های مورد استفاده می‌توانند غیرخطی باشند. در مدل GLM ، متغیرهای پیشگو $X_j (j=1,2,...,p)$ ترکیب شده و یک خط پیش بینی (LP) را ایجاد می‌نماید که وابسته به مقادیر مورد انتظار $\mu=E(Y)$ متغیر پاسخ Y می‌باشد که از طریق تابع پیوند محاسبه می‌گردد (رابطه ۱۲). که α عرض از مبدا، X متغیرهای پیش بینی کننده خطی، β ضریب رگرسیونی هر یک از متغیرهای پیشگو می‌باشد. (جدول ۱).

درخت طبقه بندی و رگرسیون

الگوریتم طبقه بندی و رگرسیون نوعی الگوریتم طبقه بندی است که برای ساخت درخت تصمیم بر اساس شاخص ناخالصی جینی عمل می‌کند. در واقع این یک درخت است که به انواع زیر درخت تصمیم گیری اشاره می‌کند. به عبارت دیگر، CART یک تک رگرسیون یا درخت طبقه بندی است که با استفاده از حداکثر قواعد تصمیم، مکرراً داده‌ها را به گروه‌هایی تقسیم می‌کند و جهت بررسی پارامتر توصیفی مورد استفاده قرار می‌گیرد. یکنواختی خروجی داده‌ها در هر مرحله با استفاده از R^2 اندازه گیری می‌شود. (رحیمی خوب، ۱۳۹۴).

رگرسیون تقویت شده درختی

مدل رگرسیون تقویت شده یک الگوریتم یادگیری ماشین بر اساس درخت های طبقه بندی و رگرسیون همراه با الگوریتم تقویت است. در این روش با تلفیق مدل های ساده دقت مدل افزایش می‌یابد (Aertsen et al., 2010). رگرسیون تقویت شده درختی از دو الگوریتم استفاده می‌کند؛ i کلاس بندی و درخت رگرسیون و ii ساخت و ترکیب چندین مدل به روش تقویت کردن. این روش مدل های دقیق تری نسبت به مدل های ساده ایجاد می‌کند بدین طریق این مدل بر مهمترین نقص مدل های ساده غلبه کرده است. ابتدا اولین درخت با استفاده از همه داده‌ها ایجاد شده و سپس درختان بعدی با استفاده از داده‌های باقیمانده و درخت اول شکل می‌گیرند (Elith et al., 2008).

جدول ۱. روابط و معادلات متغیرهای مختلف مدل های هوش مصنوعی استفاده شده در مطالعه

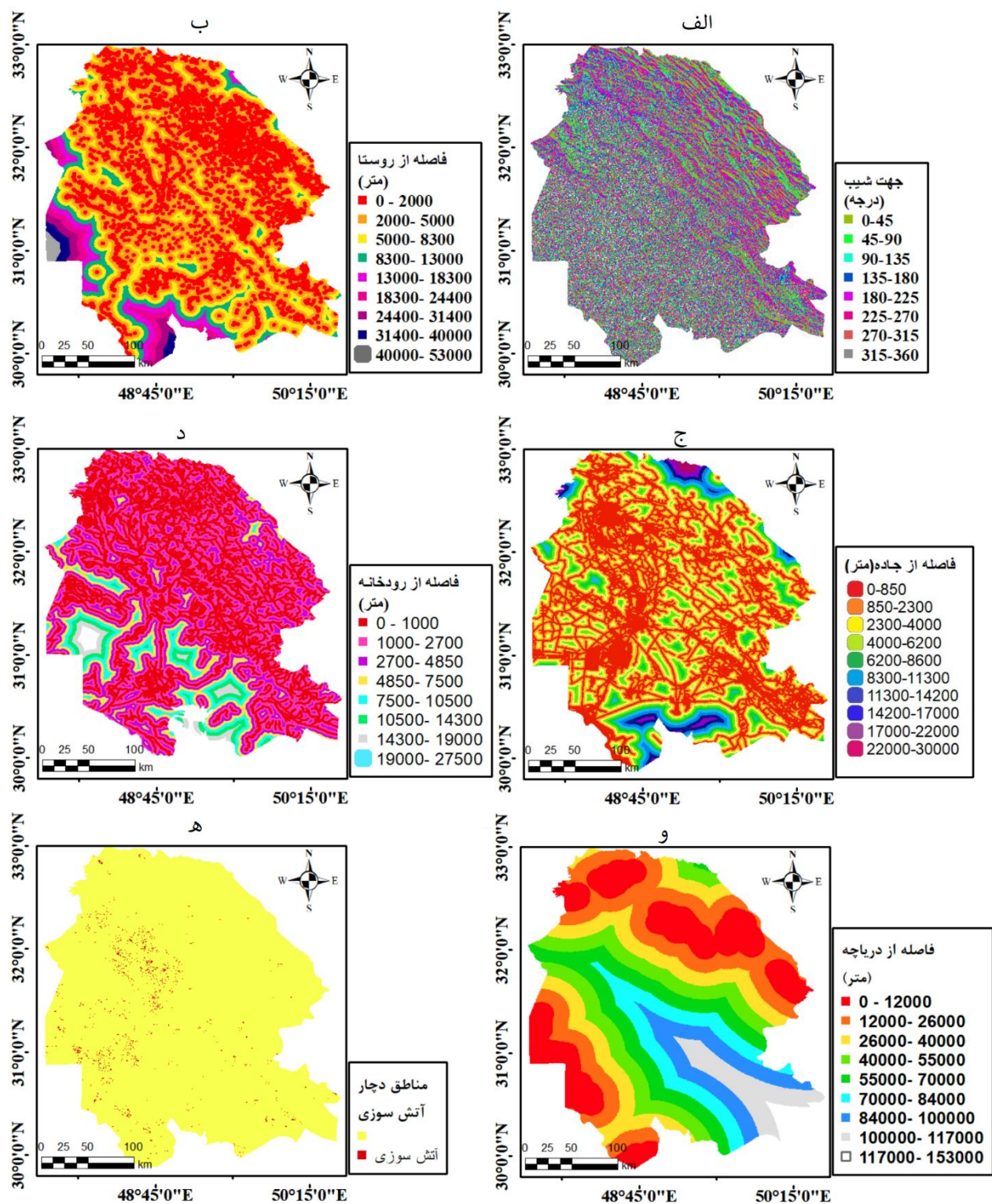
رابطه	ردیف
$Tolerance = 1 - R^2_I$	رابطه ۱
$VIF = 1/Tolerance$	رابطه ۲
$TPR = TP/(TP+FN)$	رابطه ۳
$FPR = FP/(FP+TN)$	رابطه ۴
$E = (TP+TN)/(TP+TN+FP+FN)$	رابطه ۵
$I_T(P) = \sum_{i=1}^j p_i \sum_{k \neq i}^j p_k = 1 - \sum_{i=1}^j p_i^2$	رابطه ۶
$K(x_i, x_j) = x_i^t, x_j$	رابطه ۷
$K(x_i, x_j) = (\gamma x_i^t, x_j + r)d, \gamma > 0$	رابطه ۸
$K(x_i, x_j) = \exp(-\gamma x_i - x_j ^2), \gamma > 0$	رابطه ۹

$K(x_i, x_j) = \tanh(-\gamma x_{it}, x_{jt} + r)$	رابطه ۱۰
$g(\mu) = \alpha + \sum_{j=1}^p = \int_j(x_i)$	رابطه ۱۱
$g(E(Y)) = LP = \alpha + \beta Xr$	رابطه ۱۲

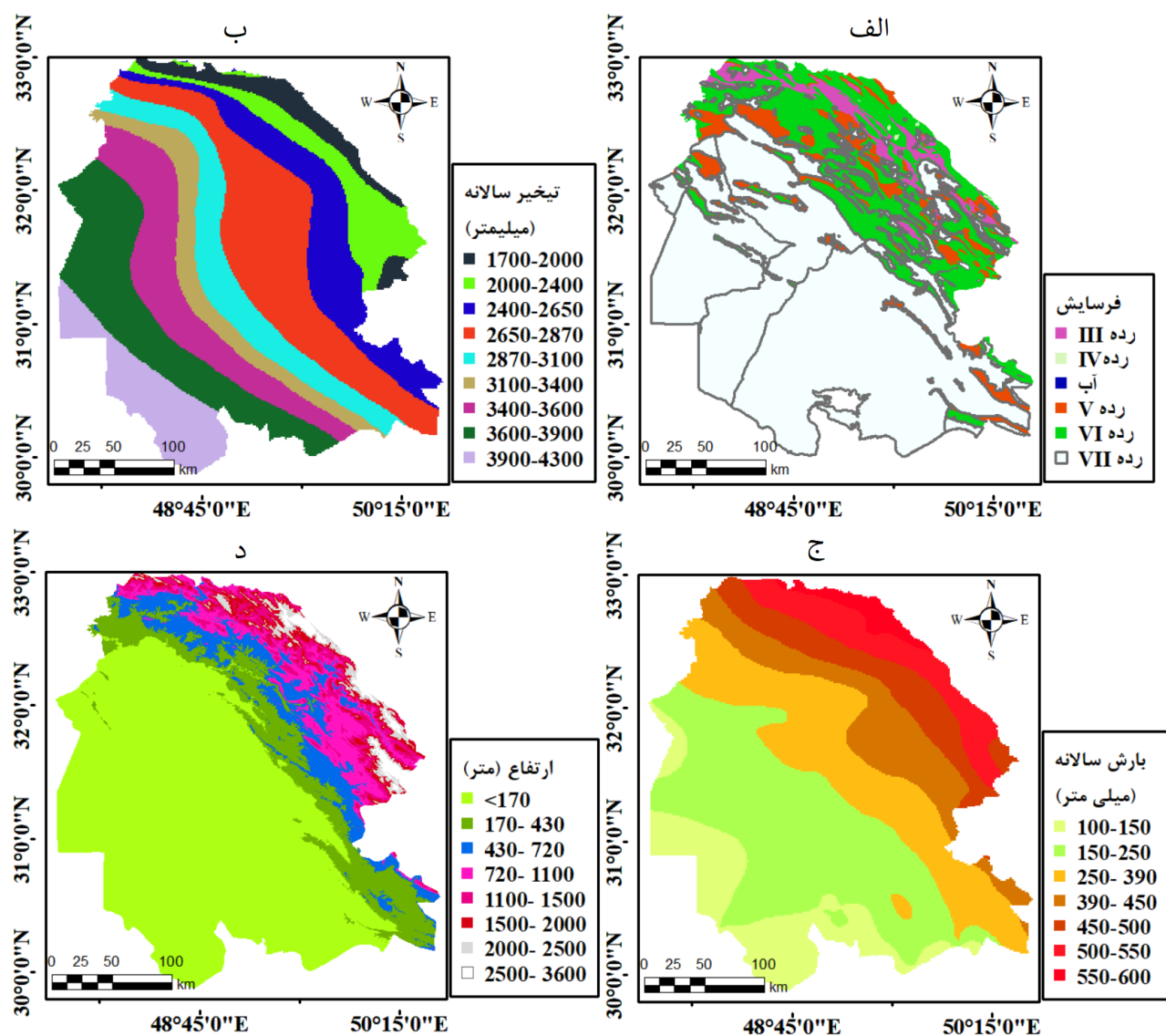
نتایج

در بخش نتایج، نقشه رقومی کلیه عوامل موثر بر پدیده جنگل‌زدایی نظیر از فاصله از روستا، فاصله از رودخانه، فاصله از دریاچه سد، فاصله از جاده، فرسایش خاک، بارش، تبخیر، آتش‌سوزی، شاخص رطوبت توپوگرافی، درصد شیب و جهت شیب تهیه و در شکل‌های ۲ و ۳ نشان داده شده است.

فیلتر استثنای نقشه

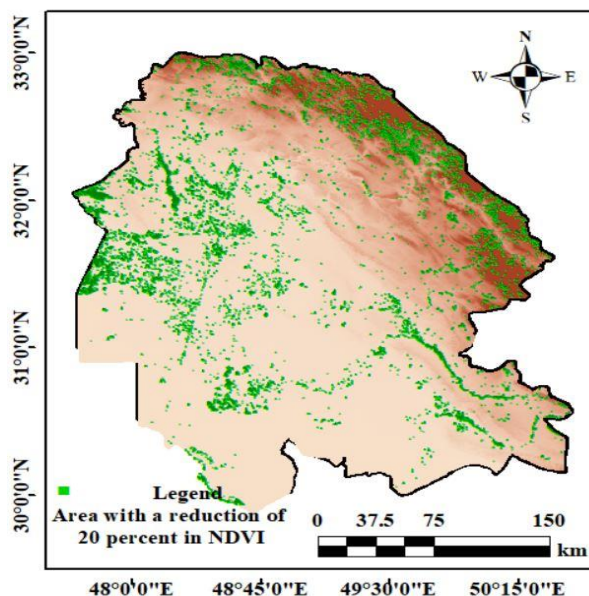


شکل ۲. لایه‌های مورد استفاده جهت تحلیل و پیش‌بینی جنگل‌زدایی، جهت شیب (الف)، فاصله از مناطق مسکونی (ب)، فاصله از جاده (ج)، فاصله از رودخانه (د)، فاصله از دریاچه سد (و)، آتش سوزی (ه).



شکل ۳. لایه‌های مورد استفاده جهت تحلیل و پیش‌بینی جنگل‌زدایی، کلاسه بندی فرسایش (الف)، تبخیر (ب)، بارش سالانه (ج)، مدل ارتفاعی رقومی (د)

با پردازش داده‌های لندست ۸، مناطق دچار جنگل‌زدایی در طول دوره زمانی ۲۰۱۳-۲۰۲۳ مطابق شکل زیر ارائه گردیده است (شکل ۴). در واقع تفاوت شاخص NDVI در مناطق جنگلی در شکل ۴، نمایانگر مناطقی است که طی ده سال دچار جنگل‌زدایی شده‌اند.



شکل ۴. مناطق جنگل‌زدایی با کاهش بیش از ۲۰ درصد شاخص پوشش گیاهی از سال ۲۰۱۳ تا ۲۰۲۳

جدول ۲. آمار جنگل‌زدایی یا استفاده از تصاویر ماهواره‌لندست ۸ طی ده سال

میزان کاهش شاخص پوشش گیاهی	درصد مساحت استان
بیش از ۱۰ درصد کاهش	۴/۰۸
بیش از ۲۰ درصد کاهش	۱/۳۴
بیش از ۳۰ درصد کاهش	۰/۶۸
بیش از ۴۰ درصد کاهش	۰/۴

مطابق جدول ۲، حدود ۴ درصد مساحت استان با حداقل ده درصد کاهش شاخص پوشش گیاهی طی ده سال مواجه بوده است. همچنین تصاویر ماهواره‌ای لندست ۸ طی سال‌های ۲۰۱۳ تا ۲۰۲۳، نشان می‌دهند که ۱/۳۴ درصد از مساحت کل استان با کاهش ۲۰ درصدی شاخص پوشش گیاهی همراه بود. در این پژوهش مناطقی که طی ده سال از ۲۰۱۳ تا ۲۰۲۳ به مقدار بیش از ۲۰ درصد دچار کاهش شاخص پوشش گیاهی شدند را به عنوان مناطق جنگل‌زدایی جهت آموزش الگوریتم‌های یادگیری ماشین مورد استفاده قرار گرفتند.

آنالیز هم خطی

نتایج ارائه شده در جدول ۳، عدم وجود مشکل هم خطی بین پارامترها را نشان می‌دهد. همه پارامترها عامل تورم واریانس کمتر از ۱۰ و خطای مجاز بیشتر از ۰/۱ دارند. بنابراین همبستگی و پدیده همخطی بین این پارامترها معنی دار نبوده و وجود ندارد، لذا برای به کارگیری در مدل‌سازی می‌توان از همه پارامترها استفاده نمود.

جدول ۳. نتایج آزمون هم خطی بین پارامترها

متغیر	محدوده خطا	ضریب تورم واریانس
فاصله از جاده	۰/۷۵	۱/۳۳
شاخص رطوبت توپوگرافی	۰/۸۷	۱/۱۶
بارندگی	۰/۳	۳/۳۸
فاصله از رودخانه	۰/۸۵	۱/۱۸
تبخیر	۰/۶۰	۱/۶۶
فاصله از دریاچه سد	۰/۸۱	۱/۲۴

۱/۵۹	۰/۶۳	فاصله از روستا
۲/۱۸	۰/۴۶	شیب
۱/۰۰	۱/۰۰	جهت شیب
۱/۰۶	۰/۹۴	آتش سوزی
۱/۸۱	۰/۵۵	فرسایش

آنالیز احتمال جنگل‌زدایی با استفاده از مدل جنگل تصادفی

مدل جنگل تصادفی به منظور بررسی جنگل‌زدایی در استان خوزستان مورد استفاده قرار گرفت. مدل جنگل تصادفی ۹۶ درصد از پهنه خوزستان را در رده پایین، ۲ درصد در رده متوسط و حدود ۲ درصد از مساحت استان را در رده زیاد نشان می‌دهد (جدول ۴). این روش طبقه‌بندی به صورت نقشه‌های پیش‌بینی احتمال جنگل‌زدایی در شکل ۵ ارائه شده است. مناطق شمال، شرق و غرب استان دارای بیشترین خطر جنگل‌زدایی هستند. با توجه به تمرکز جنگل‌های بلوط در قسمت‌های شمال و شمال شرق استان، محافظت از این مناطق از اهمیت ویژه‌ای برخوردار است. همچنین با توجه به اینکه مناطق غربی استان می‌توانند کانون گردوغبار و ایجاد آلودگی هوا باشند، از بین رفتن جنگل‌های این محدوده می‌تواند موجب تشدید گردوغبار و آلودگی هوا در استان شود.

آنالیز احتمال جنگل‌زدایی با استفاده از مدل بردار پشتیبان

مدل بردار پشتیبان نیز یکی از روش‌هایی است که در این پژوهش مورد استفاده قرار گرفته است. نتایج این مدل نیز در قالب نقشه احتمال جنگل‌زدایی ارائه گردیده است (شکل ۵ و جدول ۴). مطابق این مدل ۸۵ درصد پهنه استان بدون خطر، ۷ درصد در معرض خطر متوسط و ۸ درصد در معرض خطر زیاد جنگل‌زدایی قرار دارد. با توجه به اینکه ۱۵ درصد از استان خوزستان را جنگل‌ها تشکیل داده‌اند این مدل نزدیک به ۵۱ درصد از جنگل‌های استان را در معرض خطر زیاد نشان می‌دهد که همانند مدل جنگل تصادفی بیشتر در بخش‌های شمالی، شرقی و غربی استان قرار دارند. شاخص صحت ۰/۹۰۷ در این پژوهش بیانگر دقت کافی این مدل در بررسی تغییرات پوشش گیاهی است (Saha et al, 2022).

آنالیز جنگل‌زدایی با استفاده از مدل جمعی تعمیم یافته

مدل جمعی تعمیم یافته نیز در این پژوهش به طور موفقیت آمیز مورد استفاده قرار گرفت. نتایج به صورت سه کلاس خطر جنگل‌زدایی و در قالب نقشه ارائه شده است (شکل ۵ و جدول ۴). ۷۶ درصد رده پایین، ۱۸ درصد متوسط و ۶ درصد در رده زیاد قرار گرفته است. علاوه بر بخش‌های شمالی و غربی، بخش‌هایی از جنوب استان نیز در این مدل در معرض خطر جنگل‌زدایی قرار گرفته‌اند. جنگل‌های جنوبی استان بیشتر در دشت‌های مجاور رودخانه کارون قرار دارند که طی سال‌های اخیر با کم‌آبی رودخانه در معرض خطر قرار گرفته‌اند.

آنالیز احتمال جنگل‌زدایی با استفاده از درخت طبقه‌بندی و رگرسیون

نتایج پیش‌بینی جنگل‌زدایی استان خوزستان با استفاده از مدل درخت طبقه‌بندی و رگرسیون در سه کلاس پایین، متوسط و زیاد ارائه شده است (شکل ۵ و جدول ۴). ۹۱ درصد خطر کم، ۴ درصد خطر متوسط و ۵ درصد خطر زیاد را نشان می‌دهد. این مدل بیشتر مناطق غرب و جنوب استان را در معرض خطر جنگل‌زدایی نشان داده است. این مناطق بیشتر پوشش جنگلی تالاب‌ها و جنگل‌های مجاور رودخانه کارون هستند. کم‌آبی تالاب‌ها و رودخانه‌های استان طی دهه‌های اخیر تأثیر چشم‌گیری بر کاهش پوشش جنگلی به خصوص در مجاورت رودخانه‌ها و تالاب‌های استان داشته است.

آنالیز احتمال جنگل‌زدایی با استفاده از رگرسیون تقویت شده درختی

به منظور برآورد احتمال جنگل‌زدایی در استان خوزستان از مدل رگرسیون تقویت شده درختی نیز استفاده شده است و نتایج به صورت سه کلاس خطر جنگل‌زدایی ارائه شده است (شکل ۵ و جدول ۴). ۸۷ درصد خطر کم، ۹ درصد خطر متوسط و ۴ درصد خطر زیاد را نشان می‌دهند. این مدل نیز مانند جنگل تصادفی بیشتر مناطق شمالی و غربی و بخش‌های کمی از مناطق جنوبی استان را در معرض خطر جنگل‌زدایی نشان می‌دهد.

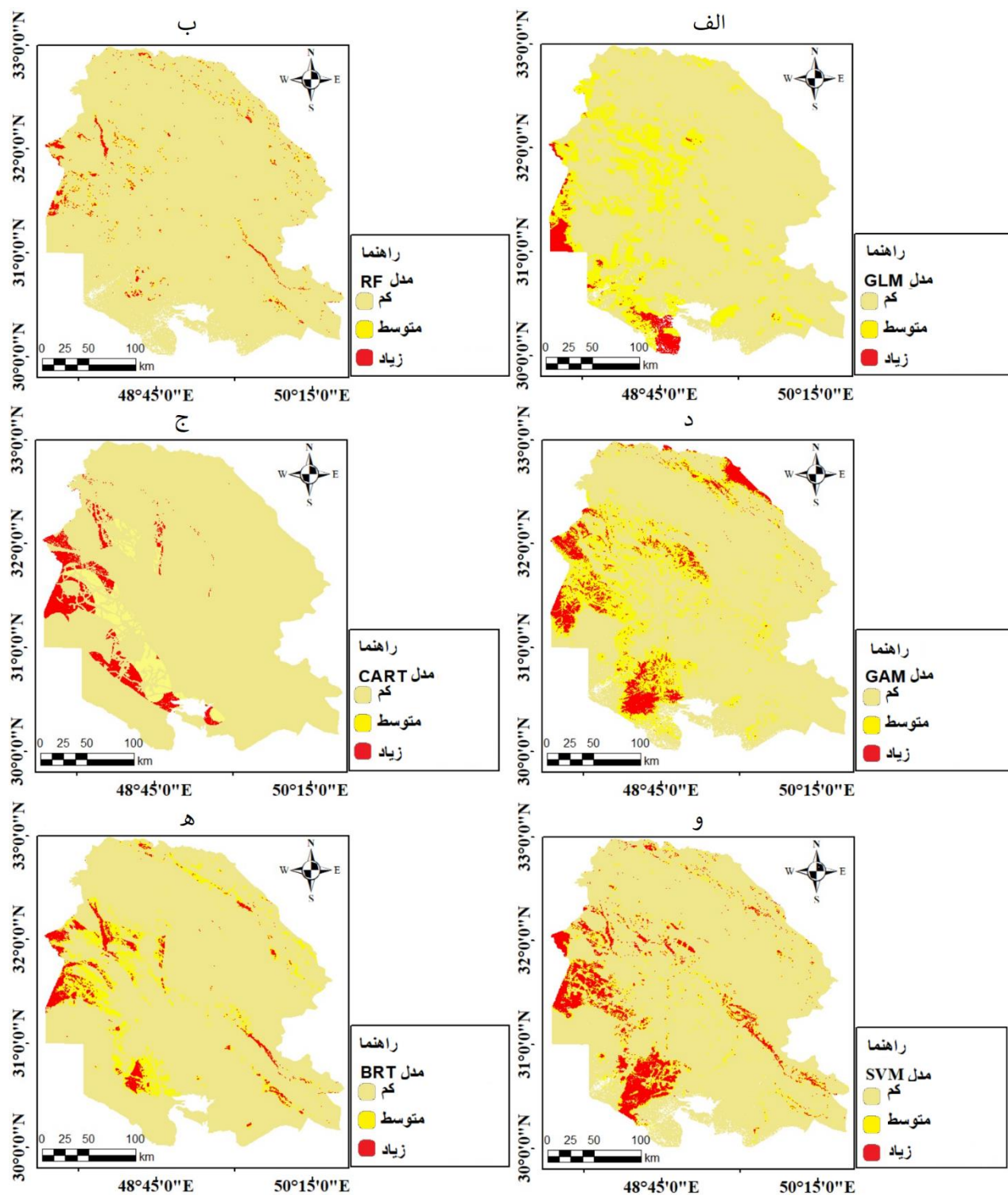
آنالیز احتمال جنگل زدایی با استفاده از مدل خطی تعمیم یافته

در این پژوهش از مدل مدل خطی تعمیم یافته نیز جهت بررسی جنگل زدایی در استان خوزستان استفاده شده است. نتایج به صورت نقشه احتمال جنگل زدایی در سه کلاس کم، متوسط و زیاد ارائه شده است (شکل ۵ و جدول ۴). ۸۱ درصد خطر کم، ۱۷ درصد خطر متوسط و ۲ درصد خطر زیاد را نشان می دهد.

جدول ۴. توزیع خطر جنگل زدایی در مدل های مختلف

کلاس	RF	GAM	GLM	SVM	BRT	CART
کم	۹۶	۷۶	۸۱	۸۵	۸۷	۹۱
متوسط	۲	۱۸	۱۷	۷	۹	۴
زیاد	۲	۶	۲	۸	۴	۵

فصل استثنای نقشه



شکل ۵. نتایج پیش‌بینی احتمال خطر با استفاده از مدل‌های مختلف خطی تعمیم یافته (الف) جنگل تصادفی (ب)، درخت طبقه بندی و رگرسیون (ج)، جمعی تعمیم یافته (د)، رگرسیون تقویت شده درختی (ه) بردار پشتیبان (و).

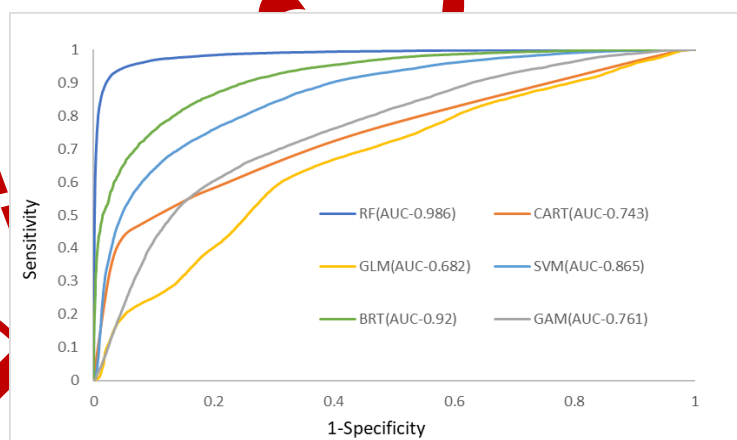
تخمین اعتبار

ارزیابی و اعتبارسنجی روش‌های یادگیری ماشین در درک بهتر مدل‌ها، بررسی علت و روند تغییرات و پیش‌بینی شرایط گیاهان در آینده به ما کمک می‌کند. اما تنها استفاده از یک روش برای اعتبارسنجی مدل کافی نیست (Saha et al, 2020). در این تحقیق روش‌های مختلف جهت اعتبارسنجی صحت مدل‌ها استفاده شده است. شش مدل مختلف یادگیری ماشین، RF، GAM، SVM،

GLM، BRT و CART با استفاده از پارامترهای مختلف از جمله نمودار ROC و AUC، Efficiency، TPR و FPR اعتبارسنجی شدند (جدول ۴) و (شکل ۶). این معیارها، توانایی پیش بینی مدل های یادگیری ماشین را نشان می دهند. مقدار AUC مدل های RF، BRT، SVM، GAM، CART و GLM به ترتیب ۰/۹۸۶، ۰/۹۲، ۰/۸۶۵، ۰/۷۶۱، ۰/۷۴۳ و ۰/۶۸۲ می باشد. نتایج نشان داد که مدل جنگل تصادفی نسبت به سایر مدل ها دارای دقت بیشتری جهت پیش بینی جنگل زدایی در مطالعه حاضر داشته است (جدول ۴). مقدار TPR مدل های RF، BRT، CART، SVM، GAM و GLM به ترتیب ۰/۹۵۳، ۰/۸۳۵، ۰/۸۲۶، ۰/۷۶۸، ۰/۷۶۸ و ۰/۶۶۸ و مقدار FPR مدل ها به ترتیب ۰/۰۴۷، ۰/۱۴۲، ۰/۲۴۰، ۰/۱۷۴، ۰/۲۳۳ و ۰/۳۰۶ برآورد شدند. این ارقام دقت کافی همه مدل های مورد استفاده در این پژوهش را جهت بررسی تغییرات پوشش جنگلی نشان می دهد (جدول ۵). با توجه به شاخص های AUC، TPR و FPR می توان نتیجه گیری کرد که همه مدل ها از دقت خوبی برخوردارند اما در مجموع مدل های جنگل تصادفی و رگرسیون تقویت شده درختی دارای دقت بیشتری نسبت به چهار مدل دیگر هستند.

جدول ۵. نتایج اعتبارسنجی مدل های یادگیری ماشین با استفاده از پارامترهای مختلف

متغیر	AUC ^۱	TPR ^۲	FPR ^۳	TNR ^۴	FNR ^۵	Efficiency	TSS
RF	۰/۹۸۶	۰/۹۵۳	۰/۰۴۷	۰/۹۵۳	۰/۰۴۷	۰/۹۵۳	۰/۹۰۶
BRT	۰/۹۲۰	۰/۸۳۵	۰/۱۴۲	۰/۸۵۸	۰/۱۶۵	۰/۸۵۱	۰/۶۹۳
SVM	۰/۸۶۵	۰/۷۶۸	۰/۱۷۴	۰/۸۲۶	۰/۲۳۲	۰/۸۰۹	۰/۵۹۴
CART	۰/۷۴۳	۰/۸۲۶	۰/۲۴۰	۰/۷۶۰	۰/۱۷۴	۰/۷۲۲	۰/۵۸۶
GAM	۰/۷۶۱	۰/۶۶۸	۰/۲۳۳	۰/۷۶۷	۰/۳۳۲	۰/۷۴۱	۰/۴۳۵
GLM	۰/۶۸۲	۰/۶۰۴	۰/۳۰۶	۰/۶۹۴	۰/۳۹۶	۰/۶۸۲	۰/۲۹۹



شکل ۶. نمودار ROC مدل های مختلف یادگیری ماشین

آنالیز حساسیت و اهمیت پارامترها

نتایج نشان می دهد که در مدل های مختلف میزان اهمیت هریک از پارامترها متفاوت است. مطابق نتایج جدول (۶)، تحلیل مدل GLM، اهمیت پارامترهای فاصله از رودخانه، فاصله از روستاها، فرسایش، بارش، شاخص رطوبت توپوگرافی، فاصله از دریاچه سد، جهت شیب، شیب، فاصله از جاده، تبخیر و آتش سوزی به ترتیب برابر با ۰/۳۶۶، ۰/۳۴۵، ۰/۳۴۳، ۰/۳۳۳، ۰/۳۳۳ و ۰/۳۳۳ است.

1. Area Under the ROC
2. True Positive Rate
3. False Positive Rate
4. True Negative Rate
5. False Negative Rate

و ۰/۳۳۲، ۰/۳۳۱، ۰/۳۳۰ و ۰/۳۲۴ نشان داد. همچنین نتایج بررسی اهمیت پارامترها با استفاده از الگوریتم RF نشان می‌دهد که اهمیت بارش، تبخیر، فاصله از دریاچه، فاصله از رودخانه، فاصله از روستا، فاصله از جاده، فرسایش، جهت شیب، شیب، آتش سوزی و شاخص رطوبت توپوگرافی به ترتیب برابر با ۰/۰۱۷، ۰/۰۱۶، ۰/۰۱۵، ۰/۰۱۵، ۰/۰۱۵، ۰/۰۱۴، ۰/۰۱۴، ۰/۰۱۴ و ۰/۰۱۳ می‌باشند. نتایج بررسی اهمیت پارامترها با استفاده از الگوریتم BRT نشان می‌دهد که اهمیت تبخیر، فاصله از دریاچه، بارش، فاصله از رودخانه، فاصله از روستا، فرسایش، آتش سوزی، فاصله از جاده جهت شیب، شاخص رطوبت توپوگرافی و شیب به ترتیب برابر با ۰/۹۳، ۰/۹۰، ۰/۸۵، ۰/۸۴، ۰/۸۲، ۰/۸۱، ۰/۸۱ و ۰/۷۸ هستند. علاوه بر این، نتایج بررسی اهمیت پارامترها با استفاده از الگوریتم CART نشان می‌دهد که اهمیت تبخیر، فاصله از روستا، بارش، فاصله از رودخانه، فرسایش، آتش سوزی، فاصله از جاده، جهت شیب، شاخص رطوبت توپوگرافی، فاصله از دریاچه و شیب به ترتیب برابر با ۰/۲۶۲، ۰/۲۵۷، ۰/۲۵۷، ۰/۲۵۷، ۰/۲۵۷، ۰/۲۵۷، ۰/۲۵۷ و ۰/۲۵۷ می‌باشند. بعلاوه نتایج بررسی اهمیت پارامترها با استفاده از الگوریتم GAM نشان می‌دهد که اهمیت فاصله از رودخانه، بارش، فاصله از دریاچه، فاصله از روستا، فاصله از جاده، شیب، فرسایش، تبخیر، جهت شیب، شاخص رطوبت توپوگرافی و آتش سوزی به ترتیب برابر با ۰/۲۷۲، ۰/۲۵۹، ۰/۲۵۴، ۰/۲۵۲، ۰/۲۴۸، ۰/۲۴۶، ۰/۲۴۵، ۰/۲۴۲، ۰/۲۴۰، ۰/۲۴۰ و ۰/۲۳۹ می‌باشند. همچنین نتایج بررسی اهمیت پارامترها با استفاده از الگوریتم SVM نشان می‌دهد که اهمیت فاصله از رودخانه، بارش، فاصله از دریاچه، فاصله از روستا، فاصله از جاده، شیب، فرسایش، تبخیر، جهت شیب، شاخص رطوبت توپوگرافی و آتش سوزی به ترتیب برابر با ۰/۱۴۷، ۰/۱۴۴، ۰/۱۴۲، ۰/۱۴۰، ۰/۱۳۹، ۰/۱۳۸، ۰/۱۳۸، ۰/۱۳۶، ۰/۱۳۵ و ۰/۱۳۵ می‌باشند. این نتایج اهمیت بالای همه پارامترهای مذکور را بر جنگل زدایی نشان می‌دهد.

پارامتر	GLM	RF	BRT	CART	GAM	SVM	Mean	Wmean
بارش	-/۰.۳۳۳	-/۰.۱۷	-/۰.۹۰	-/۰.۲۵۷	-/۰.۲۵۹	-/۰.۱۴۷	-/۰.۱۸۴	-/۰.۱۷۰
شاخص رطوبت توپوگرافی	-/۰.۳۳۳	-/۰.۱۳	-/۰.۸۰	-/۰.۲۵۷	-/۰.۲۴۰	-/۰.۱۳۵	-/۰.۱۷۶	-/۰.۱۶۲
تبخیر	-/۰.۳۳۰	-/۰.۱۷	-/۰.۹۷	-/۰.۲۷۸	-/۰.۲۴۲	-/۰.۱۳۸	-/۰.۱۸۴	-/۰.۱۷۰
آتش سوزی	-/۰.۱۲۴	-/۰.۱۴	-/۰.۸۲	-/۰.۲۵۷	-/۰.۲۳۹	-/۰.۱۳۵	-/۰.۱۷۵	-/۰.۱۶۱
فاصله از جاده	-/۰.۳۳۱	-/۰.۱۵	-/۰.۸۱	-/۰.۲۵۷	-/۰.۲۴۸	-/۰.۱۴	-/۰.۱۷۹	-/۰.۱۶۵
فاصله از دریاچه سد	-/۰.۳۳۲	-/۰.۱۶	-/۰.۹۳	-/۰.۲۴۳	-/۰.۲۵۴	-/۰.۱۴۴	-/۰.۱۸۰	-/۰.۱۶۷
فاصله از روستا	-/۰.۳۴۵	-/۰.۱۵	-/۰.۸۴	-/۰.۲۶۲	-/۰.۲۵۲	-/۰.۱۴۲	-/۰.۱۸۳	-/۰.۱۶۹
فاصله از رودخانه	-/۰.۳۶۶	-/۰.۱۵	-/۰.۸۵	-/۰.۲۵۷	-/۰.۲۷۲	-/۰.۱۵۲	-/۰.۱۹۱	-/۰.۱۷۶
فرسایش خاک	-/۰.۳۴۳	-/۰.۱۴	-/۰.۸۲	-/۰.۲۵۷	-/۰.۲۴۵	-/۰.۱۳۸	-/۰.۱۸۰	-/۰.۱۶۵
شیب	-/۰.۳۳۲	-/۰.۱۴	-/۰.۷۸	-/۰.۳۲۴	-/۰.۲۴۶	-/۰.۱۳۹	-/۰.۱۷۴	-/۰.۱۶۰
جهت شیب	-/۰.۳۳۲	-/۰.۱۴	-/۰.۸۱	-/۰.۲۵۷	-/۰.۲۴۰	-/۰.۱۳۶	-/۰.۱۷۷	-/۰.۱۶۳

به علت دقت بالای مدل‌های یادگیری ماشین در بررسی مسائل زیست محیطی و پیش بینی روابط پیچیده بین پارامترهای وابسته و پدیده‌های زیست محیطی، از این مدل‌ها (از جمله مدل خطی تعمیم یافته، شبکه عصبی مصنوعی، شبکه بیزین و جنگل تصادفی) به طور معمول در پیش‌بینی جنگل‌زدایی مورد استفاده قرار گرفتند (Fenton and Neil, 2018; Mayfield et al., 2017). در این پژوهش با استفاده از علم سنجش از دور تغییرات پوشش جنگلی استان خوزستان مورد بررسی قرار گرفته و با استفاده از یادگیری ماشین، مناطق دچار جنگل‌زدایی به عنوان یک پدیده زیست محیطی مورد تحلیل و بررسی دقیق قرار گرفت. بدین منظور از

مدل‌های یادگیری ماشین شامل جنگل تصادفی (RF)، بردار پشتیبان (SVM)، مدل جمعی تعمیم یافته (GAM)، مدل خطی تعمیم یافته (GLM)، درخت طبقه بندی و رگرسیون (CART) و رگرسیون تقویت شده درختی (BRT) استفاده شد. دقت مدل‌های مورد استفاده در پیشبینی احتمال جنگل‌زدایی با یکدیگر متفاوت است. گرچه شش مدل استفاده شده دارای دقت کافی هستند اما مقایسه مدل‌ها نشان داد که مدل‌های جنگل تصادفی و رگرسیون تقویت شده درختی نسبت به چهار مدل دیگر از دقت بیشتری برخوردار هستند. در حالی که مدل‌های جمعی تعمیم یافته و خطی تعمیم یافته نسبت به سایر مدل‌ها دقت کم‌تری را نشان دادند. در نهایت نتایج نشان دادند که در بین شش مدل مورد استفاده، مدل جنگل تصادفی (RF) دارای بیشترین دقت می‌باشد. برتری مدل جنگل تصادفی نسبت به سایر مدل‌ها در پژوهش‌های پیشین نیز ذکر شده است. Samuel (۲۰۲۴) در بررسی تغییرات جنگل‌های برزیل و Dias et al., (۲۰۲۴) در بررسی جنگل‌زدایی در منطقه آمازون برتری این مدل را نشان دادند. البته برتری این مدل در بررسی‌های صورت گرفته توسط Saha et al., (۲۰۲۲) بر روی جنگل‌های کشور هند، گزارش نشده بود. گرچه ایشان دقت کافی مدل جنگل تصادفی را گزارش دادند اما، نتایج کار آنان نشان داد که در بررسی تغییرات جنگل‌های هند مدل بردار پشتیبان دارای بیشترین دقت است. علت تفاوت نتایج این پژوهش با نتایج فعالیت‌های ساها می‌تواند تفاوت در نوع جنگل‌ها و پارامترهای مؤثر در مناطق مورد مطالعه باشد. بدیهی است روند وابستگی پارامترهای اثرگذار بر پدیده جنگل‌زدایی در مناطق مختلف جهان یکی نیست و دارای تفاوت‌های زیادی است. بنابراین لازم است در هر منطقه، مسائل زیست محیطی با استفاده از چند مدل مختلف مورد بررسی و اعتبارسنجی قرار گیرد و در پایان دقیق‌ترین مدل ملاک تصمیم‌گیری قرار گیرد.

نتایج شش مدل مورد استفاده نشان دادند که بخش چشمگیری از جنگل‌های استان خوزستان در معرض خطر جنگل‌زدایی قرار دارند و نسبت به عوامل محیطی و انسانی تأثیرپذیر و شکننده هستند. بیشترین احتمال جنگل‌زدایی در مناطق مجاور رودخانه‌ها در دشت خوزستان و نیز در سراسر کوه‌های زاگرس مشاهده شد. مدل‌های جنگل تصادفی، بردار پشتیبان، جمعی تعمیم یافته، رگرسیون تقویت شده درختی بیشتر مناطق شمال و شمال غرب که از درختان بلوط، بنه، زالزالک و بوته‌های گون و ... پوشیده شده و نیز غرب استان را که از درختان و بوته‌های تالابی پوشیده شده است، را مستعد جنگل‌زدایی نشان دادند. در حالی که دو مدل خطی تعمیم یافته و درخت طبقه بندی و رگرسیون بیشتر مناطق غربی و جنوبی استان را که انواع نخل، کنار، کهور ایرانی، گز، پده کاکله و سورمه صحرائی در مجاور تالاب و رودخانه‌ها هستند را در معرض خطر جنگل‌زدایی نشان دادند. از آنجایی که روش محاسبات توسط مدل‌های مورد استفاده با یکدیگر متفاوت است؛ بنابراین وسعت مناطق در معرض خطر در پهنه استان خوزستان در مدل‌های مختلف یکسان نیست. بررسی‌ها نشان داد که مدل جنگل‌های تصادفی نسبت به سایر مدل‌ها خطر جنگل‌زدایی را در مناطق محدودتر معرفی نمود. در حالی که مدل‌های خطی تعمیم یافته و جمعی تعمیم یافته محافظه کارانه عمل کرده و مناطق بیشتری از جنگل‌های پهنه مورد مطالعه را در معرض خطر نشان دادند. خطر جنگل‌زدایی با استفاده از شش مدل در پهنه مورد مطالعه بررسی و پیشبینی گردید. مدل‌های مورد استفاده بین ۲ تا ۸ درصد از مساحت کل استان خوزستان (۱۱ تا ۵۱ درصد مساحت جنگل‌های استان) را در معرض خطر زیاد جنگل‌زدایی نشان دادند. به طور میانگین حدود ۴ درصد از کل پهنه استان دارای احتمال بالای جنگل‌زدایی است که شامل حدود ۲۹ درصد از جنگل‌های استان می‌باشد. این آمار موید پژوهش صورت گرفته توسط عالی محمودی (۲۰۱۸) می‌باشد. نتایج کار وی نشان داد که طی سال‌های ۲۰۰۳ تا ۲۰۱۱ حدود ۱۶ درصد از جنگل‌های زاگرس از بین رفتند. با توجه به زمان بر بودن احیاء جنگل‌های زاگرس مطابق مطالعات شیرمردی و همکاران (۱۳۹۹)، اهمیت حفظ جنگل‌های زاگرس و جلوگیری از ادامه روند خطرناک جنگل‌زدایی بیش تر مشخص می‌شود. بنابراین اولویت اول برنامه ریزی در زمینه حل مسائل زیست محیطی باید جلوگیری از ایجاد و تشدید مخاطرات باشد و احیاء و جبران خسارات وارده در درجه بعدی قرار می‌گیرد.

در بررسی صورت گرفته در این پژوهش علاوه بر پارامترهای اقلیمی به بررسی عوامل تأثیرگذار دیگر از جمله تاسیسات انسان ساخت، فرسایش، پارامترهای توپوگرافی، آتش سوزی، شاخص رطوبت توپوگرافی نیز پرداخته شده است. نتایج این پژوهش نشان داد که پارامترهای مختلف بر کاهش پوشش جنگلی تأثیرگذار بودند. میزان اهمیت پارامترها در مدل‌های مورد استفاده متفاوت بوده است، به طور کلی پنج پارامتر فاصله از رودخانه، بارش، تبخیر، فاصله از دریاچه و فاصله از روستا از اهمیت بیشتری نسبت به

سایر پارامترها برخوردار هستند. می‌توان گفت که عوامل اقلیمی و انسان ساخت به طور همزمان بر تخریب جنگل‌های استان خوزستان تأثیرگذار هستند. این موضوع می‌تواند در برنامه‌ریزی‌های آینده به خصوص برنامه توسعه استان به عنوان یک اصل مهم لحاظ گردد. بیش‌ترین تغییرات پوشش جنگل‌ها در مجاورت مناطق مسکونی، جاده‌ها و در مجاورت دریاچه‌ها و رودخانه‌ها و به خصوص در مناطق دارای پوشش جنگلی متراکم تر اتفاق افتاده است. این واقعیت، تأثیر عوامل انسان ساخت را به ما نشان می‌دهد. در حالی که پارامترهای شیب، جهت شیب، شاخص رطوبت توپوگرافی و آتش سوزی نسبت به سایر پارامترها از اهمیت کم‌تری برخوردارند. آن‌چه در این پژوهش به طور قطعی قابل نتیجه‌گیری است، وجود خطر زیاد تخریب جنگل‌های جنوب غرب ایران است که در مطالعات صورت گرفته توسط پژوهشگران دیگر نیز گزارش شده بود. طبق گزارشات فائو (۲۰۲۰)، در سال ۲۰۱۴ وسعت جنگل‌های کشور ایران کاهش چشمگیری داشت. علاوه بر این نتایج این پژوهش موید گزارشات مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی کشور است که سالانه حدود ۱۲۵ هزار هکتار از جنگل‌های کشور ایران در حال تخریب هستند. البته این روند ناامیدکننده محدود به منطقه مورد مطالعه نیست، بلکه تخریب جنگل‌های سراسر جهان توسط پژوهشگران زیادی گزارش شده است. برای مثال تخریب ۹ درصد از جنگل‌های هند از سال ۱۹۹۰ تا ۲۰۱۷ توسط Saha et al., (2020) و افزایش سرعت تخریب جنگل‌های آمازون طی سال‌های ۱۹۹۰ تا ۲۰۱۸، توسط Markus et al., (2024) گزارش شده است. همچنین Buba et al., (۲۰۲۰) تخریب جنگل‌ها در منطقه نیجریه در سطح وسیع را گزارش کردند (Buba et al., 2020).

به طور کلی آنچه در این پژوهش ارائه شده است، اهمیت بالای نقش یادگیری ماشین در بررسی و حل مسئله جنگل‌زدایی با دقت بالاتر در مقایسه با روش‌های پیشین است که در مطالعات صورت گرفته توسط دیگر پژوهشگران از جمله Mayfield et al., (۲۰۱۷) به خوبی بیان شده بود. به طور کلی مطالعات صورت گرفته توسط دیگر پژوهشگران عمدتاً در مناطق غیر از محدوده مورد مطالعه در این پژوهش بوده است. این پژوهش با به کارگیری الگوریتم‌های مختلف یادگیری ماشین به بررسی و تحلیل عوامل مؤثر بر جنگل‌زدایی در منطقه جنوب غرب ایران با دقت بالا پرداخته و اهمیت هر پارامتر را در تشدید جنگل‌زدایی تعیین کرده است. همچنین محدوده‌های از این منطقه که دارای خطر جنگل‌زدایی هستند، نیز در این پژوهش معین شده‌اند.

نتیجه‌گیری

در این پژوهش با استفاده از علم سنجش از دور تغییرات پوشش جنگلی استان خوزستان مورد بررسی قرار گرفته و مناطق دچار جنگل‌زدایی به عنوان یک پدیده زیست محیطی مورد تحلیل و بررسی دقیق قرار گرفت. بدین منظور از مدل‌های یادگیری ماشین شامل جنگل تصادفی (RF)، بردار پشتیبان (SVM)، مدل جمعی تعمیم یافته (GAM)، مدل خطی تعمیم یافته (GLM)، درخت طبقه بندی و رگرسیون (CART) و رگرسیون تقویت شده درختی (BRT) استفاده شد. دقت مدل‌های مورد استفاده در پیش‌بینی احتمال جنگل‌زدایی با یکدیگر متفاوت است. گرچه شش مدل استفاده شده دارای دقت کافی هستند اما مقایسه مدل‌ها نشان می‌دهد که مدل‌های جنگل تصادفی و رگرسیون تقویت شده درختی نسبت به چهار مدل دیگر از دقت بیشتری برخوردار هستند. در حالی که مدل‌های جمعی تعمیم یافته و خطی تعمیم یافته نسبت به سایر مدل‌ها دقت کم‌تری را نشان دادند. در نهایت نتایج نشان دادند که در بین شش مدل مورد استفاده، مدل جنگل تصادفی (RF) دارای بیشترین دقت می‌باشد. از آنجایی که روش محاسبات توسط مدل‌های مورد استفاده با یکدیگر متفاوت است؛ بنابراین وسعت مناطق در معرض خطر در پهنه استان خوزستان در مدل‌های مختلف یکسان نیست. بررسی‌ها نشان می‌دهد که مدل جنگل‌های تصادفی نسبت به سایر مدل‌ها خطر جنگل‌زدایی را در مناطق محدودتر معرفی می‌کند. در حالی که مدل‌های مدل بردار پشتیبان و جمعی تعمیم یافته محافظه کارانه عمل کرده و مناطق بیشتری از جنگل‌های پهنه مورد مطالعه را در معرض خطر نشان می‌دهند. خطر جنگل‌زدایی با استفاده از شش مدل در پهنه مورد مطالعه بررسی و پیش‌بینی گردید. مدل‌های مورد استفاده بین ۲ تا ۸ درصد از مساحت کل استان خوزستان را در معرض خطر زیاد جنگل‌زدایی نشان می‌دهند. با توجه به اینکه ۱۵ درصد مساحت این استان از جنگل پوشیده شده است، این رقم معادل ۱۱ تا ۵۱ درصد جنگل‌های استان خوزستان است. به عبارت دیگر، به طور میانگین حدود ۴ درصد از کل پهنه استان یعنی حدود یک سوم جنگل‌های استان در معرض خطر بالای جنگل‌زدایی قرار دارند. بنابراین اهمیت حفظ و جلوگیری از تخریب و احیاء این جنگل‌ها با روش‌های ممکن آشکار شده است.

نتایج این پژوهش نشان می‌دهد که همه پارامترهای مختلف مورد استفاده در این پژوهش بر کاهش پوشش جنگلی تأثیرگذار بودند. میزان اهمیت پارامترها در مدل‌های مورد استفاده متفاوت بوده است، به طور کلی پنج پارامتر فاصله از رودخانه، بارش، تبخیر، فاصله از دریاچه و فاصله از روستا از اهمیت بیشتری نسبت به سایر پارامترها برخوردار هستند. می‌توان گفت که عوامل اقلیمی و انسان ساخت به طور همزمان بر تخریب جنگل‌های استان خوزستان تأثیرگذار هستند. این موضوع می‌تواند در برنامه ریزی‌های آینده به خصوص برنامه توسعه استان به عنوان یک اصل مهم لحاظ گردد. بیشترین تغییرات پوشش جنگل‌ها در مجاورت مناطق مسکونی، جاده‌ها و در مجاورت دریاچه‌ها و رودخانه‌ها و به خصوص در مناطق دارای پوشش جنگلی متراکم تر اتفاق افتاده است. این واقعیت، تأثیر عوامل انسان ساخت را به ما نشان می‌دهد. در حالی که پارامترهای شیب، جهت شیب، شاخص رطوبت توپوگرافی و آتش سوزی نسبت به سایر پارامترها از اهمیت کم تری برخوردارند.

"هیچگونه تعارض منافع بین نویسندگان وجود ندارد"

منابع

ذرتی پور، مین (۱۴۰۱). بررسی تبخیر و تعرق واقعی و توزیع مکانی نیاز آبی محصولات استراتژیک کشاورزی با استفاده از الگوریتم های سنجنش از دور و مقایسه با مدل های پایه فیزیکی تبخیر و تعرق. گزارش طرح ملی (سازمان برنامه و بودجه کشور)، شماره طرح: ۵۰۳۳۷۴، ۱۴۸ ص.

رحیمی‌خوب، علی (۱۳۹۴). مقایسه بین مدل درختی M5 و شبکه‌های عصبی برای تخمین تبخیر و تعرق مرجع در یک محیط خشک. مجله مدیریت منابع آب، ۲۸ (۲)، ۶۵۷-۶۶۹.

شعبانی، سعید، پورقاسمی، حمید، رضا، و بلاشک، تارا (۱۴۰۰). نقشه‌برداری حساسیت توده جنگلی در طول برداشت با استفاده از مدل های رگرسیون لجستیک و درخت رگرسیون تقویت شده یادگیری ماشین. مجله بوم‌شناسی، ۲۲ (۳)، e00974.

شریفانی، کریم، و امینی، مریم (۱۴۰۲). یادگیری ماشین و یادگیری عمیق: مروری بر روش‌ها و کاربردها. مجله جهانی فناوری اطلاعات و مهندسی، ۱۰ (۰۷)، ۳۸۹۷-۳۹۰۴.

شیرمردی، حمزه، علی، حیدری، قدرت الله، قربانی، جمشید، طهماسبی، پژمان، و محنتکش، عبدالمحمد (۱۳۹۹). تغییرات شاخص‌های پوشش گیاهی در اراضی زراعی با سال‌های مختلف رهاسازی در مراتع شیرمرد، استان چهارمحال و بختیاری. مجله حفاظت از اکوسیستم‌های گیاهی، ۶ (۱۳)، ۱۷۷-۱۹۶.

عالی‌محمودی سراب، سجاد (۱۳۹۷). پایش عوامل گرد و غبار و آب و هوایی مؤثر بر زوال بلوط ایرانی (Quercus branti) با استفاده از تصاویر ماهواره‌ای در جنگل‌های زاگرس جنوبی. رساله دکتری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، ایران. ۱۸۰ ص.

کشتکار، محمد. رضا (۱۴۰۳). بررسی و تحلیل تغییرات پوشش زمین در ایران طی سال‌های ۱۹۶۳ تا ۲۰۱۵ در مطالعات سند ملی آمایش سرزمین، سند گزارش پژوهشی، مرکز تحقیقات توسعه و آینده‌نگری سازمان برنامه‌ریزی و بودجه کشور، ۲۲۰ صفحه.

REFERENCES

- Aertsen, W., Kint, V., Van Orshoven, J., Özkan, K., & Muys, B. (2010). Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecological modelling*, 221(8), 1119-1130.
- Aksoy, S., Sertel, E., Roscher, R., Tanik, A., & Hamzehpour, N. (2024). Assessment of soil salinity using explainable machine learning methods and Landsat 8 images. *International Journal of Applied Earth Observation and Geoinformation*, 130, 103879.
- Alimahmoudi Sarab, S. (2018). *Monitoring dust and weather factors effecting on Quercus branti decline using satellite images in Southern Zagroos forests* (Doctoral dissertation, Gorgan University of Agricultural Sciences and Natural Resources, Iran). (In Persian).
- Bera, B., Saha, S., & Bhattacharjee, S. (2020). Forest cover dynamics (1998 to 2019) and prediction of deforestation probability using binary logistic regression (BLR) model of Silabati watershed, India. *Trees, Forests and People*, 2, 100034.

- Buba, F. N., Gajere, E. N., & Ngum, F. F. (2020). Assessing the Correlation between Forest Degradation and Climate Variability in the Oluwa Forest Reserve, Ondo State, Nigeria. *American Journal of Climate Change*, 9(4), 371-390.
- Centre for People and Forests—CPF (2012). Climate Change, Forests, and You. Gras- roots Capacity Building for REDD+ in the Asia-Pacific Region.
- Dias, F., Suhadolnik, N., Camargo, H., & Da Silva, S. (2024). Predicting the pulse of the Amazon: Machine learning insights into deforestation dynamics. *Journal of Environmental Management*, 362, 121359.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of animal ecology*, 77(4), 802-813.
- Fang, K., Shen, C., Kifer, D., & Yang, X. (2017). Prolongation of SMAP to spatiotemporally seamless coverage of continental US using a deep learning neural network. *Geophysical Research Letters*, 44(21), 11-030.
- FAO and UNEP (2020). The State of the World's Forests 2020. *Forests, Biodiversity and People*. FAO and UNEP, Rome. <https://doi.org/10.4060/ca8642en>.
- Fenton, N., & Neil, M. (2018). *Risk assessment and decision analysis with Bayesian networks*. Crc Press.
- Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of remote sensing*, 23(4), 725-749.
- Jabeen, S., Li, X., Amin, M. S., Bourahla, O., Li, S., & Jabbar, A. (2023). A review on methods and applications in multimodal deep learning. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(2s), 1-41.
- Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality & quantity*, 52, 1957-1976.
- Kang, J., Schwartz, R., Flickinger, J., & Beriwal, S. (2015). Machine learning approaches for predicting radiation therapy outcomes: a clinician's perspective. *International Journal of Radiation Oncology* Biology* Physics*, 93(5), 1127-1135.
- Keshtkar, M. (2023). Survey and analysis of changes in land cover in Iran during the years 1933 to 2015 in studies of the National Land Planning Document, *research report document*, Center for Development and Foresight Research of the National Planning and Budget Organization, 220 pp. (In Persian).
- Kumar, B. P., Babu, K. R., Anusha, B. N., & Rajasekhar, M. (2022). Geo-environmental monitoring and assessment of land degradation and desertification in the semi-arid regions using Landsat 8 OLI/TIRS, LST, and NDVI approach. *Environmental Challenges*, 8, 100578.
- Kumar, R., Nandy, S., Agarwal, R., & Kushwaha, S. P. S. (2014). Forest cover dynamics analysis and prediction modeling using logistic regression model. *Ecological indicators*, 45, 444-455.
- Maione, C., & Barbosa, R. M. (2019). Recent applications of multivariate data analysis methods in the authentication of rice and the most analyzed parameters: A review. *Critical reviews in food science and nutrition*, 59(12), 1868-1879.
- Marcus, M. S., Herguac'h, K., Coronado, E. N. H., & Gutiérrez-Vélez, V. H. (2024). Spatial distribution of degradation and deforestation of palm swamp peatlands and associated carbon emissions in the Peruvian Amazon. *Journal of Environmental Management*, 351, 119665.
- Matthew et al., 2024, Spatial distribution of degradation and deforestation of palm swamp peatlands and associated carbon emissions in the Peruvian Amazon, *Journal of Environmental Management* 351 (2024) 119665.
- Mayfield, H., Smith, C., Gallagher, M., & Hockings, M. (2017). Use of freely available datasets and machine learning methods in predicting deforestation. *Environmental modelling & software*, 87, 17-28.
- Moscovini, L., Ortenzi, L., Pallottino, F., Figorilli, S., Violino, S., Pane, C., ... & Costa, C. (2024). An open-source machine-learning application for predicting pixel-to-pixel NDVI regression from RGB calibrated images. *Computers and Electronics in Agriculture*, 216, 108536.
- Pradhan, B. (2013). A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Computers &*

- Geosciences*, 51, 350-365.
- Rahimikhoob, A. (2014). Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water resources management*, 28, 657-669. (In Persian).
- Saha, S., Bhattacharjee, S., Shit, P. K., Sengupta, N., & Bera, B. (2022). Deforestation probability assessment using integrated machine learning algorithms of Eastern Himalayan foothills (India). *Resources, Conservation & Recycling Advances*, 14, 200077.
- Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P. T. T., & Paul, G. C. (2020). Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, REPTree: A case study at the Gumani River Basin, India. *Science of the Total Environment*, 730, 139197.
- Samuel Jos 2024, Machine learning methods: Modeling net growth in the Atlantic Forest of Brazil, *Ecological Informatics*, 81(3):102564 81(3):102564.
- Senior, R. A., Hill, J. K., González del Pliego, P., Goode, L. K., & Edwards, D. P. (2017). A pantropical analysis of the impacts of forest degradation and conversion on local temperature. *Ecology and evolution*, 7(19), 7897-7908.
- Shabani, S., Pourghasemi, H. R., & Blaschke, T. (2020). Forest stand susceptibility mapping during harvesting using logistic regression and boosted regression tree machine learning models. *Global Ecology and Conservation*, 22, e00974. (In Persian).
- Sharifani, K., & Amini, M. (2023). Machine learning and deep learning: A review of methods and applications. *World Information Technology and Engineering Journal*, 10(07), 3897-3904.
- Shirmardi, H. A., Heydari, G., Ghorbani, J., Tahmasbi, P., & Mehnatkesh, A. (2019). Changes of vegetation indices in arable lands with different years of abandonment in Shirmard rangelands, Chaharmahal va Bakhtiari province. *Journal of Plant Ecosystem Conservation*, 6(13), 177-196. (In Persian).
- Takahashi, K., Kim, K., Ogata, T., & Sugano, S. (2017). Tool-body assimilation model considering grasping motion through deep learning. *Robotics and Autonomous Systems*, 91, 115-127.
- Tovar, C. L. M. (2009). Analysis of the normalized differential vegetation index (NDVI) for the detection of degradation of forest coverage in Mexico 2008–2009.
- Tuoku, L., Wu, Z., & Men, B. (2024). Impacts of climate factors and human activities on NDVI change in China. *Ecological Informatics*, 81, 102555.
- Vapnik, V.N., (1995). The Nature of Statistical Learning Theory. *Springer Science & Business Media*, Berlin, Germany.
- Wang, Z., Wang, Y., Liu, Y., Wang, F., Deng, W., & Rao, P. (2023). Spatiotemporal characteristics and natural forces of grassland NDVI changes in Qilian Mountains from a sub-basin perspective. *Ecological Indicators*, 157, 111186.
- Zhang, B., He, X., Ouyang, F., Gu, D., Dong, Y., Zhang, L., ... & Zhang, S. (2017). Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. *Cancer letters*, 403, 21-27.
- Zhao, J., Li, L., Liu, J., Yan, Y., Wang, Q., Newman, C., & Zhou, Y. (2024). A bibliometric analysis using machine learning to track paradigm shifts and analytical advances in forest ecology and forestry journal publications from 2010 to 2022. *Forest Ecosystems*, 100233.
- Zhou, C., Lin, K., Xu, D., Chen, L., Guo, Q., Sun, C., & Yang, X. (2018). Near infrared computer vision and neuro-fuzzy model-based feeding decision system for fish in aquaculture. *Computers and electronics in agriculture*, 146, 114-124.
- Zoratipour, A. (2021). Investigation of actual evapotranspiration and spatial distribution of water requirement of strategic agricultural crops using remote sensing algorithms and comparison with physical-based evapotranspiration models. *National Project Report (National Planning and Budget Organization, 2021)*, Project No.: 502374, 148p. (In Persian).